ᲡᲐᲥᲐᲠᲗᲕᲔᲚᲝᲡ ᲛᲔᲪᲜᲘᲔᲠᲔᲑᲐᲗᲐ ᲔᲠᲝᲕᲜᲣᲚᲘ ᲐᲙᲐᲓᲔᲛᲘᲘᲡ ᲛᲝᲐᲛᲑᲔ, ტ. 175, №4, 2007
BULLETIN OF THE GEORGIAN NATIONAL ACADEMY OF SCIENCES, vol. 175, no. 4, 2007

Mathematics

Selected Papers of the Symposium dedicated to the 80th Birthday of Academician Revaz Gamkrelidze (Batumi, 17-21 September, 2007)

Optimal Control of Measures

Andrei Agrachev

SISSA-ISAS, Trieste, Italy and MIAN, Moscow, Russia

ABSTRACT. We study a randomization of the standard finite dimensional optimal control problem: we just assume that boundary values of the trajectory are not fixed but have some probability distributions and try to minimize the expectation of the cost. This is actually a control version of the "optimal mass transportation". We are busy with the existence, uniqueness and characterization of minimizers. The paper is dedicated to the 80th anniversary of Revaz Valerianovich Gamkrelidze and is based on a joint work with Paul Lee. *© 2007 Bull. Georg. Natl. Acad. Sci.*

Key words: *measure, control theory, mass transportation problem.*

1. Introduction

The paper is dedicated to the 80th anniversary of my teacher Revaz Valerianovich Gamkrelidze, one of the founders of the Optimal Control Theory. Professor Gamkrelidze is the author of many bright ideas in Control Theory, one of them is the idea of generalized controls which give a proper relaxation of the optimal control problem (see book [1] and references therein). According to Gamkrelidze, generalized controls depend on time probability measures on the space of control parameters. Introduction of the generalized controls makes the space of admissible controls convex and the right-hand side of the control system depends linearly on this admissible control. Moreover, very strong approximation results demonstrate that we do not lose any information with such a relaxation. Yet the system remains nonlinear with respect to the state variable.

Later Revaz Valerianovich and myself developed the Chronological Calculus which systematically treats nonlinear (with respect to the state variable) systems as linear operator equations. Yet the state space remains nonlinear in this calculus.

The next step which we perform in this note is the *randomization* of the state variable: a "pure" state is substituted by a probability distribution. The randomization makes the state space convex and somehow regularizes the optimal control problem. If the relaxation of controls provides the existence of the optimal solution, the randomization of the state leads to its uniqueness. Actually, the randomized optimal control problem belongs to the class of so-called "mass transportation problems" (see [2]); this fact allows to use such a powerful analytic tool as Kantorovich duality in order to achieve the uniqueness and give an appropriate description of the optimal solution.

We start from a semi-heuristic calculation in the space of volume forms in order to demonstrate the flavor of the problem and to give an idea of what kind of solution we could expect. Then we formulate rigorous results whose proofs can be found in [3].

2. Optimal Displacement

Let M be a smooth manifold and U be a closed subset of another manifold. We consider a standard optimal control problem with the integral cost and fixed endpoint:

$$\min \int_0^1 \varphi_{u)t)}(q(t))dt \,,\ \dot{q}(t) = f_{u(t)}(q(t)) \,,\ q(0) = q_0 \,,\ q(1) = q_1 \,. \tag{1}$$

where $q \in M\,;\ u \in U$, and the Lagrangian $(q,u) \mapsto \varphi_u(q)$ is smooth as well as the mapping $(q,u) \mapsto f_u(q)$, where $f_u(q) \in T_q M \subset TM$. This problem induces the following "optimal displacement" problem on the space of volume forms (i. e. positive measures with smooth densities). The state space for this new problem is the space of volume forms, although, eventually, it can be extended to a more general class of nonnegative measures. Control parameters are mappings $u : M \to U$ (say, Lipschitz mappings); we will use the symbol f_u for the vector field $q \mapsto f_{u(q)}(q)$ Control functions are measurable bounded with respect to t and Lipschitz with respect to q mappings $(t,q) \mapsto u_t(q)$, $t \in [0,1]$, $q \in M$. The Lagrangian of the optimal displacement problem is:

$$l(\mu,u) = \int_M \varphi_u \mu$$

and the control system:

$$\dot{\mu}_t = -L_{f_{u_t}} \mu_t \tag{2}$$

where L_{f_u} is the Lie derivative. Then the measure μ_t is the pushforward of μ_0 by the $F_{u_t}^t = \overrightarrow{\exp} \int_0^t f_{u_\tau} d\tau$ on M

generated by the differential equation $\dot{q} = f_{u_t}(q)$, i. e. $\mu_t = F_{u_t *}^t(\mu_0)$. The problem is to minimize

$$\int_0^1 l(\mu_t, u_t)dt = \int_M \left(\int_0^1 \varphi_{u_t} \circ F_{u_t}^t dt \right) \mu_0$$

for all (μ_t, u_t) such that (2) is satisfied and the "endpoints" μ_0 , μ_1 are fixed. We would like to follow the Hamiltonian approach of the Pontryagin Maximum Principle in order to characterize possible solutions to the minimization problem (see [1], [4] for the original coordinate treatment and [5], [6] for the intrinsic presentation). The dual space to the space of measures is the space of scalar functions, so that the "cotangent bundle" to our state space is the direct sum of the space of measures $\{\mu\}$ and the space of functions $\{p\}$. The depending on control "Hamiltonian of the optimal control problem":

$$b_u(p,\mu) = -< p, L_{f_u}\mu > -\ell(\mu,u) = -\int pL_{f_u}\mu - \int \varphi_u \mu \,.$$

We integrate by parts and obtain:

$$b_u(p,\mu) = \int (< dp, f_u > -\varphi_u)\mu = \int h_u(dp)\mu$$

where $h_u(z) =< z, f_u(q) > -\varphi_u(q)$, $z \in T_q^* M$, $q \in M$. Hence

$$B(p,\mu) \overset{def}{=} \max_u b_u(p,\mu) = \int H(dp)\mu$$

where

$$H(z) \overset{def}{=} \max_{u \in U} h_u(z) \,,\ z \in T^* M \,, \tag{3}$$

is the maximized Hamiltonian of the standard optimal control problem (1).

Now find the "vertical derivative" $\frac{\partial B}{\partial p}$. This is easy:

$$<\frac{\partial B}{\partial p}, p'> = \int <dp', d^v_{dp}H> \mu = -\int p' L_{(d^v_{dp}H)}\mu \, .$$

So, we have:

$$\frac{\partial B}{\partial p}(p,\mu) = -L_{(d^v_{dp}H)}\mu \, , \; \frac{\partial B}{\partial \mu} = H(dp)$$

We see that the "Hamiltonian system" for B is actually reduced to the Hamilton-Jacobi equation

$$\dot{p} = -H(dp) \tag{4}$$

plus the transport equation:

$$\dot{\mu} = -L_{(d^v_{dp}H)}\mu$$

a quite natural and predictable result. Now we have good candidates for the optimal displacement: the measure should be transformed by the flow on M defined by the characteristic curves associated to the Cauchy problem for the Hamilton-Jacobi equation (4). Recall that characteristic curves are projections to M of solutions to the Hamiltonian system on T^*M with the Hamiltonian function H.

3. Existence and Uniqueness of Optimal Map

Here we formulate the main existence and uniqueness result. The proof is essentially analytic and uses Kantorovich duality (see [3] and [2]); anyway, I hope that simple calculation of the previous section explains why the result is natural.

In fact, we deal with a more general optimal displacement problem than one considered in the previous section. Now μ_0 and μ_1 are arbitrary Borel probability measures on M and admissible controls are Borel maps $u : M \to L^\infty([0,1], U)$. Let $F^t_{u(q)} = \overrightarrow{\exp}\int_0^t f_{u(q)(\tau)} d\tau$ the problem is to minimize

$$F^t_{u(q)} = \int_M \left(\int_0^1 \varphi_{u(q)(t)} F_{u(q)}(q)\, dt \right) d\mu_0(q)$$

for all admissible controls u such that the "endpoints" μ_0 and $\mu_1 = F^1_{u*}(\mu_0)$ are fixed.

We set

$$c(x,y) = \min\{\int_0^1 \varphi_{u(t)}(q(t))dt, : \dot{q}(t) = f_{u(t)}(q(t)), q(0) = x, q(1) = y\}$$

the optimal cost for the standard finite dimensional problem (1). Function c is defined on the subset of $M \times M$ formed by the endpoints for which there exists a minimum. We denote by $\vec{H}$ the Hamiltonian vector field of the function H (see (3)) and by $e^{t\vec{H}}$ the flow on T^*M generated by this vector field if H is smooth and $\vec{H}$ is complete. Finally, let $\pi : T^*M \to M$ be the standard projection.

Theorem 3.1. *Assume that Hamiltonian H is of class C^2, the field $\vec{H}$ is complete, measures μ_0, μ_1 have compact supports and the cost function c is well-defined and Lipschitz in a neighborhood of* $\sup p(\mu_0) \times \sup p(\mu_1)$. If *$\mu_0$ is absolutely continuous (w. r. t. the Lebesgue measure) then there exists a unique up to μ_0-measure zero optimal displacement F^t_u. Moreover, there exists a Lipschitz function a on M such that*

$$F^t_u = \pi\left(e^{t\vec{H}}(d_q a)\right), \; 0 \le t \le 1,$$

for μ_0-almost all $q \in M$.

4. Regularity of Control Cost

In Theorem 3.1, we prove the existence and uniqueness of optimal maps under certain regularity conditions on the cost. In this section, we give simple conditions which guarantee this regularity. Here we consider only affine with respect to control systems. In other words, $f_u(q) = X_0(q) + \sum_{i=1}^{k} u_i X_i(q)$, where $X_0, X_1, \ldots, X_k$ are fixed smooth vector fields on the manifold M, $u = (u_1, \ldots, u_k)$ and $U = \mathbf{R}^k$ (see [6]).

The Cauchy problem for system

$$f_u(q) = X_0(q(t)) + \sum_{i=1}^{k} u_i(t) X_i(q(t)) \tag{5}$$

is correctly stated for any locally integrable vector-function $u(\cdot)$ and we assume, throughout this section, that system (5) is complete, i. e. all solutions of the system are defined on the whole semi-axis $[0, \infty)$.

This completeness assumption is automatically satisfied if M is a compact manifold or M is a Lie group and the fields X_i are left-invariant or if M is a closed submanifold of the Euclidean space and $|X_i(q)| \leq c(1+|q|)$, $i = 0,1,\ldots,k$.

We need some basic notions of the geometric control theory, see [5] for detail. Fix a point q_0 in the manifold M. Consider the endpoint map $End_{q_0} : L^1([0,1], \mathbf{R}^k) \to M$ defined by $End_{q_0}(u(\cdot)) = q(1)$, where $q(\cdot)$ is the admissible path corresponding to the control $u(\cdot)$ and initial condition $q(0) = q_0$. It is known that the map End_{q_0} is a smooth mapping. The critical points of the map End_q are called singular controls. Admissible paths corresponding to the singular controls are called singular trajectories.

We also need the Hessian of the mapping End_{q_0} at the critical point. Let E be a Banach space which is an everywhere dense subspace of a Hilbert space H. Consider a mapping $\Phi : E \to \mathbf{R}^n$ such that the restriction of this map $\Phi|_W$ to any finite dimensional subspace W of the Banach space E is C_2. Moreover, we assume that the first and second derivatives of all the restrictions $\Phi|_W$ are continuous in the Hilbert space topology on the bounded subsets of E. In other words,

$$\Phi(v+w) - \Phi(v) = D_v\Phi(w) + \frac{1}{2} D_v^2\Phi(w) + o(|w|^2), \quad w \in W$$

where $D_v\Phi$ is a linear and $D_v^2\Phi$ quadratic mappings from E to $\mathbf{R}^n$. Moreover, $D_v\Phi|_W$ and $D_v^2\Phi|_W$ continuously depend on v in the topology of H while v is contained in a ball of E.

The Hessian $Hess_v\Phi : \ker D_v\Phi \to \operatorname{coker} D_v$ of the function Φ is the restriction of $D_v^2\Phi$ to the kernel of $D_v\Phi$ with values considered up to the image of $D_v\Phi$. Hessian is a part of $D_v^2\Phi$ which survives smooth changes of variables in E and $\mathbf{R}^n$. Let ξ be a covector in the dual space $\mathbf{R}^{n^*}$ such that $\xi D_v\Phi = 0$, then $\xi Hess_v\Phi$ is a well-defined real quadratic form on $\ker D_v\Phi$. We denote the Morse index of this quadratic form by $ind(\xi Hess_v\Phi)$. Recall that the Morse index of a quadratic form is the supremum of dimensions of the subspaces where the form is negative definite.

Definition 4.1. A critical point v of Φ is called *sharp* if there exists a covector $\xi \neq 0$ such that $\xi D_v\Phi = 0$ and $ind(\xi Hess_v\Phi) < +\infty$.

Needless to say, the spaces E, H and $\mathbf{R}^n$ can be substituted by smooth manifolds (Banach, Hilbert and n-dimensional) in all this terminology.

Going back to the control system (5), let $(u(\cdot), q(\cdot))$ be an admissible pair for this system. We say that the control $u(\cdot)$ and the path $q(\cdot)$ are sharp if $u(\cdot)$ is a sharp critical point of the mapping $End_{q(0)}$.

One necessary condition for control and path to be sharp is the so-called Goh condition.

Proposition 4.2 (Goh condition). *If* $\xi Hess_{u(\cdot)} End_{q(0)} < +\infty$, *then*

$$< \xi(t), X_i(q(t)) >=< \xi(t), [X_i, X_j](q(t)) >= 0, \ i, j = 1,...,k, \ 0 \le t \le 1,$$

where $\xi(t) = P_{t,1}^{*}\xi$ and $P_{t,1} = \overrightarrow{\exp}\int_t^1 f_{u(\tau)} d\tau$.

See [5, 8] and references therein for the proof and other effective necessary and sufficient conditions of sharpness.

Now turn to the optimal control problem. We assume that the Lagrangian $\varphi_u(q)$ is strictly positive, strongly convex with respect to *u* and satisfies standard growth conditions, i.e. for any compact $K \subset M$ there exist constants *a, b* such that $| d_q \varphi | \le a(\varphi_u(q) + | u |) + b \quad \forall u \in \mathbf{R}^k, \ q \in K$ and $\frac{| u |}{\varphi_u(q)}$ tends to zero as $| u |$ tends to zero, uniformly on K. These properties guarantee smoothness of the Hamiltonian

$$H(\xi, q) = \max_{u \in R^k} \left(< \xi, X_0(q) + \sum u_i X_i(q) > - \varphi_u(q) \right).$$

We are now ready to state the main result of of this section.

Theorem 4.3. *Assume that the system (5) does not admit sharp controls, then the set*

$$D = \{(x, End_x(u(\cdot))) \,|\, x \in M, u \in L^{\infty}([0,1], R^k)\}$$

open in the product $M \times M$. *Moreover, the optimal* $\cos t\,(x, y) \mapsto c(x, y)$ *is locally Lipschitz on the set D.*

5. Applications: Optimal Displacement on Sub-Riemannian Manifolds

In this section, we will apply the results in the previous sections to some sub-Riemannian manifolds. First, let us recall some basic definitions.

Let Δ and Δ' be two (possibly singular) distributions on a manifold *M*. Define the distribution $[\Delta, \Delta']$ by

$$[\Delta, \Delta'] = span\{[v, w] \,|\, v \text{ is a section of } \Delta, w \text{ is a section of } \Delta'\}$$

Define inductively the following distributions: $[\Delta, \Delta'] = \Delta^2$ and $\Delta^k = [\Delta, \Delta^{k-1}]$. A distribution is called *k*-generating if $\Delta^k = TM$ and the smallest such *k* is called the degree of nonholonomy. Also, the distribution is called bracket generating if it is *k*-generating for some *k*.

If * is a bracket generating distribution, then it defines a flag of distribution by

$$\Delta \subset \Delta^2 \subset ... \subset TM.$$

The growth vector of the distribution Δ at the point *x* is defined by $(\dim \Delta_x, \dim \Delta_x^2, ..., \dim T_x M)$. Let $x(\cdot) : [a, b] \to M$ be an admissible curve, that is a Lipschitz curve almost everywhere tangent to Δ. The following classical result on bracket generating distributions is the starting point of sub-Riemannian geometry.

Theorem 5.1. (Rashevskii and Chow) *Given any two points x and y on the manifold M with a bracket generating distribution, there exists an admissible curve joining the two points.*

Using Rashevskii-Chow Theorem, we can define the sub-Riemannian distance *d*. Let <,> be a fibre inner product on the distribution Δ, called sub-Riemannian metric. The length of an admissible curve $q(\cdot)$ is defined in the usual way: $length(q(\cdot)) = \int_a^b \sqrt{< \dot{q}(t), \dot{q}(t) >} dt$. The sub-Riemannian distance *d*(*x; y*) between two points *x* and *y* is defined by the infimum of the length of all admissible curves joining *x* and *y*. There is a quantitative version of Chow-Rashevskii Theorem, called Ball-Box Theorem, which gives Hölder continuity of the sub-Riemannian distance. See [9] for details.

Corollary 5.2. *Let d be the metric of a complete sub-Riemannian space with distribution* Δ. *Function* d^2 *is locally Lipschitz if and only if the distribution is 2-generating.*

Proof. Indeed, d^2 is the optimal cost of the optimal control problem for a linear with respect to control system and quadratic with respect to control Lagrangian; this is a special case of the problem considered in the previous section. The systems with 2-generating distributions do not admit sharp paths because these systems are not compatible with the Goh condition. On the other hand, constant paths (points) are sharp minimizers in the case of distributions whose nonholonomy degree is greater than 2 and the Ball-Box Theorem implies that d^2 is not locally Lipschitz at the diagonal in this case.

The locally Lipschitz property of the distance d out of the diagonal is guaranteed for a much bigger class of distributions. In particular, it is proved in [10] that a generic distribution of $rank > 2$ does not admit nonconstant sharp trajectories. In the class of Carnot groups, the following estimates are valid: generic n-dimensional Carnot group with rank k distribution does not admit nonconstant sharp trajectories if $n \le (k-1)k+1$ and has nonconstant sharp length minimizing trajectories if $n \ge (k-1)\left(\frac{k^2}{3}+\frac{5k}{6}+1\right)$. Recall that a simply-connected Lie group endowed with a left-invariant distribution V_1 is a Carnot group if the Lie algebra $\mathbf{g}$ is a graded nilpotent Lie algebra such that it is Lie generated by the block with lowest grading (i.e. $\mathbf{g} = V_1 \oplus V_2 \oplus \ldots \oplus V_k$, $[V_i, V_j] = V_{i+j}$, $V_i = 0$ if $i > k$ and V_1 Lie-generates $\mathbf{g}$).

Clearly, if the cost is locally Lipschitz out of the diagonal, then the statement of Theorem 4.1 remains valid with the extra assumption $\mathrm{supp}(\mu) \cap \mathrm{supp}(\nu) = \emptyset$.

მათემატიკა

აკადემიკოს რევაზ გამყრელიძის დაბადებიდან მე-80 წლისთავისადმი მიძღვნილი სიმპოზიუმის მასალები (ბათუმი, 17-21 სექტემბერი, 2007)

ზომის ოპტიმალური მართვა

ა. აგრაჩევი

SISSA-ISAS, ტრიესტი, იტალია, MIAN, მოსკოვი, რუსეთი

ნაშრომში შევისწავლით სტანდარტული სასრულგანზომილებიანი ოპტიმალური მართვის ამოცანის რანდომიზაციას. ვუშვებთ, რომ ტრაექტორიის სასაზღვრო მნიშვნელობები არ არის ფიქსირებული, მაგრამ ცნობილია მათი მათემატიკური ლოდინი და ხდება მოსალოდნელი ფასის მინიმიზაციის მიღწევა. ეს ამოცანა არის “მასის ოპტიმალური გადატანის” ამოცანა, დასმული მართვის თეორიის თვალსაზრისით. ვიკვლევთ ამონახსნის არსებობასა და ერთადერთობას, აგრეთვე დავახასიათებთ იმ ფუნქციონალს, რომელიც აზდენს მინიმიზაციას.

ნაშრომი ეძღვნება რევაზ გამყრელიძის დაბადების 80 წლისთავს და ეფუძნება ავტორისა და პაულ ლის ერთობლივ ნაშრომს.

REFERENCES

1. *R. V. Gamkrelidze* (1978), Principles of Optimal Control Theory, New York: Plenum Publishing Corporation.
2. *C. Villani* (2003), Topics in Mass Transportation, AMS. – Providence (Rhode Island).
3. *A. A. Agrachev, P. Lee* (2007), Control theory and optimal mass transport. Preprint, SISSA: **68**, M., 32 p.
4. *L. S. Pontryagin, V. G. Boltyanskij, R. V. Gamkrelidze, E. F. Mishchenko* (1964), The Mathematical Theory of Optimal Processes, Oxford: Pergamon Press.
5. *A. A. Agrachev, Y. L. Sachkov* (2004), Control Theory from the Geometric Viewpoint: Springer.
6. *A. A. Agrachev, R. V. Gamkrelidze* (2006), The Pontryagin Maximum Principle 50 years later. – Proc. Steklov Math. Inst., Dynamical Systems: Modelling, Optimization, and Control, 4-12.
7. *L. Ambrosio, S. Rigot* (2004), Optimal mass transportation in the Heisenberg group. – J. Func. Anal. **208**, 261-301.
8. *A. A. Agrachev, A. V. Sarychev* (1996), Abnormal sub-Riemannian Geodesics, Morse index and rigidity. – Annales de l'Institut Henri Poincaré-Analyse non linéaire, v.13, 635-690.
9. *R. Montgomery* (2002), A tour of sub-Riemannian geometries, their geodesics and applications, AMS.
10. *A. A. Agrachev, J.P. Gauthier* (2001), On the subanalyticity of Carnot-Caratheodory distances. – Ann. I. H. Poincaré - AN 18, 359-382.