A. A. Agrachev[1] **Yu. L. Sachkov**[2]

Control Theory from the Geometric Viewpoint*

[1]SISSA-ISAS, 2-4 Via Beirut, Trieste 34014, Italy
& Steklov Mathematical Institute, 8 ul. Gubkina, Moscow 117966, Russia
E-mail: agrachev@ma.sissa.it
[2]Program Systems Institute, Pereslavl-Zalessky 152140, Russia
E-mail: sachkov@sys.botik.ru
*Work partially supported by Russian Foundation for Basic Research, Project No. 02-01-00506

Preface

This book presents some facts and methods of the Mathematical Control Theory treated from the geometric point of view. The book is mainly based on graduate courses given by the first coauthor in the years 2000–2001 at the International School for Advanced Studies, Trieste, Italy. Mathematical prerequisites are reduced to standard courses of Analysis and Linear Algebra plus some basic Real and Functional Analysis. No preliminary knowledge of Control Theory or Differential Geometry is required.

What this book is about? The classical deterministic physical world is described by smooth dynamical systems: the future in such a system is completely determined by the initial conditions. Moreover, the near future changes smoothly with the initial data. If we leave room for "free will" in this fatalistic world, then we come to control systems. We do so by allowing certain parameters of the dynamical system to change freely at every instant of time. That is what we routinely do in real life with our body, car, cooker, as well as with aircraft, technological processes etc. We try to *control* all these dynamical systems!

Smooth dynamical systems are governed by differential equations. In this book we deal only with finite dimensional control systems: they are governed by ordinary differential equations on finite dimensional smooth manifolds. A control system for us is thus a family of ordinary differential equations. The family is parametrized by *control parameters.* All differential equations of the family are defined on one and the same manifold which is called the *state space* of the control system. We may select any admissible values of the control parameters (i.e. select any dynamical system from the family) and we are free to change these values at every time instant. The way of selection, which is a function of time, is called a *control* or a *control function.*

As soon as a control is fixed, the control system turns into a nonautonomous ordinary differential equation. A solution of such an equation is uniquely determined by the initial condition and is called an admissible trajectory of the control system (associated with a given control). Thus, an admissible trajectory is a curve in the state space. The initial condition (initial state) is just a starting point of the trajectory; different controls provide, generally speaking, different admissible trajectories started from a fixed state. All these trajectories fill the *attainable (reachable) set* of the given initial state.

To characterize the states reachable from a given initial one is the first natural problem to study in Control Theory: the Controllability Problem. As soon as the possibility to reach a certain state is established, we try to do it in the best way. Namely, we try to steer the initial state to the final one as fast as possible, or try to find the shortest admissible trajectory connecting the initial and the final states, or to minimize some other cost. This is the Optimal Control Problem. These two problems are our leading lights throughout the book.

Why Geometry? The right-hand side of the ordinary differential equation is a vector field and the dynamical system governed by the equation is the flow generated by this vector field. Hence a control system is a family of vector fields.

The features of control systems we study do not change under transformations induced by smooth transformations of the state space. Moreover, our systems admit a wide class of reparametrizations of the family of vector fields, which are called feedback transformations in Control Theory and gauge transformations in Geometry and Mathematical Physics. This is a formal reason why the intrinsic geometric language and geometric methods are relevant to Control Theory.

There is another more fundamental reason. As we mentioned, a dynamical system is a flow (a one-parametric group of transformations of the state space) generated by a vector field. An admissible trajectory of the control system associated to a constant control is a trajectory of the corresponding flow. Admissible trajectories associated with a piecewise constant control are realized by the composition of elements of the flows corresponding to the values of the control function. The arbitrary control case is realized via an approximation by piecewise constant controls. We see that the structure of admissible trajectories and attainable sets is intimately related to the group of transformations generated by the dynamical systems involved. In turn, groups of transformations form the heart of Geometry.

Now, what could be the position of Control techniques and the Control way of thinking in Geometry and, more generally, in the study of basic structures of the world around us? It is usual in Control Theory for the dimension of attainable sets to be essentially greater than the dimension of the sets of admissible velocities. In particular, a generic pair of vector fields on an n-dimensional manifold provides n-dimensional attainable sets, where n is as big as we want. In other words, constraints on velocities do not imply state constraints. Such a situation is traditionally indicated by saying that constraints are "nonholonomic". Control theory is a discipline that systematically studies various types of behavior under nonholonomic constraints and provides adequate methods for the investigation of variational problems with nonholonomic constraints.

The first chapter of the book is of introductory nature: we recall what smooth manifolds and ordinary differential equations on manifolds are, and define control systems. Chapter 2 is devoted to an operator calculus that creates great flexibility in handling of nonlinear control systems. In Chapters 3 and 4 we introduce a simple and extremely popular in applications class of linear systems and give an effective characterization of systems that can be made linear by a smooth transformation of the state space. Chapters 5–7 are devoted to the fundamental Orbit Theorem of Nagano and Sussmann and its applications. The Orbit Theorem states that any orbit of the group generated by a family of flows is an immersed submanifold. Chapter 8 contains general results on the structure of attainable sets starting from a simple test to guarantee that these sets are full dimensional. In Chapter 9 we introduce feedback transformations, give a feedback classification of linear systems, and effectively characterize systems that can be made linear by feedback and state transformations.

The rest of the book is mainly devoted to the Optimal Control. In Chapter 10 we state the optimal control problem, give its geometric interpretation, and discuss the existence of solutions. Chapter 11 contains basic facts on differential forms and Hamiltonian systems; we need this information to investigate optimal control problems. Chapter 12 is devoted to the intrinsic formulation and detailed proof of the Pontryagin Maximum Principle, a key result in the Optimal Control Theory. Chapters 13–16 contain numerous applications of the Pontryagin Maximum Principle including a curious property of Hamiltonian systems with convex Hamiltonians and more or less complete theories of linear time-optimal problems and linear–quadratic problems with finite horizons. In Chapter 17 we discuss a Hamiltonian version of the theory of fields of extremals, which is suitable for applications in the Optimal Control, and introduce the Hamilton–Jacobi equation. Chapters 18 and 19 are devoted to the moving frames technique for optimal control problems and to problems on Lie groups. The definition and basic facts on Lie groups are given in Chapter 18: they are simple corollaries of the general geometric control techniques developed in previous chapters. Chapters 20 and 21 contain the theory of the Second Variation with second order necessary and sufficient optimality conditions for regular and singular extremals. The short Chapter 22 presents an instructive reduction procedure, which establishes a connection between singular and regular extremals. In Chapter 23 we introduce and compute (in simplest low dimensional cases) the curvature, a remarkable feedback invariant of optimal control problems. Finally in Chapter 24 we discuss the control of a classical nonholonomic system: two bodies rolling one on another without slipping or twisting. The Appendix contains proofs of some results formulated in Chapter 2.

This is a very brief overview of the contents of the book. In each chapter we try to stay at textbook level, i.e. to present just the first basic results with some applications. The topic of practically every chapter has an extensive development, sometimes rather impressive. In order to study these topics deeper the reader is referred to research papers.

Geometric Control Theory is a broad subject and many important topics are not even mentioned in the book. In particular, we do not study the feedback stabilization problem and the huge theory of control systems with outputs including fundamental concepts of Observability and Realization. For this and other material see books on Control listed in the Bibliography.

Acknowledgments.

We wish to thank our Teachers Professor R. V. Gamkrelidze and Professor A. F. Filippov for their Lessons on Mathematics and Control Theory, and for encouragement during our work on this book.

We are grateful to Ulysse Serres for valuable remarks that helped us to improve exposition.

We are indebted to Irina and Elena: without patience and constant support of our wives, this book could not be written.

A. A. Agrachev, Yu. L. Sachkov

Moscow – Pereslavl-Zalessky – Trieste
November 2002

Contents

1 Vector fields and control systems **1**
1.1 Smooth manifolds 1
1.2 Vector fields on smooth manifolds 3
1.3 Smooth ODEs and flows on manifolds 6
1.4 Control systems 9

2 Elements of Chronological Calculus **15**
2.1 Points, diffeomorphisms, and vector fields 15
2.2 Seminorms and $C^\infty(M)$-topology 19
2.3 Families of functionals and operators 20
2.4 Chronological exponential 22
2.5 Action of diffeomorphisms on vector fields 30
2.6 Commutation of flows 33
2.7 Variations formula 34
2.8 Derivative of flow 36

3 Linear systems **39**
3.1 Cauchy's formula for linear systems 39
3.2 Controllability of linear systems 41

4 State linearizability **45**
4.1 Local linearizability 45
4.2 Global linearizability 49

5 The Orbit Theorem **55**
5.1 Formulation of the Orbit Theorem 55
5.2 Immersed submanifolds 56
5.3 Corollaries of the Orbit Theorem 57
5.4 Proof of the Orbit Theorem 59
5.5 Analytic case 62
5.6 Frobenius Theorem 65
5.7 State equivalence of control systems 66

6 Rotations of the rigid body **71**
6.1 State space 71
6.2 Euler equations 74
6.3 Phase portrait 77
6.4 Controlled rigid body: orbits 79

7 Control of configurations **85**
7.1 Model 85
7.2 Two free points 88
7.3 Three free points 89
7.4 Broken line 92

8 Attainable sets **95**
8.1 Full-rank systems 95
8.2 Compatible systems and relaxations 98
8.3 Poisson stability 100
8.4 Controlled rigid body: attainable sets 102

9 Feedback and state equivalence **105**
9.1 Feedback equivalence 105
9.2 Linear systems with scalar control 107
9.3 Linear systems with vector control parameters 110
9.4 State-feedback linearizability 114

10 Optimal control problem **121**
10.1 Problem statement 121
10.2 Reduction to study of attainable sets 122
10.3 Compactness of attainable sets 124
10.4 Time-optimal problem 126
10.5 Relaxations 127

11 Exterior Calculus **129**
11.1 Differential 1-forms 129
11.2 Differential k-forms 132
11.3 Exterior differential 136
11.4 Lie derivative of differential forms 137
11.5 Elements of Symplectic Geometry 140

12 Pontryagin Maximum Principle **151**
12.1 Geometric statement of PMP and discussion 151
12.2 Proof of PMP 155
12.3 Geometric statement of PMP for free time 160
12.4 PMP for optimal control problems 162
12.5 PMP with general boundary conditions 165

13 Examples of optimal control problems 173
13.1 The fastest stop of a train at a station 173
13.2 Control of a linear oscillator 176
13.3 The cheapest stop of a train 178
13.4 Control of a linear oscillator with cost 180
13.5 Dubins car 181

14 Convex Hamiltonians 187

15 Linear time-optimal problem 191
15.1 Problem statement 191
15.2 Geometry of polytopes 192
15.3 Bang-bang theorem 192
15.4 Uniqueness of optimal controls and extremals 195
15.5 Switchings of optimal control 197

16 Linear-quadratic problem 203
16.1 Problem statement and assumptions 203
16.2 Existence of optimal control 204
16.3 Extremals 207
16.4 Conjugate points 208

17 Sufficient optimality conditions 215
17.1 Sufficient optimality conditions 215
17.2 Hamilton-Jacobi equation 222
17.3 Dynamic programming 223

18 Geometric optimal control problems 225
18.1 Trivialized cotangent bundle 225
18.2 Lie groups 233
18.3 Hamiltonian systems on Lie groups 237

19 Problems on compact Lie groups 243
19.1 Riemannian problem 243
19.2 A sub-Riemannian problem 245
19.3 Control of quantum systems 249
19.4 A time-optimal problem on SO(3) 261

20 Second order optimality conditions 269
20.1 Hessian 269
20.2 Local openness of mappings 272
20.3 Differentiation of the endpoint mapping 280
20.4 Necessary optimality conditions 284
20.5 Applications 293
20.6 Single-input case 296

21 Jacobi equation **307**
21.1 Regular case: derivation of Jacobi equation 308
21.2 Singular case: derivation of Jacobi equation 312
21.3 Necessary optimality conditions 316
21.4 Transformation of Jacobi equation 317
21.5 Sufficient optimality conditions 320

22 Reduction **329**
22.1 Reduction 329
22.2 Rigid body control 332
22.3 Angular velocity control 333

23 Curvature **335**
23.1 Curvature of 2-dimensional systems 335
23.2 Curvature of 3-dimensional affine systems 345

24 Rolling bodies **349**
24.1 Geometric model 349
24.2 Two-dimensional Riemannian geometry 350
24.3 Admissible velocities 354
24.4 Controllability 355
24.5 Length minimization problem 358

A Appendix **363**
A.1 Homomorphisms and operators in $C^\infty(M)$ 363
A.2 Remainder term 365

Bibliography **369**

Index **371**

Chapter 1

Vector fields and control systems on smooth manifolds

1.1 Smooth manifolds

We give just a brief outline of basic notions related to the smooth manifolds. For a consistent presentation, see an introductory chapter to any textbook on analysis on manifolds, e.g. [24].

In the sequel, "smooth" (manifold, mapping, vector field etc.) means C^∞.

Definition 1.1. A subset $M \subset \mathbb{R}^n$ is called a *smooth k-dimensional submanifold* of $\mathbb{R}^n$, $k \leq n$, if any point $x \in M$ has a neighborhood O_x in $\mathbb{R}^n$ in which M is described in one of the following ways:

(1) there exists a smooth vector-function

$$F : O_x \to \mathbb{R}^{n-k}, \qquad \operatorname{rank} \left.\frac{dF}{dx}\right|_x = n - k$$

such that

$$O_x \cap M = F^{-1}(0);$$

(2) there exists a smooth vector-function

$$f : V_0 \to \mathbb{R}^n$$

from a neighborhood of the origin $0 \in V_0 \subset \mathbb{R}^k$ with

$$f(0) = x, \qquad \operatorname{rank} \left.\frac{df}{dx}\right|_0 = k$$

such that

$$O_x \cap M = f(V_0)$$

and $f : V_0 \to O_x \cap M$ is a homeomorphism;

(3) there exists a smooth vector-function

$$\Phi \ : \ O_x \to O_0 \subset \mathbb{R}^n$$

onto a neighborhood of the origin $0 \in O_0 \subset \mathbb{R}^n$ with

$$\operatorname{rank} \left. \frac{d\,\Phi}{d\,x} \right|_x = n$$

such that

$$\Phi(O_x \cap M) = \mathbb{R}^k \cap O_0.$$

Exercise 1.1. Prove that three local descriptions of a smooth submanifold given in (1)–(3) are mutually equivalent.

Remarks. (1) There are two topologically different one-dimensional manifolds: the line $\mathbb{R}^1$ and the circle S^1. The sphere S^2 and the torus $T^2 = S^1 \times S^1$ are two-dimensional manifolds. The torus can be viewed as a sphere with a handle. Any compact orientable two-dimensional manifold is topologically a sphere with p handles, $p = 0, 1, 2, \ldots$.

(2) Smooth manifolds appear naturally already in the basic analysis. For example, the circle S^1 and the torus T^2 are natural domains of periodic and doubly periodic functions respectively. On the sphere S^2, it is convenient to consider restriction of homogeneous functions of 3 variables.

So a smooth submanifold is a subset in $\mathbb{R}^n$ which can locally be defined by a regular system of smooth equations and by a smooth regular parametrization.

In spite of the intuitive importance of the image of manifolds as subsets of a Euclidean space, it is often convenient to consider manifolds independently of any embedding in $\mathbb{R}^n$. An abstract manifold is defined as follows.

Definition 1.2. A *smooth k-dimensional manifold M* is a Hausdorff paracompact topological space endowed with a smooth structure: M is covered by a system of open subsets

$$M = \cup_\alpha O_\alpha$$

called coordinate neighborhoods, in each of which is defined a homeomorphism

$$\Phi_\alpha \ : \ O_\alpha \to \mathbb{R}^k$$

called a local coordinate system such that all compositions

$$\Phi_{\alpha_1} \circ \Phi_{\alpha_2}^{-1} \ : \Phi_{\alpha_2}(O_{\alpha_1} \cap O_{\alpha_2}) \subset \mathbb{R}^k \to \Phi_{\alpha_1}(O_{\alpha_1} \cap O_{\alpha_2}) \subset \mathbb{R}^k$$

are diffeomorphisms.

Remark. For a smooth submanifold in $\mathbb{R}^n$, the abstract definition holds.

Conversely, any connected smooth abstract manifold can be considered as a smooth submanifold in $\mathbb{R}^n$. Before precise formulation of this statement, we give two definitions.

Definition 1.3. Let M and N be k- and l-dimensional smooth manifolds respectively. A mapping

$$f \ : \ M \to N$$

is called *smooth* if it is smooth in coordinates. That is, let $M = \cup_\alpha O_\alpha$ and $N = \cup_\beta V_\beta$ be coverings of M and N by coordinate neighborhoods and

$$\Phi_\alpha \ : \ O_\alpha \to \mathbb{R}^k, \qquad \Psi_\beta \ : \ V_\beta \to \mathbb{R}^l$$

the corresponding coordinate mappings. Then all compositions

$$\Psi_\beta \circ f \circ \Phi_\alpha^{-1} \ : \ \Phi_\alpha(O_\alpha \cap f^{-1}(V_\beta)) \subset \mathbb{R}^k \to \Psi_\beta(f(O_\alpha) \cap V_\beta) \subset \mathbb{R}^l$$

should be smooth.

Definition 1.4. A smooth manifold M is called *diffeomorphic* to a smooth manifold N if there exists a homeomorphism

$$f \ : \ M \to N$$

such that both f and its inverse f^{-1} are smooth mappings. Such mapping f is called a *diffeomorphism*.

The set of all diffeomorphisms $f \ : \ M \to M$ of a smooth manifold M is denoted by $\operatorname{Diff} M$.

A smooth mapping $f \ : \ M \to N$ is called an *embedding* of M into N if $f \ : \ M \to f(M)$ is a diffeomorphism. A mapping $f \ : \ M \to N$ is called *proper* if $f^{-1}(K)$ is compact for any compactum $K \Subset N$.

Theorem 1.1 (Whitney). *Any smooth connected k-dimensional manifold can be properly embedded into $\mathbb{R}^{2k+1}$.*

Summing up, we may say that a smooth manifold is a space which looks locally like a linear space but without fixed linear structure, so that all smooth coordinates are equivalent. The manifolds, not linear spaces, form an adequate framework for the modern nonlinear analysis.

1.2 Vector fields on smooth manifolds

The tangent space to a smooth manifold at a point is a linear approximation of the manifold in the neighborhood of this point.

Definition 1.5. Let M be a smooth k-dimensional submanifold of $\mathbb{R}^n$ and $x \in M$ its point. Then the *tangent space* to M at the point x is a k-dimensional linear subspace

$$T_x M \subset \mathbb{R}^n$$

defined as follows for cases (1)–(3) of Definition 1.1:

(1) $\quad T_x M = \operatorname{Ker} \left.\dfrac{dF}{dx}\right|_x,$

(2) $\quad T_x M = \operatorname{Im} \left.\dfrac{df}{dx}\right|_0,$

(3) $\quad T_x M = \left(\left.\dfrac{d\Phi}{dx}\right|_x\right)^{-1} \mathbb{R}^k.$

Remark. The tangent space is a coordinate-invariant object since smooth changes of variables lead to linear transformations of the tangent space.

In an abstract way, the tangent space to a manifold at a point is the set of velocity vectors to all smooth curves in the manifold that start from this point.

Definition 1.6. Let $\gamma(\cdot)$ be a smooth curve in a smooth manifold M starting from a point $x \in M$:

$$\gamma \ : \ (-\varepsilon, \varepsilon) \to M \text{ a smooth mapping}, \qquad \gamma(0) = x.$$

The *tangent vector*

$$\left.\frac{d\gamma}{dt}\right|_{t=0} = \dot{\gamma}(0)$$

to the curve $\gamma(\cdot)$ at the point x is the equivalence class of all smooth curves in M starting from x and having the same 1-st order Taylor polynomial as $\gamma(\cdot)$, for any coordinate system in a neighborhood of x.

Definition 1.7. The *tangent space* to a smooth manifold M at a point $x \in M$ is the set of all tangent vectors to all smooth curves in M starting at x:

$$T_x M = \left\{ \left.\frac{d\gamma}{dt}\right|_{t=0} \mid \gamma \ : \ (-\varepsilon, \varepsilon) \to M \text{ smooth}, \gamma(0) = x \right\}.$$

Exercise 1.2. Let M be a smooth k-dimensional manifold and $x \in M$. Show that the tangent space $T_x M$ has a natural structure of a linear k-dimensional space.

Definition 1.8. A *smooth vector field* on a smooth manifold M is a smooth mapping

$$x \in M \mapsto V(x) \in T_x M$$

that associates to any point $x \in M$ a tangent vector $V(x)$ at this point.

In the sequel we denote by $\operatorname{Vec} M$ the set of all smooth vector fields on a smooth manifold M.

Definition 1.9. A *smooth dynamical system*, or an *ordinary differential equation* (*ODE*), on a smooth manifold M is an equation of the form

$$\frac{dx}{dt} = V(x), \qquad x \in M,$$

or, equivalently,

$$\dot{x} = V(x), \qquad x \in M,$$

where $V(x)$ is a smooth vector field on M. A solution to this system is a smooth mapping

$$\gamma : I \to M,$$

where $I \subset \mathbb{R}$ is an interval, such that

$$\frac{d\gamma}{dt} = V(\gamma(t)) \quad \forall\, t \in I.$$

Definition 1.10. Let $\Phi : M \to N$ be a smooth mapping between smooth manifolds M and N. The *differential* of Φ at a point $x \in M$ is a linear mapping

$$D_x\Phi : T_xM \to T_{\Phi(x)}N$$

defined as follows:

$$D_x\Phi\left(\left.\frac{d\gamma}{dt}\right|_{t=0}\right) = \left.\frac{d}{dt}\right|_{t=0}\Phi(\gamma(t)),$$

where

$$\gamma : (-\varepsilon, \varepsilon) \subset \mathbb{R} \to M, \qquad \gamma(0) = x,$$

is a smooth curve in M starting at x.

Now we apply smooth mappings to vector fields. Let $V \in \operatorname{Vec} M$ be a vector field on M and

$$\dot{x} = V(x) \tag{1.1}$$

the corresponding ODE. To find the action of a diffeomorphism

$$\Phi : M \to N, \qquad \Phi : x \mapsto y = \Phi(x)$$

on the vector field $V(x)$, take a solution $x(t)$ of (1.1) and compute the ODE satisfied by the image $y(t) = \Phi(x(t))$:

$$\dot{y}(t) = \frac{d}{dt}\Phi(x(t)) = (D_x\Phi)\,\dot{x}(t) = (D_x\Phi)\,V(x(t)) = (D_{\Phi^{-1}(y)}\Phi)\,V(\Phi^{-1}(y(t))).$$

So the required ODE is

$$\dot{y} = \left(D_{\Phi^{-1}(y)}\Phi\right)V(\Phi^{-1}(y)). \tag{1.2}$$

The right-hand side here is the transformed vector field on N induced by the diffeomorphism Φ:

$$(\Phi_* V)(y) \stackrel{\text{def}}{=} \left(D_{\Phi^{-1}(y)}\Phi\right)V(\Phi^{-1}(y)).$$

The notation Φ_{*x} is used, along with $D_x\Phi$, for differential of a mapping Φ at a point x.

Remark. In general, a smooth mapping Φ induces transformation of tangent vectors, not of vector fields. In order that $D\Phi$ transform vector fields to vector fields, Φ should be a diffeomorphism.

1.3 Smooth differential equations and flows on manifolds

Theorem 1.2. *Consider a smooth ODE*

$$\dot{x} = V(x), \qquad x \in M \subset \mathbb{R}^n, \tag{1.3}$$

on a smooth submanifold M of $\mathbb{R}^n$. For any initial point $x_0 \in M$, there exists a unique solution

$$x(t, x_0), \qquad t \in (a, b), \quad a < 0 < b,$$

of equation (1.3) *with the initial condition*

$$x(0, x_0) = x_0,$$

defined on a sufficiently short interval (a, b). The mapping

$$(t, x_0) \mapsto x(t, x_0)$$

is smooth. In particular, the domain (a, b) of the solution $x(\cdot, x_0)$ can be chosen smoothly depending on x_0.

Proof. We prove the theorem by reduction to its classical analog in $\mathbb{R}^n$.

The statement of the theorem is local. We rectify the submanifold M in the neighborhood of the point x_0:

$$\Phi : O_{x_0} \subset \mathbb{R}^n \to O_0 \subset \mathbb{R}^n,$$
$$\Phi(O_{x_0} \cap M) = \mathbb{R}^k.$$

Consider the restriction $\varphi = \Phi|_M$. Then a curve $x(t)$ in M is a solution to (1.3) if and only if its image $y(t) = \varphi(x(t))$ in $\mathbb{R}^k$ is a solution to the induced system:

$$\dot{y} = \Phi_* V(y), \qquad y \in \mathbb{R}^k.$$

□

Theorem 1.3. *Let $M \subset \mathbb{R}^n$ be a smooth submanifold and let*

$$\dot{x} = V(x), \qquad x \in \mathbb{R}^n, \tag{1.4}$$

be a system of ODEs in $\mathbb{R}^n$ such that

$$x \in M \;\Rightarrow\; V(x) \in T_x M.$$

Then for any initial point $x_0 \in M$, the corresponding solution $x(t, x_0)$ to (1.4) *with $x(0, x_0) = x_0$ belongs to M for all sufficiently small $|t|$.*

Proof. Consider the restricted vector field:

$$f = V|_M.$$

By the existence theorem for M, the system

$$\dot{x} = f(x), \qquad x \in M,$$

has a solution $x(t, x_0)$, $x(0, x_0) = x_0$, with

$$x(t, x_0) \in M \qquad \text{for small } |t|. \tag{1.5}$$

On the other hand, the curve $x(t, x_0)$ is a solution of (1.4) with the same initial condition. Then inclusion (1.5) proves the theorem. □

Definition 1.11. A vector field $V \in \operatorname{Vec} M$ is called *complete*, if for all $x_0 \in M$ the solution $x(t, x_0)$ of the Cauchy problem

$$\dot{x} = V(x), \qquad x(0, x_0) = x_0 \tag{1.6}$$

is defined for all $t \in \mathbb{R}$.

Example 1.1. The vector field $V(x) = x$, $x \in \mathbb{R}$, is complete, but the vector field $V(x) = x^2$, $x \in \mathbb{R}$, is not complete.

Proposition 1.1. *Suppose that there exists $\varepsilon > 0$ such that for any $x_0 \in M$ the solution $x(t, x_0)$ to* (1.6) *is defined for $t \in (-\varepsilon, \varepsilon)$. Then the vector field $V(x)$ is complete.*

Remark. In this proposition it is required that there exists $\varepsilon > 0$ common for all initial points $x_0 \in M$. In general, ε may be not bounded away from zero for all $x_0 \in M$. E.g., for the vector field $V(x) = x^2$ we have $\varepsilon \to 0$ as $x_0 \to \infty$.

Proof. Suppose that the hypothesis of the proposition is true. Then we can introduce the following family of mappings in M:

$$\begin{aligned} P^t &: M \to M, \qquad t \in (-\varepsilon, \varepsilon), \\ P^t &: x_0 \mapsto x(t, x_0). \end{aligned}$$

$P^t(x_0)$ is the shift of a point $x_0 \in M$ along the trajectory of the vector field $V(x)$ for time t.

By Theorem 1.2, all mappings P^t are smooth. Moreover, the family $\{ P^t \mid t \in (-\varepsilon, \varepsilon) \}$ is a smooth family of mappings.

A very important property of this family is that it forms a local one-parameter group, i.e.,

$$P^t(P^s(x)) = P^s(P^t(x)) = P^{t+s}(x), \quad x \in M, \quad t,\ s,\ t+s \in (-\varepsilon, \varepsilon).$$

Indeed, the both curves in M:

$$t \mapsto P^t(P^s(x)) \quad \text{and} \quad t \mapsto P^{t+s}(x)$$

satisfy the ODE $\dot{x} = V(x)$ with the same initial value $P^0(P^s(x)) = P^{0+s}(x) = P^s(x)$. By uniqueness, $P^t(P^s(x)) = P^{t+s}(x)$. The equality for $P^s(P^t(x))$ is obtained by switching t and s.

So we have the following local group properties of the mappings P^t:

$$\begin{aligned}
&P^t \circ P^s = P^s \circ P^t = P^{t+s}, \qquad t,\ s,\ t+s \in (-\varepsilon, \varepsilon),\\
&P^0 = \mathrm{Id},\\
&P^{-t} \circ P^t = P^t \circ P^{-t} = \mathrm{Id}, \qquad t \in (-\varepsilon, \varepsilon),\\
&P^{-t} = \left(P^t\right)^{-1}, \qquad t \in (-\varepsilon, \varepsilon).
\end{aligned}$$

In particular, all P^t are diffeomorphisms.

Now we extend the mappings P^t for all $t \in \mathbb{R}$. Any $t \in \mathbb{R}$ can be represented as

$$t = \frac{\varepsilon}{2}K + \tau, \qquad 0 \le \tau < \frac{\varepsilon}{2}, \quad K = 0, \pm 1, \pm 2, \ldots.$$

We set

$$P^t \overset{\mathrm{def}}{=} P^\tau \circ \underbrace{P^{\pm\varepsilon/2} \circ \cdots \circ P^{\pm\varepsilon/2}}_{|K| \text{ times}}, \qquad \pm = \operatorname{sgn} t.$$

Then the curve

$$t \mapsto P^t(x_0), \qquad t \in \mathbb{R},$$

is a solution to Cauchy problem (1.6). □

Definition 1.12. For a complete vector field $V \in \operatorname{Vec} M$, the mapping

$$t \mapsto P^t, \qquad t \in \mathbb{R},$$

is called the *flow* generated by V.

Remark. It is useful to imagine a vector field $V \in \operatorname{Vec} M$ as a field of velocity vectors of a moving liquid in M. Then the flow P^t takes any particle of the liquid from a position $x \in M$ and transfers it for a time $t \in \mathbb{R}$ to the position $P^t(x) \in M$.

Simple sufficient conditions for completeness of a vector field are given in terms of compactness.

Proposition 1.2. *Let $K \subset M$ be a compact subset, and let $V \in \operatorname{Vec} M$. Then there exists $\varepsilon_K > 0$ such that for all $x_0 \in K$ the solution $x(t, x_0)$ to Cauchy problem* (1.6) *is defined for all $t \in (-\varepsilon_K, \varepsilon_K)$.*

Proof. By Theorem 1.2, domain of the solution $x(t, x_0)$ can be chosen continuously depending on x_0. The diameter of this domain has a positive infimum $2\varepsilon_K$ for x_0 in the compact set K. □

Corollary 1.1. *If a smooth manifold M is compact, then any vector field $V \in \operatorname{Vec} M$ is complete.*

Corollary 1.2. *Suppose that a vector field $V \in \operatorname{Vec} M$ has a compact support:*

$$\operatorname{supp} V \stackrel{\text{def}}{=} \overline{\{ x \in M \mid V(x) \neq 0 \}} \text{ is compact.}$$

Then V is complete.

Proof. Indeed, by Proposition 1.2, there exists $\varepsilon > 0$ such that all trajectories of V starting in $\operatorname{supp} V$ are defined for $t \in (-\varepsilon, \varepsilon)$. But $V|_{M \setminus \operatorname{supp} V} = 0$, and all trajectories of V starting outside of $\operatorname{supp} V$ are constant, thus defined for all $t \in \mathbb{R}$. By Proposition 1.1, the vector field V is complete. $\square$

Remark. If we are interested in the behavior of (trajectories of) a vector field $V \in \operatorname{Vec} M$ in a compact subset $K \subset M$, we can suppose that V is complete. Indeed, take an open neighborhood O_K of K with the compact closure $\overline{O_K}$. We can find a function $a \in C^\infty(M)$ such that

$$a(x) = \begin{cases} 1, & x \in K, \\ 0, & x \in M \setminus O_K. \end{cases}$$

Then the vector field $a(x)V(x) \in \operatorname{Vec} M$ is complete since it has a compact support. On the other hand, in K the vector fields $a(x)V(x)$ and $V(x)$ coincide, thus have the same trajectories.

1.4 Control systems

For dynamical systems, the future $x(t, x_0)$, $t > 0$, is completely determined by the present state $x_0 = x(0, x_0)$. The law of transformation $x_0 \mapsto x(t, x_0)$ is the flow P^t, i.e., dynamics of the system

$$\dot{x} = V(x), \qquad x \in M, \tag{1.7}$$

it is determined by one vector field $V(x)$.

In order to be able to affect dynamics, to control it, we consider a family of dynamical systems

$$\dot{x} = V_u(x), \qquad x \in M, \quad u \in U, \tag{1.8}$$

with a family of vector fields V_u parametrized by a parameter $u \in U$. A system of the form (1.8) is called a *control system*. The variable u is a *control parameter*, and the set U is the *space of control parameters*. A priori we do not impose any restrictions on U, it is an arbitrary set, although, typically U will be a subset of a smooth manifold. The variable x is the *state*, and the manifold M is the *state space* of control system (1.8).

In control theory we can change dynamics of control system (1.8) at any moment of time by changing values of $u \in U$. For any $u \in U$, the corresponding vector field $V_u \in \operatorname{Vec} M$ generates the flow, which is denoted by P_u^t.

A typical problem of control theory is to find the set of points that can be reached from an initial point $x_0 \in M$ by choosing various values of $u \in U$ and switching from one value to another time to time (for dynamical system (1.7), this reachable set is just the semitrajectory $x(t, x_0) = P^t(x_0)$, $t \geq 0$). Suppose that we start from a point $x_0 \in M$ and use the following control strategy for control system (1.8): first we choose some control parameter $u_1 \in U$, then we switch to another control parameter $u_2 \in U$. Which points in M can be reached with such control strategy? With the control parameter u_1, we can reach points of the form

$$\{ P^{t_1}_{u_1}(x_0) \mid t_1 \geq 0 \},$$

and the whole set of reachable points has the form

$$\{ P^{t_2}_{u_2} \circ P^{t_1}_{u_1}(x_0) \mid t_1,\, t_2 \geq 0 \},$$

a piece of a 2-dimensional surface.

A natural next question is: what points can be reached from x_0 by any kind of control strategies?

Before studying this question, consider a particular control system that gives a simplified model of a car.

Example 1.2. We suppose that the state of a car is determined by the position of its center of mass $x = (x^1, x^2) \in \mathbb{R}^2$ and orientation angle $\theta \in S^1$ relative to the positive direction of the axis x^1. Thus the state space of our system is a nontrivial 3-dimensional manifold, a solid torus

$$M = \{ q = (x, \theta) \mid x \in \mathbb{R}^2,\ \theta \in S^1 \} = \mathbb{R}^2 \times S^1.$$

Suppose that two kinds of motion are possible: we can drive the car forward and backwards with some fixed linear velocity $u_1 \in \mathbb{R}$, and we can turn the car around its center of mass with some fixed angular velocity $u_2 \in \mathbb{R}$. We can combine these two kinds of motion in an admissible way.

The dynamical system that describes the linear motion with a velocity $u_1 \in \mathbb{R}$ has the form

$$\begin{cases} \dot{x}^1 = u_1 \cos\theta, \\ \dot{x}^2 = u_1 \sin\theta, \\ \dot{\theta} = 0. \end{cases} \tag{1.9}$$

Rotation with an angular velocity $u_2 \in \mathbb{R}$ is described as

$$\begin{cases} \dot{x}^1 = 0, \\ \dot{x}^2 = 0, \\ \dot{\theta} = u_2. \end{cases} \tag{1.10}$$

The control parameter $u = (u_1, u_2)$ can take any values in the given subset $U \subset \mathbb{R}^2$. If we write ODEs (1.9) and (1.10) in the vector form:

$$\dot{q} = u_1 V_1(q), \qquad \dot{q} = u_2 V_2(q),$$

where

$$V_1(q) = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}, \qquad V_2(q) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{1.11}$$

then our model reads

$$\dot q = V_u(q) = u_1 V_1(q) + u_2 V_2(q), \qquad q \in M, \quad u \in U.$$

This model can be rewritten in the complex form:

$$\begin{aligned} & z = x^1 + ix^2 \in \mathbb{C}, \\ & \dot z = u_1 e^{i\theta}, \\ & \dot\theta = u_2, \\ & (u_1, u_2) \in U, \qquad (z, \theta) \in \mathbb{C} \times S^1. \end{aligned}$$

Remark. Control system (1.8) is often written in another form:

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

We prefer the notation $V_u(x)$, which stresses that for a fixed $u \in U$, V_u is a single object — a vector field on M.

Now we return to the study of the points reachable by trajectories of a control system from an initial point.

Definition 1.13. The *attainable set* (or *reachable set*) of control system (1.8) with piecewise-constant controls from a point $x_0 \in M$ for a time $t \geq 0$ is defined as follows:

$$\mathcal{A}_{x_0}(t) = \{ P_{u_k}^{\tau_k} \circ \ldots \circ P_{u_1}^{\tau_1}(x_0) \mid \tau_i \geq 0, \ \sum_{i=1}^{k} \tau_i = t, \ u_i \in U, \ k \in \mathbb{N} \}.$$

The attainable set from x_0 for arbitrary nonnegative time of motion has the form

$$\mathcal{A}_{x_0} = \cup_{t \geq 0} \mathcal{A}_{x_0}(t).$$

For simplicity, consider first the smallest nontrivial space of control parameters consisting of two indices:

$$U = \{1, 2\}$$

(even this simple case shows essential features of the reachability problem). Then the attainable set for arbitrary nonnegative times has the form:

$$\mathcal{A}_{x_0} = \{ P_2^{\tau_k} \circ P_1^{\tau_{k-1}} \circ \ldots \circ P_2^{\tau_2} \circ P_1^{\tau_1}(x_0) \mid \tau_i \geq 0, \ k \in \mathbb{N} \}.$$

This expression suggests that the attainable set $\mathcal{A}_{x_0}$ depends heavily upon commutator properties of the flows P_1^t and P_2^s.

Consider first the trivial commutative case, i.e., suppose that the flows commute:

$$P_1^t \circ P_2^s = P_2^s \circ P_1^t \quad \forall t,\ s \in \mathbb{R}.$$

Then the attainable set can be evaluated precisely: since

$$P_2^{\tau_k} \circ P_1^{\tau_{k-1}} \circ \ldots \circ P_2^{\tau_2} \circ P_1^{\tau_1} = P_2^{\tau_k + \ldots + \tau_2} \circ P_1^{\tau_{k-1} + \ldots + \tau_1},$$

then

$$\mathcal{A}_{x_0} = \{\, P_2^s \circ P_1^t(x_0) \mid t,\ s \geq 0 \,\}.$$

So in the commutative case the attainable set by two control parameters is a piece of a smooth two-dimensional surface, possibly with singularities. It is easy to see that if the number of control parameters is $k \geq 2$ and the corresponding flows $P_1^{t_1}, \ldots, P_k^{t_k}$ commute, then $\mathcal{A}_{x_0}$ is, in general, a piece of a k-dimensional manifold, and, in particular, $\dim \mathcal{A}_{x_0} \leq k$.

But this commutative case is exceptional and occurs almost never in real control systems.

Example 1.3. In the model of a car considered above the control dynamics is defined by two vector fields (1.11) on the 3-dimensional manifold $M = \mathbb{R}^2_x \times S^1_\theta$. It is obvious that from any initial configuration $q_0 = (x_0, \theta_0) \in M$ we can drive the car to any terminal configuration $q_1 = (x_1, \theta_1) \in M$ by alternating linear motions and rotations (both with fixed velocities). So any point in the 3-dimensional manifold M can be reached by means of 2 vector fields V_1, V_2. This is due to noncommutativity of these fields (i.e., of their flows).

Given an arbitrary pair of vector fields $V_1, V_2 \in \operatorname{Vec} M$, how can one recognize their commuting properties without finding the flows P_1^t, P_2^s explicitly, i.e., without integration of the ODEs $\dot{x} = V_1(x)$, $\dot{x} = V_2(x)$?

If the flows P_1^t, P_2^s commute, then the curve

$$\gamma(s,t) = P_1^{-t} \circ P_2^s \circ P_1^t(x) = P_2^s(x), \qquad t,\ s \in \mathbb{R}, \tag{1.12}$$

does not depend on t. It is natural to suggest that a lower-order term in the Taylor expansion of (1.12) at $t = s = 0$ is responsible for commuting properties of flows of the vector fields V_1, V_2 at the point x. The first-order derivatives

$$\left.\frac{\partial \gamma}{\partial t}\right|_{s=t=0} = 0, \qquad \left.\frac{\partial \gamma}{\partial s}\right|_{s=t=0} = V_2(x)$$

are obviously useless, as well as the pure second-order derivatives

$$\left.\frac{\partial^2 \gamma}{\partial t^2}\right|_{s=t=0} = 0, \qquad \left.\frac{\partial^2 \gamma}{\partial s^2}\right|_{s=t=0} = \left.\frac{\partial}{\partial s}\right|_{s=0} V_2(P_2^s(x)).$$

The required derivative should be the mixed second-order one

$$\left.\frac{\partial^2 \gamma}{\partial t \partial s}\right|_{s=t=0}.$$

It turns out that this derivative is a tangent vector to M. It is called the *Lie bracket* of the vector fields V_1, V_2 and is denoted by $[V_1, V_2](x)$:

$$[V_1, V_2](x) \stackrel{\text{def}}{=} \left.\frac{\partial^2}{\partial t \partial s}\right|_{t=s=0} P_1^{-t} \circ P_2^s \circ P_1^t(x) \in T_x M. \tag{1.13}$$

The vector field $[V_1, V_2] \in \operatorname{Vec} M$ determines commuting properties of V_1 and V_2 (it is often called *commutator* of vector fields V_1, V_2).

An effective formula for computing Lie bracket of vector fields in local coordinates is given in the following statement.

Proposition 1.3. *Let $M = \mathbb{R}^n$ and $V_1, V_2 \in \operatorname{Vec} M$. Then*

$$[V_1, V_2](x) = \frac{d\,V_2}{d\,x} V_1(x) - \frac{d\,V_1}{d\,x} V_2(x). \tag{1.14}$$

The proof is left to the reader as an exercise.

Another way to define Lie bracket of vector fields V_1, V_2 is to consider the path

$$\gamma(t) = P_2^{-t} \circ P_1^{-t} \circ P_2^t \circ P_1^t(x).$$

Exercise 1.3. Show that

$$\gamma(t) = x + [V_1, V_2](x) t^2 + o(t^2), \qquad t \to 0,$$

i.e., $[V_1, V_2](x)$ is the velocity vector of the C^1 curve $\gamma(\sqrt{t})$. In particular, this proves that $[V_1, V_2](x)$ is indeed a tangent vector to M:

$$[V_1, V_2](x) \in T_x M.$$

In the next chapter we will develop an efficient algebraic way to do similar calculations without any coordinates.

In the commutative case, the set of reachable points does not depend on the number of switches of a control strategy used. In the general noncommutative case, on the contrary, the greater number of switches, the more points can be reached.

Suppose that we can move along vector fields $\pm V_1$ and $\pm V_2$. Then, infinitesimally, we can move in the new direction $\pm[V_1, V_2]$, which is in general linearly independent of the initial ones $\pm V_1$, $\pm V_2$. Using the same switching control strategy with the vector fields $\pm V_1$ and $\pm[V_1, V_2]$, we add one more infinitesimal direction of motion $\pm[V_1, [V_1, V_2]]$. Analogously, we can obtain $\pm[V_2, [V_1, V_2]]$. Iterating this procedure with the new vector fields obtained at previous steps, we can have a Lie bracket of arbitrarily high order as an infinitesimal direction of motion with a sufficiently large number of switches.

Example 1.4. Compute the Lie bracket of the vector fields

$$V_1(q) = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}, \quad V_2(q) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad q = \begin{pmatrix} x_1 \\ x_2 \\ \theta \end{pmatrix} \in \mathbb{R}^2_{(x_1,x_2)} \times S^1_\theta$$

appearing in the model of a car. Recall that the field V_1 generates the forward motion, and V_2 the counterclockwise rotation of the car. By (1.14), we have

$$\begin{aligned}[][V_1, V_2](q) &= \frac{d\,V_2}{d\,q} V_1(q) - \frac{d\,V_1}{d\,q} V_2(q) = -\begin{pmatrix} 0 & 0 & -\sin\theta \\ 0 & 0 & \cos\theta \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \sin\theta \\ -\cos\theta \\ 0 \end{pmatrix}. \end{aligned}$$

The vector field $[V_1, V_2]$ generates the motion of the car in the direction perpendicular to orientation of the car. This is a typical maneuver in parking a car: the sequence of 4 motions with the same small amplitude of the form

motion forward $\to$ rotation counterclockwise $\to$ motion backward $\to$
$\to$ rotation clockwise

results in motion to the right (in the main term).

We show this explicitly by computing the Lie bracket $[V_1, V_2]$ as in Example 1.3:

$$\begin{aligned} P_2^{-t} \circ P_1^{-t} \circ P_2^{t} \circ P_1^{t} \begin{pmatrix} x_1 \\ x_2 \\ \theta \end{pmatrix} &= \begin{pmatrix} x_1 + t(\cos\theta - \cos(\theta + t)) \\ x_2 + t(\sin\theta - \sin(\theta + t)) \\ \theta \end{pmatrix} \\ = \begin{pmatrix} x_1 \\ x_2 \\ \theta \end{pmatrix} + t^2 \begin{pmatrix} \sin\theta \\ -\cos\theta \\ 0 \end{pmatrix} &+ o(t^2), \qquad t \to 0, \end{aligned}$$

and we have once more

$$[V_1, V_2](q) = \begin{pmatrix} \sin\theta \\ -\cos\theta \\ 0 \end{pmatrix}. \tag{1.15}$$

Of course, we can also compute this Lie bracket by definition as in (1.13):

$$\begin{aligned} P_1^{-t} \circ P_2^{s} \circ P_1^{t} \begin{pmatrix} x_1 \\ x_2 \\ \theta \end{pmatrix} &= \begin{pmatrix} x_1 + t(\cos\theta - \cos(\theta + s)) \\ x_2 + t(\sin\theta - \sin(\theta + s)) \\ \theta + s \end{pmatrix} \\ = \begin{pmatrix} x_1 \\ x_2 \\ \theta \end{pmatrix} + s \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + ts \begin{pmatrix} \sin\theta \\ -\cos\theta \\ 0 \end{pmatrix} &+ O(t^2 + s^2)^{3/2}, \quad t, s \to 0, \end{aligned}$$

and the Lie bracket (1.15) follows.

Chapter 2

Elements of Chronological Calculus

We introduce an operator calculus that will allow us to work with nonlinear systems and flows as with linear ones, at least at the formal level. The idea is to replace a nonlinear object, a smooth manifold M, by a linear, although infinite-dimensional one: the commutative algebra of smooth functions on M (for details, see [1], [2]). For basic definitions and facts of functional analysis used in this chapter, one can consult e.g. [20].

2.1 Points, diffeomorphisms, and vector fields

In this section we identify points, diffeomorphisms, and vector fields on the manifold M with functionals and operators on the algebra $C^\infty(M)$.

Addition, multiplication, and product with constants are defined in the algebra $C^\infty(M)$, as usual, pointwise: if $a, b \in C^\infty(M)$, $q \in M$, $\alpha \in \mathbb{R}$, then

$$
\begin{aligned}
(a+b)(q) &= a(q) + b(q), \\
(a \cdot b)(q) &= a(q) \cdot b(q), \\
(\alpha \cdot a)(q) &= \alpha \cdot a(q).
\end{aligned}
$$

Any point $q \in M$ defines a linear functional

$$\widehat{q} : C^\infty(M) \to \mathbb{R}, \qquad \widehat{q}a = a(q), \ a \in C^\infty(M).$$

The functionals $\widehat{q}$ are homomorphisms of the algebras $C^\infty(M)$ and $\mathbb{R}$:

$$
\begin{aligned}
\widehat{q}(a+b) &= \widehat{q}a + \widehat{q}b, \qquad a, \ b \in C^\infty(M), \\
\widehat{q}(a \cdot b) &= (\widehat{q}a) \cdot (\widehat{q}b), \qquad a, \ b \in C^\infty(M), \\
\widehat{q}(\alpha \cdot a) &= \alpha \cdot \widehat{q}a, \qquad \alpha \in \mathbb{R}, \ a \in C^\infty(M).
\end{aligned}
$$

So to any point $q \in M$, there corresponds a nontrivial homomorphism of algebras $\widehat{q} : C^\infty(M) \to \mathbb{R}$. It turns out that there exists an inverse correspondence.

Proposition 2.1. *Let $\varphi : C^\infty(M) \to \mathbb{R}$ be a nontrivial homomorphism of algebras. Then there exists a point $q \in M$ such that $\varphi = \widehat{q}$.*

We prove this proposition in Appendix A.

Remark. Not only the manifold M can be reconstructed as a set from the algebra $C^\infty(M)$. One can recover topology on M from the weak topology in the space of functionals on $C^\infty(M)$:

$$\lim_{n\to\infty} q_n = q \quad \text{ if and only if } \quad \lim_{n\to\infty} \widehat{q}_n a = \widehat{q} a \quad \forall a \in C^\infty(M).$$

Moreover, the smooth structure on M is also recovered from $C^\infty(M)$, actually, "by definition": a real function on the set $\{\widehat{q} \mid q \in M\}$ is smooth if and only if it has a form $\widehat{q} \mapsto \widehat{q}a$ for some $a \in C^\infty(M)$.

Any diffeomorphism $P : M \to M$ defines an automorphism of the algebra $C^\infty(M)$:

$$\widehat{P} : C^\infty(M) \to C^\infty(M), \qquad \widehat{P} \in \operatorname{Aut}(C^\infty(M)),$$
$$(\widehat{P}a)(q) = a(P(q)), \qquad q \in M, \quad a \in C^\infty(M),$$

i.e., $\widehat{P}$ acts as a change of variables in a function a. Conversely, any automorphism of $C^\infty(M)$ has such a form.

Proposition 2.2. *Any automorphism $A : C^\infty(M) \to C^\infty(M)$ has a form of $\widehat{P}$ for some $P \in \operatorname{Diff} M$.*

Proof. Let $A \in \operatorname{Aut}(C^\infty(M))$. Take any point $q \in M$. Then the composition

$$\widehat{q} \circ A : C^\infty(M) \to \mathbb{R}$$

is a nonzero homomorphism of algebras, thus it has the form $\widehat{q_1}$, $q_1 \in M$. We denote $q_1 = P(q)$ and obtain

$$\widehat{q} \circ A = \widehat{P(q)} = \widehat{q} \circ \widehat{P} \qquad \forall q \in M,$$

i.e.,

$$A = \widehat{P},$$

and P is the required diffeomorphism. □

Now we characterize tangent vectors to M as functionals on $C^\infty(M)$. Tangent vectors to M are velocity vectors to curves in M, and points of M are identified with linear functionals on $C^\infty(M)$; thus we should obtain linear functionals on $C^\infty(M)$, but not homomorphisms into $\mathbb{R}$. To understand, which functionals on $C^\infty(M)$ correspond to tangent vectors to M, take a smooth curve $q(t)$ of points in M. Then the corresponding curve of functionals $\widehat{q}(t) = \widehat{q(t)}$ on $C^\infty(M)$ satisfies the multiplicative rule

$$\widehat{q}(t)(a \cdot b) = \widehat{q}(t)a \cdot \widehat{q}(t)b, \qquad a,\ b \in C^\infty(M).$$

We differentiate this equality at $t = 0$ and obtain that the velocity vector to the curve of functionals

$$\xi \stackrel{\text{def}}{=} \left.\frac{d\,\widehat{q}}{d\,t}\right|_{t=0}, \qquad \xi : C^\infty(M) \to \mathbb{R},$$

satisfies the Leibniz rule:

$$\xi(ab) = \xi(a)b(q(0)) + a(q(0))\xi(b).$$

Consequently, to each tangent vector $v \in T_q M$ we should put into correspondence a linear functional

$$\xi : C^\infty(M) \to \mathbb{R}$$

such that

$$\xi(ab) = (\xi a)b(q) + a(q)(\xi b), \qquad a,\ b \in C^\infty(M). \tag{2.1}$$

But there is a linear functional $\xi = \widehat{v}$ naturally related to any tangent vector $v \in T_q M$, the directional derivative along v:

$$\widehat{v}a = \left.\frac{d}{d\,t}\right|_{t=0} a(q(t)), \qquad q(0) = q, \quad \dot{q}(0) = v,$$

and such functional satisfies Leibniz rule (2.1).

Now we show that this rule characterizes exactly directional derivatives.

Proposition 2.3. *Let $\xi : C^\infty(M) \to \mathbb{R}$ be a linear functional that satisfies Leibniz rule* (2.1) *for some point $q \in M$. Then $\xi = \widehat{v}$ for some tangent vector $v \in T_q M$.*

Proof. Notice first of all that any functional ξ that meets Leibniz rule (2.1) is local, i.e., it depends only on values of functions in an arbitrarily small neighborhood O_q of the point q:

$$\tilde{a}|_{O_q} = a|_{O_q} \quad \Rightarrow \quad \xi\tilde{a} = \xi a, \qquad a,\ \tilde{a} \in C^\infty(M).$$

Indeed, take a cut function $b \in C^\infty(M)$ such that $b|_{M\setminus O_q} \equiv 1$ and $b(q) = 0$. Then $(\tilde{a} - a)b = \tilde{a} - a$, thus

$$\xi(\tilde{a} - a) = \xi((\tilde{a} - a)b) = \xi(\tilde{a} - a)\,b(q) + (\tilde{a} - a)(q)\,\xi b = 0.$$

So the statement of the proposition is local, and we prove it in coordinates.

Let $(x_1, \dots, x_n)$ be local coordinates on M centered at the point q. We have to prove that there exist $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ such that

$$\xi = \sum_{i=1}^{n} \alpha_i \left.\frac{\partial}{\partial\,x_i}\right|_0 .$$

First of all,

$$\xi(1) = \xi(1 \cdot 1) = (\xi 1) \cdot 1 + 1 \cdot (\xi 1) = 2\xi(1),$$

thus $\xi(1) = 0$. By linearity, $\xi(\text{const}) = 0$.

In order to find the action of ξ on an arbitrary smooth function, we expand it by Hadamard Lemma:

$$a(x) = a(0) + \sum_{i=1}^{n} \int_0^1 \frac{\partial a}{\partial x_i}(tx) x_i \, dt = a(0) + \sum_{i=1}^{n} b_i(x) x_i,$$

where

$$b_i(x) = \int_0^1 \frac{\partial a}{\partial x_i}(tx) \, dt$$

are smooth functions. Now

$$\xi a = \sum_{i=1}^{n} \xi(b_i x_i) = \sum_{i=1}^{n} ((\xi b_i) x_i(0) + b_i(0)(\xi x_i)) = \sum_{i=1}^{n} \alpha_i \frac{\partial a}{\partial x_i}(0),$$

where we denote $\alpha_i = \xi x_i$ and make use of the equality $b_i(0) = \dfrac{\partial a}{\partial x_i}(0)$. □

So tangent vectors $v \in T_q M$ can be identified with directional derivatives $\widehat{v} : C^\infty(M) \to \mathbb{R}$, i.e., linear functionals that meet Leibniz rule (2.1).

Now we characterize vector fields on M. A smooth vector field on M is a family of tangent vectors $v_q \in T_q M$, $q \in M$, such that for any $a \in C^\infty(M)$ the mapping $q \mapsto v_q a$, $q \in M$, is a smooth function on M.

To a smooth vector field $V \in \operatorname{Vec} M$ there corresponds a linear operator

$$\widehat{V} : C^\infty(M) \to C^\infty(M)$$

that satisfies the Leibniz rule

$$\widehat{V}(ab) = (\widehat{V}a)b + a(\widehat{V}b), \qquad a,\ b \in C^\infty(M),$$

the directional derivative (Lie derivative) along V.

A linear operator on an algebra meeting the Leibniz rule is called a *derivation* of the algebra, so the Lie derivative $\widehat{V}$ is a derivation of the algebra $C^\infty(M)$. We show that the correspondence between smooth vector fields on M and derivations of the algebra $C^\infty(M)$ is invertible.

Proposition 2.4. *Any derivation of the algebra $C^\infty(M)$ is the directional derivative along some smooth vector field on M.*

Proof. Let $D : C^\infty(M) \to C^\infty(M)$ be a derivation. Take any point $q \in M$. We show that the linear functional

$$d_q \stackrel{\text{def}}{=} \widehat{q} \circ D : C^\infty(M) \to \mathbb{R}$$

is a directional derivative at the point q, i.e., satisfies Leibniz rule (2.1):

$$\begin{aligned} d_q(ab) &= \widehat{q}(D(ab)) = \widehat{q}((Da)b + a(Db)) = \widehat{q}(Da)b(q) + a(q)\widehat{q}(Db) = \\ &(d_q a)b(q) + a(q)(d_q b), \qquad a,\ b \in C^\infty(M). \end{aligned}$$

□

So we can identify points $q \in M$, diffeomorphisms $P \in \operatorname{Diff} M$, and vector fields $V \in \operatorname{Vec} M$ with nontrivial homomorphisms $\widehat{q} : C^\infty(M) \to \mathbb{R}$, automorphisms $\widehat{P} : C^\infty(M) \to C^\infty(M)$, and derivations $\widehat{V} : C^\infty(M) \to C^\infty(M)$ respectively.

For example, we can write a point $P(q)$ in the operator notation as $\widehat{q} \circ \widehat{P}$. Moreover, in the sequel we omit hats and write $q \circ P$. This does not cause ambiguity: if q is to the right of P, then q is a point, P a diffeomorphism, and $P(q)$ is the value of the diffeomorphism P at the point q. And if q is to the left of P, then q is a homomorphism, P an automorphism, and $q \circ P$ a homomorphism of $C^\infty(M)$. Similarly, $V(q) \in T_qM$ is the value of the vector field V at the point q, and $q \circ V : C^\infty(M) \to \mathbb{R}$ is the directional derivative along the vector $V(q)$.

2.2 Seminorms and $C^\infty(M)$-topology

We introduce seminorms and topology on the space $C^\infty(M)$.

By Whitney's Theorem, a smooth manifold M can be properly embedded into a Euclidean space $\mathbb{R}^N$ for sufficiently large N. Denote by h_i, $i = 1, \dots, N$, the smooth vector field on M that is the orthogonal projection from $\mathbb{R}^N$ to M of the constant basis vector field $\frac{\partial}{\partial x_i} \in \operatorname{Vec}(\mathbb{R}^N)$. So we have N vector fields $h_1, \dots, h_N \in \operatorname{Vec} M$ that span the tangent space T_qM at each point $q \in M$.

We define the family of seminorms $\|\cdot\|_{s,K}$ on the space $C^\infty(M)$ in the following way:

$$\|a\|_{s,K} = \sup\left\{ |h_{i_l} \circ \cdots \circ h_{i_1} a(q)| \mid q \in K,\ 1 \le i_1, \dots, i_l \le N,\ 0 \le l \le s \right\},$$
$$a \in C^\infty(M), \quad s \ge 0, \quad K \Subset M.$$

This family of seminorms defines a topology on $C^\infty(M)$. A local base of this topology is given by the subsets

$$\left\{ a \in C^\infty(M) \mid \|a\|_{n,K_n} < \frac{1}{n} \right\}, \qquad n \in \mathbb{N},$$

where K_n, $n \in \mathbb{N}$, is a chained system of compacta that cover M: $K_n \subset K_{n+1}$, $\cup_{n=1}^\infty K_n = M$.

This topology on $C^\infty(M)$ does not depend on embedding of M into $\mathbb{R}^N$. It is called the *topology of uniform convergence of all derivatives on compacta*, or just *$C^\infty(M)$-topology*. This topology turns $C^\infty(M)$ into a Fréchet space (a complete, metrizable, locally convex topological vector space).

A sequence of functions $a_k \in C^\infty(M)$ converges to $a \in C^\infty(M)$ as $k \to \infty$ if and only if

$$\lim_{k\to\infty} \|a_k - a\|_{s,K} = 0 \quad \forall\ s \geq 0,\ K \Subset M.$$

For vector fields $V \in \operatorname{Vec} M$, we define the seminorms

$$\|V\|_{s,K} = \sup \left\{ \|Va\|_{s,K} \mid \|a\|_{s+1,K} = 1 \right\}, \qquad s \geq 0, \quad K \Subset M. \tag{2.2}$$

One can prove that any vector field $V \in \operatorname{Vec} M$ has finite seminorms $\|V\|_{s,K}$, and that there holds an estimate of the action of a diffeomorphism $P \in \operatorname{Diff} M$ on a function $a \in C^\infty(M)$:

$$\|Pa\|_{s,K} \leq C_{s,P} \|a\|_{s,P(K)}, \qquad s \geq 0, \quad K \Subset M. \tag{2.3}$$

Thus vector fields and diffeomorphisms are linear continuous operators on the topological vector space $C^\infty(M)$.

2.3 Families of functionals and operators

In the sequel we will often consider one-parameter families of points, diffeomorphisms, and vector fields that satisfy various regularity properties (e.g. differentiability or absolute continuity) with respect to the parameter. Since we treat points as functionals, and diffeomorphisms and vector fields as operators on $C^\infty(M)$, we can introduce regularity properties for them in the weak sense, via the corresponding properties for one-parameter families of functions

$$t \mapsto a_t, \qquad a_t \in C^\infty(M), \quad t \in \mathbb{R}.$$

So we start from definitions for families of functions.

Continuity and *differentiability* of a family of functions a_t w.r.t. parameter t are defined in a standard way since $C^\infty(M)$ is a topological vector space. A family a_t is called *measurable* w.r.t. t if the real function $t \mapsto a_t(q)$ is measurable for any $q \in M$. A measurable family a_t is called *locally integrable* if

$$\int_{t_0}^{t_1} \|a_t\|_{s,K}\, dt < \infty \qquad \forall\ s \geq 0, \quad K \Subset M, \quad t_0,\ t_1 \in \mathbb{R}.$$

A family a_t is called *absolutely continuous* w.r.t. t if

$$a_t = a_{t_0} + \int_{t_0}^{t} b_\tau\, d\tau$$

for some locally integrable family of functions b_t. A family a_t is called *Lipschitzian* w.r.t. t if

$$\|a_t - a_\tau\|_{s,K} \leq C_{s,K} |t - \tau| \qquad \forall s \geq 0, \quad K \Subset M, \quad t,\ \tau \in \mathbb{R},$$

and *locally bounded* w.r.t. t if

$$\|a_t\|_{s,K} \leq C_{s,K,I}, \qquad \forall\ s \geq 0, \quad K \Subset M, \quad I \Subset \mathbb{R}, \quad t \in I,$$

where $C_{s,K}$ and $C_{s,K,I}$ are some constants depending on s, K, and I.

Now we can define regularity properties of families of functionals and operators on $C^\infty(M)$. A family of linear functionals or linear operators on $C^\infty(M)$

$$t \mapsto A_t, \qquad t \in \mathbb{R},$$

has some property (i.e., is *continuous*, *differentiable*, *measurable*, *locally integrable*, *absolutely continuous*, *Lipschitzian*, *locally bounded* w.r.t. t) if the family

$$t \mapsto A_t a, \qquad t \in \mathbb{R},$$

has the same property for any $a \in C^\infty(M)$.

A locally bounded w.r.t. t family of vector fields

$$t \mapsto V_t, \qquad V_t \in \operatorname{Vec} M, \quad t \in \mathbb{R},$$

is called a *nonautonomous vector field*, or simply a *vector field*, on M. An absolutely continuous w.r.t. t family of diffeomorphisms

$$t \mapsto P^t, \qquad P^t \in \operatorname{Diff} M, \quad t \in \mathbb{R},$$

is called a *flow* on M. So, for a nonautonomous vector field V_t, the family of functions $t \mapsto V_t a$ is locally integrable for any $a \in C^\infty(M)$. Similarly, for a flow P^t, the family of functions $(P^t a)(q) = a(P^t(q))$ is absolutely continuous w.r.t. t for any $a \in C^\infty(M)$.

Integrals of measurable locally integrable families, and derivatives of differentiable families are also defined in the weak sense:

$$\int_{t_0}^{t_1} A_t \, dt \ : a \mapsto \int_{t_0}^{t_1} (A_t a) \, dt, \qquad a \in C^\infty(M),$$

$$\frac{d}{dt} A_t \ : a \mapsto \frac{d}{dt}(A_t a), \qquad a \in C^\infty(M).$$

One can show that if A_t and B_t are continuous families of operators on $C^\infty(M)$ which are differentiable at t_0, then the family $A_t \circ B_t$ is continuous, moreover, differentiable at t_0, and satisfies the Leibniz rule:

$$\left.\frac{d}{dt}\right|_{t_0} (A_t \circ B_t) = \left(\left.\frac{d}{dt}\right|_{t_0} A_t \right) \circ B_{t_0} + A_{t_0} \circ \left(\left.\frac{d}{dt}\right|_{t_0} B_t \right),$$

see the proof in Appendix A.

If families A_t and B_t of operators (functionals) are absolutely continuous, then the composition $A_t \circ B_t$ (resp. the product $A_t B_t$) is absolutely continuous as well. For an absolute continuous family of functions a_t, the family $A_t a_t$ is also absolutely continuous, and the Leibniz rule holds for it as well.

2.4 Chronological exponential

In this section we consider a *nonautonomous ODE* of the form

$$\dot{q} = V_t(q), \qquad q(0) = q_0, \tag{2.4}$$

where V_t is a locally bounded nonautonomous vector field on M, and study the flow determined by this field.

2.4.1 ODEs with discontinuous right-hand side

To obtain local solutions to Cauchy problem (2.4) on a manifold M, we reduce it to a Cauchy problem in a Euclidean space. For details about nonautonomous differential equations in $\mathbb{R}^n$ with right-hand side discontinuous in t, see e.g. [11].

Choose local coordinates $x = (x^1, \dots, x^n)$ in a neighborhood O_{q_0} of the point q_0:

$$\begin{aligned} &\Phi : O_{q_0} \subset M \to O_{x_0} \subset \mathbb{R}^n, \qquad \Phi : q \mapsto x, \\ &\Phi(q_0) = x_0. \end{aligned}$$

In these coordinates, the field V_t reads

$$(\Phi_* V_t)(x) = \widetilde{V}_t(x) = \sum_{i=1}^{n} v_i(t,x) \frac{\partial}{\partial x^i}, \qquad x \in O_{x_0}, \quad t \in \mathbb{R}, \tag{2.5}$$

and problem (2.4) takes the form

$$\dot{x} = \widetilde{V}_t(x), \quad x(0) = x_0, \qquad x \in O_{x_0} \subset \mathbb{R}^n. \tag{2.6}$$

Since the nonautonomous vector field $V_t \in \operatorname{Vec} M$ is locally bounded, the components $v_i(t,x)$, $i = 1, \dots, n$, of its coordinate representation (2.5) are:

(1) measurable and locally bounded w.r.t. t for any fixed $x \in O_{x_0}$,

(2) smooth w.r.t. x for any fixed $t \in \mathbb{R}$,

(3) differentiable in x with locally bounded partial derivatives:

$$\left| \frac{\partial v_i}{\partial x}(t,x) \right| \leq C_{I,K}, \qquad t \in I \Subset \mathbb{R},\ x \in K \Subset O_{x_0},\ i = 1, \dots, n.$$

By the classical Carathéodory Theorem (see e.g. [9]), Cauchy problem (2.6) has a unique solution, i.e., a vector-function $x(t, x_0)$, Lipschitzian w.r.t. t and smooth w.r.t. x_0, and such that:

(1) ODE (2.6) is satisfied for almost all t,

(2) initial condition holds: $x(0, x_0) = x_0$.

Then the pull-back of this solution from $\mathbb{R}^n$ to M

$$q(t, q_0) = \Phi^{-1}(x(t, x_0)),$$

is a solution to problem (2.4) in M. The mapping $q(t, q_0)$ is Lipschitzian w.r.t. t and smooth w.r.t. q_0, it satisfies almost everywhere the ODE and the initial condition in (2.4).

For any $q_0 \in M$, the solution $q(t, q_0)$ to Cauchy problem (2.4) can be continued to a maximal interval $t \in J_{q_0} \subset \mathbb{R}$ containing the origin and depending on q_0.

We will assume that the solutions $q(t, q_0)$ are defined for all $q_0 \in M$ and all $t \in \mathbb{R}$, i.e., $J_{q_0} = \mathbb{R}$ for any $q_0 \in M$. Then the nonautonomous field V_t is called *complete*. This holds, e.g., when all the fields V_t, $t \in \mathbb{R}$, vanish outside of a common compactum in M (in this case we say that the nonautonomous vector field V_t has a *compact support*).

2.4.2 Definition of the right chronological exponential

Equation (2.4) rewritten as a linear equation for Lipschitzian w.r.t. t families of functionals on $C^\infty(M)$:

$$\dot{q}(t) = q(t) \circ V_t, \qquad q(0) = q_0, \tag{2.7}$$

is satisfied for the family of functionals

$$q(t, q_0) \, : \, C^\infty(M) \to \mathbb{R}, \qquad q_0 \in M, \quad t \in \mathbb{R}.$$

We prove later that this Cauchy problem has no other solutions (see Proposition 2.5). Thus the flow defined as

$$P^t \; : \; q_0 \mapsto q(t, q_0) \tag{2.8}$$

is a unique solution of the operator Cauchy problem

$$\dot{P}^t = P^t \circ V_t, \qquad P^0 = \mathrm{Id}, \tag{2.9}$$

(where Id is the identity operator) in the class of Lipschitzian flows on M. The flow P^t determined in (2.8) is called the *right chronological exponential* of the field V_t and is denoted as

$$P^t = \overrightarrow{\exp} \int_0^t V_\tau \, d\tau.$$

Now we develop an asymptotic series for the chronological exponential, which justifies such a notation.

2.4.3 Formal series expansion

We rewrite differential equation in (2.7) as an integral one:

$$q(t) = q_0 + \int_0^t q(\tau) \circ V_\tau \, d\tau \tag{2.10}$$

then substitute this expression for $q(t)$ into the right-hand side

$$
\begin{aligned}
&= q_0 + \int_0^t \left(q_0 + \int_0^{\tau_1} q(\tau_2) \circ V_{\tau_2} \, d\tau_2 \right) \circ V_{\tau_1} \, d\tau_1 \\
&= q_0 \circ \left(\mathrm{Id} + \int_0^t V_\tau \, dt \right) + \iint\limits_{0 \le \tau_2 \le \tau_1 \le t} q(\tau_2) \circ V_{\tau_2} \circ V_{\tau_1} \, d\tau_2 \, d\tau_1 ,
\end{aligned}
$$

repeat this procedure iteratively, and obtain the decomposition:

$$
\begin{aligned}
q(t) = q_0 \circ \Bigg(\mathrm{Id} &+ \int_0^t V_\tau \, d\tau + \iint\limits_{\Delta_2(t)} V_{\tau_2} \circ V_{\tau_1} \, d\tau_2 \, d\tau_1 + \ldots + \\
&\int\limits_{\Delta_n(t)} \cdots \int V_{\tau_n} \circ \cdots \circ V_{\tau_1} \, d\tau_n \, \ldots \, d\tau_1 \Bigg) + \\
&\quad \int\limits_{\Delta_{n+1}(t)} \cdots \int q(\tau_{n+1}) \circ V_{\tau_{n+1}} \circ \cdots \circ V_{\tau_1} \, d\tau_{n+1} \, \ldots \, d\tau_1 . \qquad (2.11)
\end{aligned}
$$

Here

$$\Delta_n(t) = \{ (\tau_1, \ldots, \tau_n) \in \mathbb{R}^n \mid 0 \le \tau_n \le \cdots \le \tau_1 \le t \}$$

is the n-dimensional simplex. Purely formally passing in (2.11) to the limit $n \to \infty$, we obtain a formal series for the solution $q(t)$ to problem (2.7):

$$q_0 \circ \left(\mathrm{Id} + \sum_{n=1}^{\infty} \int\limits_{\Delta_n(t)} \cdots \int V_{\tau_n} \circ \cdots \circ V_{\tau_1} \, d\tau_n \, \ldots \, d\tau_1 \right),$$

thus for the solution P^t to problem (2.9):

$$\mathrm{Id} + \sum_{n=1}^{\infty} \int\limits_{\Delta_n(t)} \cdots \int V_{\tau_n} \circ \cdots \circ V_{\tau_1} \, d\tau_n \, \ldots \, d\tau_1 . \qquad (2.12)$$

Exercise 2.1. We obtained the previous series expansion under the condition $t > 0$, although the chronological exponential is defined for all values of t. Show that the flow $\overrightarrow{\exp} \int_0^t V_\tau \, d\tau$ admits for $t < 0$ the series expansion

$$\mathrm{Id} + \sum_{n=1}^{\infty} \int\limits_{\Delta_n(-t)} \cdots \int (-V_{\tau_n}) \circ \cdots \circ (-V_{\tau_1}) \, d\tau_n \, \ldots \, d\tau_1 .$$

This series is similar to (2.12), so in the sequel we restrict ourselves by the study of the case $t > 0$.

2.4.4 Estimates and convergence of the series

Unfortunately, these series never converge on $C^\infty(M)$ in the weak sense (if $V_t \not\equiv 0$): there always exists a smooth function on M, on which they diverge. Although, one can show that (2.12) is an asymptotic series for the chronological exponential $P^t = \overrightarrow{\exp} \int_0^t V_\tau \, d\tau$. There holds the following bound of the remainder term: denote the m-th partial sum of series (2.12) as

$$S_m(t) = \mathrm{Id} + \sum_{n=1}^{m-1} \int \cdots \int_{\Delta_n(t)} V_{\tau_n} \circ \cdots \circ V_{\tau_1} \, d\tau_n \, \ldots \, d\tau_1,$$

then for any $a \in C^\infty(M)$, $s \geq 0$, $K \Subset M$

$$\begin{aligned} &\left\| \left(\overrightarrow{\exp} \int_0^t V_\tau \, d\tau - S_m(t) \right) a \right\|_{s,K} \\ &\quad \leq C e^{C \int_0^t \|V_\tau\|_{s,K'} \, d\tau} \frac{1}{m!} \left(\int_0^t \|V_\tau\|_{s+m-1,K'} \, d\tau \right)^m \|a\|_{s+m,K'} \qquad (2.13) \\ &\quad = O(t^m), \qquad t \to 0, \end{aligned}$$

where $K' \Subset M$ is some compactum containing K. We prove estimate (2.13) in Appendix, Section A.2. It follows from estimate (2.13) that

$$\left\| \left(\overrightarrow{\exp} \int_0^t \varepsilon V_\tau \, d\tau - S_m^\varepsilon(t) \right) a \right\|_{s,K} = O(\varepsilon^m), \qquad \varepsilon \to 0,$$

where $S_m^\varepsilon(t)$ is the m-th partial sum of series (2.12) for the field εV_t.

Thus we have an asymptotic series expansion:

$$\overrightarrow{\exp} \int_0^t V_\tau \, d\tau \approx \mathrm{Id} + \sum_{n=1}^{\infty} \int \cdots \int_{\Delta_n(t)} V_{\tau_n} \circ \cdots \circ V_{\tau_1} \, d\tau_n \, \ldots \, d\tau_1. \qquad (2.14)$$

In the sequel we will use terms of the zeroth, first, and second orders of the series obtained:

$$\overrightarrow{\exp} \int_0^t V_\tau \, d\tau \approx \mathrm{Id} + \int_0^t V_\tau \, d\tau + \iint_{0 \leq \tau_2 \leq \tau_1 \leq t} V_{\tau_2} \circ V_{\tau_1} \, d\tau_2 \, d\tau_1 + \cdots .$$

We prove that the asymptotic series converges to the chronological exponential on any normed subspace $L \subset C^\infty(M)$ where V_t is well-defined and bounded:

$$V_t L \subset L, \qquad \|V_t\| = \sup\{\|V_t a\| \mid a \in L, \ \|a\| \leq 1\} < \infty. \qquad (2.15)$$

We apply operator series (2.14) to any $a \in L$ and bound terms of the series obtained:

$$a + \sum_{n=1}^{\infty} \int \cdots \int_{\Delta_n(t)} V_{\tau_n} \circ \cdots \circ V_{\tau_1} \, a \, d\tau_n \, \ldots \, d\tau_1. \qquad (2.16)$$

We have

$$\left\| \int_{\Delta_n(t)} \!\!\cdots\!\! \int V_{\tau_n} \circ \cdots \circ V_{\tau_1} \, a \, d\tau_n \, \dots \, d\tau_1 \right\|$$

$$\leq \int_{0 \leq \tau_n \leq \cdots \leq \tau_1 \leq t} \!\!\cdots\!\! \int \|V_{\tau_n}\| \cdot \cdots \cdot \|V_{\tau_1}\| \, d\tau_n \, \dots \, d\tau_1 \cdot \|a\|$$

by symmetry w.r.t. permutations of indices $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$

$$= \int_{0 \leq \tau_{\sigma(n)} \leq \cdots \leq \tau_{\sigma(1)} \leq t} \!\!\cdots\!\! \int \|V_{\tau_n}\| \cdot \cdots \cdot \|V_{\tau_1}\| \, d\tau_n \, \dots \, d\tau_1 \cdot \|a\|$$

passing to the integral over cube

$$= \frac{1}{n!} \int_0^t \cdots \int_0^t \|V_{\tau_n}\| \cdot \cdots \cdot \|V_{\tau_1}\| \, d\tau_n \, \dots \, d\tau_1 \cdot \|a\|$$

$$= \frac{1}{n!} \left(\int_0^t \|V_\tau\| \, d\tau \right)^n \cdot \|a\|.$$

So series (2.16) is majorized by the exponential series, thus the operator series (2.14) converges on L.

Series (2.16) can be differentiated termwise, thus it satisfies the same ODE as the function $P^t a$:

$$\dot{a}_t = V_t a_t, \qquad a_0 = a.$$

Consequently,

$$P^t a = \left(\mathrm{Id} + \sum_{n=1}^{\infty} \int_{\Delta_n(t)} \!\!\cdots\!\! \int V_{\tau_n} \circ \cdots \circ V_{\tau_1} \, d\tau_n \, \dots \, d\tau_1 \right) a.$$

So in the case (2.15) the asymptotic series converges to the chronological exponential and there holds the bound

$$\|P^t a\| \leq e^{\int_0^t \|V_\tau\| \, d\tau} \|a\|, \qquad a \in L.$$

Moreover, one can show that the bound and convergence hold not only for locally bounded, but also for integrable on $[0, t]$ vector fields:

$$\int_0^t \|V_\tau\| \, d\tau < \infty.$$

Notice that conditions (2.15) are satisfied for any finite-dimensional V_t-invariant subspace $L \subset C^\infty(M)$. In particular, this is the case when $M = \mathbb{R}^n$, L is the space of linear vector fields, and V_t is a linear vector field on $\mathbb{R}^n$.

If M and V_t are real analytic, then series (2.14) converges for sufficiently small t (we leave this without proof).

2.4.5 Left chronological exponential

Consider the inverse operator $Q^t = (P^t)^{-1}$ to the right chronological exponential $P^t = \overrightarrow{\exp} \int_0^t V_\tau \, d\tau$. We find an ODE for the flow Q^t by differentiation of the identity

$$P^t \circ Q^t = \mathrm{Id}\,.$$

Leibniz rule yields

$$\dot{P}^t \circ Q^t + P^t \circ \dot{Q}^t = 0,$$

thus, in view of ODE (2.9) for the flow P^t,

$$P^t \circ V_t \circ Q^t + P^t \circ \dot{Q}^t = 0.$$

We multiply this equality by Q^t from the left and obtain

$$V_t \circ Q^t + \dot{Q}^t = 0.$$

That is, the flow Q^t is a solution of the Cauchy problem

$$\frac{d}{dt} Q^t = -V_t \circ Q^t, \qquad Q^0 = \mathrm{Id}, \tag{2.17}$$

which is dual to Cauchy problem (2.9) for P^t. The flow Q^t is called the *left chronological exponential* and is denoted as

$$Q^t = \overleftarrow{\exp} \int_0^t (-V_\tau) \, d\tau.$$

We find an asymptotic expansion for the left chronological exponential in the same way as for the right one, by successive substitutions into the right-hand side:

$$\begin{aligned}
Q^t &= \mathrm{Id} + \int_0^t (-V_\tau) \circ Q^\tau \, d\tau \\
&= \mathrm{Id} + \int_0^t (-V_\tau) \, d\tau + \iint\limits_{\Delta_2(t)} (-V_{\tau_1}) \circ (-V_{\tau_2}) \circ Q^{\tau_2} \, d\tau_2 \, d\tau_1 = \cdots \\
&= \mathrm{Id} + \sum_{n=1}^{m-1} \int\limits_{\Delta_n(t)} \cdots \int (-V_{\tau_1}) \circ \cdots \circ (-V_{\tau_n}) \, d\tau_n \, \ldots \, d\tau_1 \\
&\quad + \int\limits_{\Delta_m(t)} \cdots \int (-V_{\tau_1}) \circ \cdots \circ (-V_{\tau_m}) \circ Q^{\tau_m} \, d\tau_m \, \ldots \, d\tau_1
\end{aligned}$$

For the left chronological exponential holds an estimate of the remainder term as (2.13) for the right one, and the series obtained is asymptotic:

$$\overleftarrow{\exp} \int_0^t (-V_\tau) \, d\tau \approx \mathrm{Id} + \sum_{n=1}^{\infty} \int\limits_{\Delta_n(t)} \cdots \int (-V_{\tau_1}) \circ \cdots \circ (-V_{\tau_n}) \, d\tau_n \, \ldots \, d\tau_1.$$

Remarks. (1) Notice that the reverse arrow in the left chronological exponential $\overleftarrow{\exp}$ corresponds to the reverse order of the operators $(-V_{\tau_1}) \circ \cdots \circ (-V_{\tau_n})$, $\tau_n \leq \ldots \leq \tau_1$.

(2) The right and left chronological exponentials satisfy the corresponding differential equations:

$$\frac{d}{dt} \overrightarrow{\exp} \int_0^t V_\tau \, d\tau = \overrightarrow{\exp} \int_0^t V_\tau \, d\tau \circ V_t,$$
$$\frac{d}{dt} \overleftarrow{\exp} \int_0^t (-V_\tau) \, d\tau = -V_t \circ \overleftarrow{\exp} \int_0^t (-V_\tau) \, d\tau.$$

The directions of arrows correlate with the direction of appearance of operators V_t, $-V_t$ in the right-hand side of these ODEs.

(3) If the initial value is prescribed at a moment of time $t_0 \neq 0$, then the lower limit of integrals in the chronological exponentials is t_0.

(4) There holds the following obvious rule for composition of flows:

$$\overrightarrow{\exp} \int_{t_0}^{t_1} V_\tau \, d\tau \circ \overrightarrow{\exp} \int_{t_1}^{t_2} V_\tau \, d\tau = \overrightarrow{\exp} \int_{t_0}^{t_2} V_\tau \, d\tau.$$

Exercise 2.2. Prove that

$$\overrightarrow{\exp} \int_{t_0}^{t_1} V_\tau \, d\tau = \left(\overrightarrow{\exp} \int_{t_1}^{t_0} V_\tau \, d\tau \right)^{-1} = \overleftarrow{\exp} \int_{t_1}^{t_0} (-V_\tau) \, d\tau. \qquad (2.18)$$

2.4.6 Uniqueness for functional and operator ODEs

We saw that equation (2.7) for Lipschitzian families of functionals has a solution $q(t) = q_0 \circ \overrightarrow{\exp} \int_0^t V_\tau \, d\tau$. We can prove now that this equation has no other solutions.

Proposition 2.5. *Let V_t be a complete locally bounded nonautonomous vector field on M. Then Cauchy problem (2.7) has a unique solution in the class of Lipschitzian families of functionals on $C^\infty(M)$.*

Proof. Let a Lipschitzian family of functionals q_t be a solution to problem (2.7). Then

$$\frac{d}{dt} \left(q_t \circ (P^t)^{-1} \right) = \frac{d}{dt} \left(q_t \circ Q^t \right) = q_t \circ V_t \circ Q^t - q_t \circ V_t \circ Q^t = 0,$$

thus $q_t \circ Q^t \equiv \text{const}$. But $Q_0 = \text{Id}$, consequently, $q_t \circ Q^t \equiv q_0$, hence

$$q_t = q_0 \circ P^t = q_0 \circ \overrightarrow{\exp} \int_0^t V_\tau \, d\tau$$

is a unique solution of Cauchy problem (2.7). □

Similarly, the both operator equations $\dot{P}^t = P^t \circ V_t$ and $\dot{Q}^t = -V_t \circ Q^t$ have no other solutions in addition to the chronological exponentials.

2.4.7 Autonomous vector fields

In the special case of *autonomous vector fields*,

$$V_t \equiv V \in \operatorname{Vec} M,$$

the flow generated by a complete field is called the *exponential* and is denoted as e^{tV}. The asymptotic series for the exponential takes the form

$$e^{tV} \approx \sum_{n=0}^{\infty} \frac{t^n}{n!} V^n = \operatorname{Id} + tV + \frac{t^2}{2} V \circ V + \cdots ,$$

i.e, it is the standard exponential series.

The exponential of an autonomous vector field satisfies the ODEs

$$\frac{d}{dt} e^{tV} = e^{tV} \circ V = V \circ e^{tV}, \qquad e^{tV}\big|_{t=0} = \operatorname{Id}.$$

We apply the asymptotic series for exponential to find the Lie bracket of autonomous vector fields $V, W \in \operatorname{Vec} M$. We compute the first nonconstant term in the asymptotic expansion at $t = 0$ of the curve:

$$\begin{aligned}
q(t) &= q \circ e^{tV} \circ e^{tW} \circ e^{-tV} \circ e^{-tW} \\
&= q \circ \left(\operatorname{Id} + tV + \frac{t^2}{2} V^2 + \cdots \right) \circ \left(\operatorname{Id} + tW + \frac{t^2}{2} W^2 + \cdots \right) \\
&\qquad \circ \left(\operatorname{Id} - tV + \frac{t^2}{2} V^2 + \cdots \right) \circ \left(\operatorname{Id} - tW + \frac{t^2}{2} W^2 + \cdots \right) \\
&= q \circ \left(\operatorname{Id} + t(V + W) + \frac{t^2}{2} (V^2 + 2V \circ W + W^2) + \cdots \right) \\
&\qquad \circ \left(\operatorname{Id} - t(V + W) + \frac{t^2}{2} (V^2 + 2V \circ W + W^2) + \cdots \right) \\
&= q \circ \left(\operatorname{Id} + t^2 (V \circ W - W \circ V) + \cdots \right).
\end{aligned}$$

So the Lie bracket of the vector fields as operators (directional derivatives) in $C^\infty(M)$ is

$$[V, W] = V \circ W - W \circ V.$$

This proves the formula in local coordinates: if

$$V = \sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i}, \qquad W = \sum_{i=1}^{n} b_i \frac{\partial}{\partial x_i}, \qquad a_i,\ b_i \in C^\infty(M),$$

then

$$[V, W] = \sum_{i,j=1}^{n} \left(a_j \frac{\partial b_i}{\partial x_j} - b_j \frac{\partial a_i}{\partial x_j} \right) \frac{\partial}{\partial x_i} = \frac{dW}{dx} V - \frac{dV}{dx} W.$$

Similarly,

$$\begin{aligned} q \circ e^{tV} \circ e^{sW} \circ e^{-tV} &= q \circ (\mathrm{Id} + tV + \cdots) \circ (\mathrm{Id} + sW + \cdots) \circ (\mathrm{Id} - tV + \cdots) \\ &= q \circ (\mathrm{Id} + sW + ts[V, W] + \cdots), \end{aligned}$$

and

$$q \circ [V, W] = \left. \frac{\partial^2}{\partial s \partial t} \right|_{s=t=0} q \circ e^{tV} \circ e^{sW} \circ e^{-tV}.$$

2.5 Action of diffeomorphisms on vector fields

We have already found counterparts to points, diffeomorphisms, and vector fields among functionals and operators on $C^\infty(M)$. Now we consider action of diffeomorphisms on vector fields.

Take a tangent vector $v \in T_q M$ and a diffeomorphism $P \in \mathrm{Diff}\, M$. The tangent vector $P_* v \in T_{P(q)} M$ is the velocity vector of the image of a curve starting from q with the velocity vector v. We claim that

$$P_* v = v \circ P, \qquad v \in T_q M, \quad P \in \mathrm{Diff}\, M, \tag{2.19}$$

as functionals on $C^\infty(M)$. Take a curve

$$q(t) \in M, \qquad q(0) = q, \qquad \left. \frac{d}{dt} \right|_{t=0} q(t) = v,$$

then

$$\begin{aligned} P_* v \, a &= \left. \frac{d}{dt} \right|_{t=0} a(P(q(t))) = \left(\left. \frac{d}{dt} \right|_{t=0} q(t) \right) \circ P a \\ &= v \circ P a, \qquad a \in C^\infty(M). \end{aligned}$$

Now we find expression for $P_* V$, $V \in \mathrm{Vec}\, M$, as a derivation of $C^\infty(M)$. We have

$$\begin{aligned} q \circ P \circ P_* V &= P(q) \circ P_* V = (P_* V)(P(q)) = P_*(V(q)) = V(q) \circ P \\ &= q \circ V \circ P, \quad q \in M, \end{aligned}$$

thus

$$P \circ P_* V = V \circ P,$$

i.e.,

$$P_* V = P^{-1} \circ V \circ P, \qquad P \in \mathrm{Diff}\, M, \ V \in \mathrm{Vec}\, M.$$

So diffeomorphisms act on vector fields as similarities. In particular, diffeomorphisms preserve compositions:

$$P_*(V \circ W) = P^{-1} \circ (V \circ W) \circ P = (P^{-1} \circ V \circ P) \circ (P^{-1} \circ W \circ P) = P_* V \circ P_* W,$$

and Lie brackets of vector fields:

$$P_*[V, W] = P_*(V \circ W - W \circ V) = P_*V \circ P_*W - P_*W \circ P_*V = [P_*V, P_*W].$$

If $B : C^\infty(M) \to C^\infty(M)$ is an automorphism, then the standard algebraic notation for the corresponding similarity is $\mathrm{Ad}\, B$:

$$(\mathrm{Ad}\, B)V \stackrel{\text{def}}{=} B \circ V \circ B^{-1}.$$

That is,

$$P_* = \mathrm{Ad}\, P^{-1}, \qquad P \in \mathrm{Diff}\, M.$$

Now we find an infinitesimal version of the operator Ad. Let P^t be a flow on M,

$$P^0 = \mathrm{Id}, \qquad \left.\frac{d}{dt}\right|_{t=0} P^t = V \in \mathrm{Vec}\, M.$$

Then

$$\left.\frac{d}{dt}\right|_{t=0} (P^t)^{-1} = -V,$$

so

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} (\mathrm{Ad}\, P^t)W &= \left.\frac{d}{dt}\right|_{t=0} (P^t \circ W \circ (P^t)^{-1}) = V \circ W - W \circ V \\ &= [V, W], \qquad W \in \mathrm{Vec}\, M. \end{aligned}$$

Denote

$$\mathrm{ad}\, V = \mathrm{ad}\left(\left.\frac{d}{dt}\right|_{t=0} P^t \right) \stackrel{\text{def}}{=} \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}\, P^t,$$

then

$$(\mathrm{ad}\, V)W = [V, W], \qquad W \in \mathrm{Vec}\, M.$$

Differentiation of the equality

$$\mathrm{Ad}\, P^t\, [X, Y] = [\mathrm{Ad}\, P^t\, X, \mathrm{Ad}\, P^t\, Y] \qquad X, Y \in \mathrm{Vec}\, M,$$

at $t = 0$ gives *Jacobi identity* for Lie bracket of vector fields:

$$(\mathrm{ad}\, V)[X, Y] = [(\mathrm{ad}\, V)X, Y] + [X, (\mathrm{ad}\, V)Y],$$

which may also be written as

$$[V, [X, Y]] = [[V, X], Y] + [X, [V, Y]], \qquad V, X, Y \in \mathrm{Vec}\, M,$$

or, in a symmetric way

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0, \qquad X, Y, Z \in \mathrm{Vec}\, M. \tag{2.20}$$

The set $\mathrm{Vec}\, M$ is a vector space with an additional operation — Lie bracket, which has the properties:

(1) bilinearity:

$$[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z],$$
$$[X, \alpha Y + \beta Z] = \alpha[X, Y] + \beta[X, Z], \qquad X, Y, Z \in \operatorname{Vec} M, \quad \alpha, \beta \in \mathbb{R},$$

(2) skew-symmetry:

$$[X, Y] = -[Y, X], \qquad X, Y \in \operatorname{Vec} M,$$

(3) Jacobi identity (2.20).

In other words, the set $\operatorname{Vec} M$ of all smooth vector fields on a smooth manifold M forms a *Lie algebra*.

Consider the flow $P^t = \overrightarrow{\exp} \int_0^t V_\tau \, d\tau$ of a nonautonomous vector field V_t. We find an ODE for the family of operators $\operatorname{Ad} P^t = (P^t)_*^{-1}$ on the Lie algebra $\operatorname{Vec} M$.

$$\begin{aligned} \frac{d}{dt}(\operatorname{Ad} P^t)X &= \frac{d}{dt}\left(P^t \circ X \circ (P^t)^{-1}\right) \\ &= P^t \circ V_t \circ X \circ (P^t)^{-1} - P^t \circ X \circ V_t \circ (P^t)^{-1} \\ &= (\operatorname{Ad} P^t)[V_t, X] = (\operatorname{Ad} P^t) \operatorname{ad} V_t \, X, \qquad X \in \operatorname{Vec} M. \end{aligned}$$

Thus the family of operators $\operatorname{Ad} P^t$ satisfies the ODE

$$\frac{d}{dt} \operatorname{Ad} P^t = (\operatorname{Ad} P^t) \circ \operatorname{ad} V_t \tag{2.21}$$

with the initial condition

$$\operatorname{Ad} P^0 = \operatorname{Id}. \tag{2.22}$$

So the family $\operatorname{Ad} P^t$ is an invertible solution for the Cauchy problem

$$\dot{A}_t = A_t \circ \operatorname{ad} V_t, \quad A_0 = \operatorname{Id}$$

for operators $A_t : \operatorname{Vec} M \to \operatorname{Vec} M$. We can apply the same argument as for the analogous problem (2.9) for flows to derive the asymptotic expansion

$$\begin{aligned} \operatorname{Ad} P^t \approx \operatorname{Id} + \int_0^t \operatorname{ad} V_\tau \, d\tau + \cdots \\ + \int \cdots \int_{\Delta_n(t)} \operatorname{ad} V_{\tau_n} \circ \cdots \circ \operatorname{ad} V_{\tau_1} \, d\tau_n \ldots d\tau_1 + \cdots \end{aligned} \tag{2.23}$$

then prove uniqueness of the solution, and justify the following notation:

$$\overrightarrow{\exp} \int_0^t \operatorname{ad} V_\tau \, d\tau \stackrel{\text{def}}{=} \operatorname{Ad} P^t = \operatorname{Ad}\left(\overrightarrow{\exp} \int_0^t V_\tau \, d\tau\right).$$

Similar identities for the left chronological exponential are

$$\overleftarrow{\exp} \int_0^t \operatorname{ad}(-V_\tau)\, d\tau \stackrel{\text{def}}{=} \operatorname{Ad}\left(\overleftarrow{\exp} \int_0^t (-V_\tau)\, d\tau\right)$$
$$\approx \operatorname{Id} + \sum_{n=1}^{\infty} \int \cdots \int_{\Delta_n(t)} (-\operatorname{ad} V_{\tau_1}) \circ \cdots \circ (-\operatorname{ad} V_{\tau_n})\, d\tau_n \, \ldots \, d\tau_1 .$$

For the asymptotic series (2.23), there holds an estimate of the remainder term similar to estimate (2.13) for the flow P^t. Denote the partial sum

$$T_m = \operatorname{Id} + \sum_{n=1}^{m-1} \int \cdots \int_{\Delta_n(t)} \operatorname{ad} V_{\tau_n} \circ \cdots \circ \operatorname{ad} V_{\tau_1}\, d\tau_n \, \ldots \, d\tau_1 ,$$

then for any $X \in \operatorname{Vec} M$, $s \geq 0$, $K \Subset M$

$$\begin{aligned} &\left\| \left(\operatorname{Ad} \overrightarrow{\exp} \int_0^t V_\tau \, d\tau - T_m \right) X \right\|_{s,K} \\ &\leq C_1 e^{C_1 \int_0^t \|V_\tau\|_{s+1,K'}\, d\tau} \frac{1}{m!} \left(\int_0^t \|V_\tau\|_{s+m,K'}\, d\tau \right)^m \|X\|_{s+m,K'} \qquad (2.24) \\ &= O(t^m), \qquad t \to 0, \end{aligned}$$

where $K' \Subset M$ is some compactum containing K.

For autonomous vector fields, we denote

$$e^{t \operatorname{ad} V} \stackrel{\text{def}}{=} \operatorname{Ad} e^{tV},$$

thus the family of operators $e^{t \operatorname{ad} V} : \operatorname{Vec} M \to \operatorname{Vec} M$ is the unique solution to the problem

$$\dot{A}_t = A_t \circ \operatorname{ad} V, \qquad A_0 = \operatorname{Id},$$

which admits the asymptotic expansion

$$e^{t \operatorname{ad} V} \approx \operatorname{Id} + t \operatorname{ad} V + \frac{t^2}{2} \operatorname{ad}^2 V + \cdots .$$

Exercise 2.3. Let $P \in \operatorname{Diff} M$, and let V_t be a nonautonomous vector field on M. Prove that

$$P \circ \overrightarrow{\exp} \int_0^t V_\tau \, d\tau \circ P^{-1} = \overrightarrow{\exp} \int_0^t (\operatorname{Ad} P\, V_\tau)\, d\tau. \qquad (2.25)$$

2.6 Commutation of flows

Let $V_t \in \operatorname{Vec} M$ be a nonautonomous vector field and $P^t = \overrightarrow{\exp} \int_0^t V_\tau \, d\tau$ the corresponding flow. We are interested in the question: under what conditions the flow P^t preserves a vector field $W \in \operatorname{Vec} M$:

$$P^t{}_* W = W \qquad \forall t,$$

or, which is equivalent,

$$(P^t)_*^{-1} W = W \qquad \forall t.$$

Proposition 2.6.

$$P_*^t W = W \quad \forall t \quad \Leftrightarrow \quad [V_t, W] = 0 \quad \forall t.$$

Proof. We have

$$\begin{aligned}
\frac{d}{dt}(P_t)_*^{-1} W &= \frac{d}{dt} \operatorname{Ad} P^t W = \left(\frac{d}{dt} \ \overrightarrow{\exp} \int_0^t \operatorname{ad} V_\tau \, d\tau \right) W \\
&= \left(\overrightarrow{\exp} \int_0^t \operatorname{ad} V_\tau \, d\tau \circ \operatorname{ad} V_\tau \right) W = \left(\overrightarrow{\exp} \int_0^t \operatorname{ad} V_\tau \, d\tau \right) [V_t, W] \\
&= (P^t)_*^{-1} [V_t, W],
\end{aligned}$$

thus $(P^t)_*^{-1} W \equiv W$ if and only if $[V_t, W] \equiv 0$. □

In general, flows do not commute, neither for nonautonomous vector fields V_t, W_t:

$$\overrightarrow{\exp} \int_0^{t_1} V_\tau \, d\tau \circ \overrightarrow{\exp} \int_0^{t_2} W_\tau \, d\tau \neq \overrightarrow{\exp} \int_0^{t_2} W_\tau \, d\tau \circ \overrightarrow{\exp} \int_0^{t_1} V_\tau \, d\tau,$$

nor for autonomous vector fields V, W:

$$e^{t_1 V} \circ e^{t_2 W} \neq e^{t_2 W} \circ e^{t_1 V}.$$

In the autonomous case, commutativity of flows is equivalent to commutativity of vector fields:

$$e^{t_1 V} \circ e^{t_2 W} = e^{t_2 W} \circ e^{t_1 V}, \quad t_1, t_2 \in \mathbb{R}, \qquad \Leftrightarrow \qquad [V, W] = 0.$$

We already showed that commutativity of vector fields is necessary for commutativity of flows. Let us prove that it is sufficient. Indeed,

$$\left(\operatorname{Ad} e^{t_1 V}\right) W = e^{t_1 \operatorname{ad} V} W = W.$$

Taking into account equality (2.25), we obtain

$$e^{t_1 V} \circ e^{t_2 W} \circ e^{-t_1 V} = e^{t_2 \left(\operatorname{Ad} e^{t_1 V}\right) W} = e^{t_2 W}.$$

2.7 Variations formula

Consider an ODE of the form

$$\dot{q} = V_t(q) + W_t(q). \tag{2.26}$$

We think of V_t as an initial vector field and W_t as its perturbation. Our aim is to find a formula for the flow Q^t of the new field $V_t + W_t$ as a perturbation of the flow $P^t = \overrightarrow{\exp} \int_0^t V_\tau \, d\tau$ of the initial field V_t. In other words, we wish to have a decomposition of the form

$$Q^t = \overrightarrow{\exp} \int_0^t (V_\tau + W_\tau) \, d\tau = C_t \circ P^t.$$

We proceed as in the method of variation of parameters; we substitute the previous expression to ODE (2.26):

$$\begin{aligned} \frac{d}{dt} Q^t &= Q^t \circ (V_t + W_t) \\ &= \dot{C}_t \circ P^t + C_t \circ P^t \circ V_t \\ &= \dot{C}_t \circ P^t + Q^t \circ V_t, \end{aligned}$$

cancel the common term $Q^t \circ V_t$:

$$Q^t \circ W_t = \dot{C}_t \circ P^t,$$

and write down the ODE for the unknown flow C_t:

$$\begin{aligned} \dot{C}_t &= Q^t \circ W_t \circ \left(P^t\right)^{-1} \\ &= C_t \circ P^t \circ W_t \circ \left(P^t\right)^{-1} \\ &= C_t \circ \left(\operatorname{Ad} P^t\right) W_t \\ &= C_t \circ \left(\overrightarrow{\exp} \int_0^t \operatorname{ad} V_\tau \, d\tau \right) W_t, \\ C_0 &= \operatorname{Id}. \end{aligned}$$

This operator Cauchy problem is of the form (2.9), thus it has a unique solution:

$$C_t = \overrightarrow{\exp} \int_0^t \left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} V_\theta \, d\theta \right) W_\tau \, d\tau.$$

Hence we obtain the required decomposition of the perturbed flow:

$$\overrightarrow{\exp} \int_0^t (V_\tau + W_\tau) \, d\tau = \overrightarrow{\exp} \int_0^t \left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} V_\theta \, d\theta \right) W_\tau \, d\tau \circ \overrightarrow{\exp} \int_0^t V_\tau \, d\tau. \tag{2.27}$$

This equality is called the *variations formula*. It can be written as follows:

$$\overrightarrow{\exp} \int_0^t (V_\tau + W_\tau) \, d\tau = \overrightarrow{\exp} \int_0^t \left(\operatorname{Ad} P^\tau\right) W_\tau \, d\tau \circ P^t.$$

So the perturbed flow is a composition of the initial flow P^t with the flow of the perturbation W_t twisted by P^t.

Now we obtain another form of the variations formula, with the flow P^t to the left of the twisted flow. We have

$$\begin{aligned}
\overrightarrow{\exp} \int_0^t (V_\tau + W_\tau)\, d\tau &= \overrightarrow{\exp} \int_0^t (\operatorname{Ad} P^\tau)\, W_\tau\, d\tau \circ P^t \\
&= P^t \circ \left(P^t\right)^{-1} \circ \overrightarrow{\exp} \int_0^t (\operatorname{Ad} P^\tau)\, W_\tau\, d\tau \circ P^t \\
&= P^t \circ \overrightarrow{\exp} \int_0^t \left(\operatorname{Ad} \left(P^t\right)^{-1} \circ \operatorname{Ad} P^\tau\right) W_\tau\, d\tau \\
&= P^t \circ \overrightarrow{\exp} \int_0^t \left(\operatorname{Ad} \left(\left(P^t\right)^{-1} \circ P^\tau\right)\right) W_\tau\, d\tau.
\end{aligned}$$

Since

$$\left(P^t\right)^{-1} \circ P^\tau = \overrightarrow{\exp} \int_t^\tau V_\theta\, d\theta,$$

we obtain

$$\begin{aligned}
\overrightarrow{\exp} \int_0^t (V_\tau + W_\tau)\, d\tau &= P^t \circ \overrightarrow{\exp} \int_0^t \left(\overrightarrow{\exp} \int_t^\tau \operatorname{ad} V_\theta\, d\theta\right) W_\tau\, d\tau \\
&= \overrightarrow{\exp} \int_0^t V_\tau\, d\tau \circ \overrightarrow{\exp} \int_0^t \left(\overrightarrow{\exp} \int_t^\tau \operatorname{ad} V_\theta\, d\theta\right) W_\tau\, d\tau.
\end{aligned} \tag{2.28}$$

For autonomous vector fields $V, W \in \operatorname{Vec} M$, the variations formulas (2.27), (2.28) take the form:

$$e^{t(V+W)} = \overrightarrow{\exp} \int_0^t e^{\tau \operatorname{ad} V} W\, d\tau \circ e^{tV} = e^{tV} \circ \overrightarrow{\exp} \int_0^t e^{(\tau - t) \operatorname{ad} V} W\, d\tau. \tag{2.29}$$

In particular, for $t = 1$ we have

$$e^{V+W} = \overrightarrow{\exp} \int_0^1 e^{\tau \operatorname{ad} V} W\, d\tau \circ e^V.$$

2.8 Derivative of flow with respect to parameter

Let $V_t(s)$ be a nonautonomous vector field depending smoothly on a real parameter s. We study dependence of the flow of $V_t(s)$ on the parameter s.

We write

$$\overrightarrow{\exp} \int_0^t V_\tau(s + \varepsilon)\, d\tau = \overrightarrow{\exp} \int_0^t \left(V_\tau(s) + \delta_{V_\tau}(s, \varepsilon)\right)\, d\tau \tag{2.30}$$

with the perturbation $\delta_{V_\tau}(s, \varepsilon) = V_\tau(s + \varepsilon) - V_\tau(s)$. By the variations formula (2.27), the previous flow is equal to

$$\overrightarrow{\exp} \int_0^t \left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} V_\theta(s)\, d\theta\right) \delta_{V_\tau}(s, \varepsilon)\, d\tau \circ \overrightarrow{\exp} \int_0^t V_\tau(s)\, d\tau.$$

Now we expand in ε:

$$\begin{aligned}
\delta_{V_\tau}(s,\varepsilon) &= \varepsilon \frac{\partial}{\partial s} V_\tau(s) + O(\varepsilon^2), \qquad \varepsilon \to 0, \\
W_\tau(s,\varepsilon) &\stackrel{\text{def}}{=} \left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} V_\theta(s)\, d\theta \right) \delta_{V_\tau}(s,\varepsilon) \\
&= \varepsilon \left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} V_\theta(s)\, d\theta \right) \frac{\partial}{\partial s} V_\tau(s) + O(\varepsilon^2), \qquad \varepsilon \to 0,
\end{aligned}$$

thus

$$\begin{aligned}
\overrightarrow{\exp} \int_0^t W_\tau(s,\varepsilon)\, d\tau &= \operatorname{Id} + \int_0^t W_\tau(s,\varepsilon)\, d\tau + O(\varepsilon^2) \\
&= \operatorname{Id} + \varepsilon \int_0^t \left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} V_\theta(s)\, d\theta \right) \frac{\partial}{\partial s} V_\tau(s)\, d\tau + O(\varepsilon^2).
\end{aligned}$$

Finally,

$$\begin{aligned}
&\overrightarrow{\exp} \int_0^t V_\tau(s+\varepsilon)\, d\tau = \overrightarrow{\exp} \int_0^t W_{s,\tau}(\varepsilon)\, d\tau \circ \overrightarrow{\exp} \int_0^t V_\tau(s)\, d\tau \\
&\quad = \overrightarrow{\exp} \int_0^t V_\tau(s)\, d\tau \\
&\qquad + \varepsilon \int_0^t \left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} V_\theta(s)\, d\theta \right) \frac{\partial}{\partial s} V_\tau(s)\, d\tau \circ \overrightarrow{\exp} \int_0^t V_\tau(s)\, d\tau + O(\varepsilon^2),
\end{aligned}$$

that is,

$$\begin{aligned}
&\frac{\partial}{\partial s} \overrightarrow{\exp} \int_0^t V_\tau(s)\, d\tau \\
&\qquad = \int_0^t \left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} V_\theta(s)\, d\theta \right) \frac{\partial}{\partial s} V_\tau(s)\, d\tau \circ \overrightarrow{\exp} \int_0^t V_\tau(s)\, d\tau. \qquad (2.31)
\end{aligned}$$

Similarly, we obtain from the variations formula (2.28) the equality

$$\begin{aligned}
&\frac{\partial}{\partial s} \overrightarrow{\exp} \int_0^t V_\tau(s)\, d\tau \\
&\qquad = \overrightarrow{\exp} \int_0^t V_\tau(s)\, d\tau \circ \int_0^t \left(\overrightarrow{\exp} \int_t^\tau \operatorname{ad} V_\theta(s)\, d\theta \right) \frac{\partial}{\partial s} V_\tau(s)\, d\tau. \qquad (2.32)
\end{aligned}$$

For an autonomous vector field depending on a parameter $V(s)$, formula (2.31) takes the form

$$\frac{\partial}{\partial s} e^{tV(s)} = \int_0^t e^{\tau \operatorname{ad} V(s)} \frac{\partial V}{\partial s}\, d\tau \circ e^{tV(s)},$$

and at $t = 1$:

$$\frac{\partial}{\partial s} e^{V(s)} = \int_0^1 e^{\tau \operatorname{ad} V(s)} \frac{\partial V}{\partial s}\, d\tau \circ e^{V(s)}. \qquad (2.33)$$

Proposition 2.7. *Assume that*

$$\left[\int_0^t V_\tau \, d\tau, V_t\right] = 0 \qquad \forall t. \tag{2.34}$$

Then

$$\overrightarrow{\exp} \int_0^t V_\tau \, d\tau = e^{\int_0^t V_\tau \, d\tau} \qquad \forall t.$$

That is, we state that under the commutativity assumption (2.34), the chronological exponential $\overrightarrow{\exp} \int_0^t V_\tau \, d\tau$ coincides with the flow $Q^t = e^{\int_0^t V_\tau \, d\tau}$ defined as follows:

$$Q^t = Q_1^t,$$
$$\frac{\partial Q_s^t}{\partial s} = \int_0^t V_\tau \, d\tau \circ Q_s^t, \qquad Q_0^t = \mathrm{Id}\,.$$

Proof. We show that the exponential in the right-hand side satisfies the same ODE as the chronological exponential in the left-hand side. By (2.33), we have

$$\frac{d}{dt} e^{\int_0^t V_\tau \, d\tau} = \int_0^1 e^{\tau \operatorname{ad} \int_0^t V_\theta \, d\theta} \, V_t \, d\tau \circ e^{\int_0^t V_\tau \, d\tau}.$$

In view of equality (2.34),

$$e^{\tau \operatorname{ad} \int_0^t V_\theta \, d\theta} \, V_t = V_t,$$

thus

$$\frac{d}{dt} e^{\int_0^t V_\tau \, d\tau} = V_t \circ e^{\int_0^t V_\tau \, d\tau}.$$

By equality (2.34), we can permute operators in the right-hand side:

$$\frac{d}{dt} e^{\int_0^t V_\tau \, d\tau} = e^{\int_0^t V_\tau \, d\tau} \circ V_t.$$

Notice the initial condition

$$\left. e^{\int_0^t V_\tau \, d\tau} \right|_{t=0} = \mathrm{Id}\,.$$

Now the statement follows since the Cauchy problem for flows

$$\dot{A}_t = A_t \circ V_t, \qquad A_0 = \mathrm{Id}$$

has a unique solution:

$$A_t = e^{\int_0^t V_\tau \, d\tau} = \overrightarrow{\exp} \int_0^t V_\tau \, d\tau.$$

□

Chapter 3

Linear systems

Linear control systems have the form

$$\dot{x} = Ax + c + \sum_{i=1}^{m} u_i b_i, \qquad x \in \mathbb{R}^n, \quad u = (u_1, \ldots, u_m) \in \mathbb{R}^m, \tag{3.1}$$

where A is a constant real $n \times n$ matrix and c, b_1, ..., b_m are constant vectors in $\mathbb{R}^n$.

3.1 Cauchy's formula for linear systems

Let $u(t) = (u_1(t), \ldots, u_m(t))$ be locally integrable functions. Then the solution of (3.1) corresponding to this control and satisfying the initial condition

$$x(0, x_0) = x_0$$

is given by *Cauchy's formula*:

$$x(t, x_0) = e^{tA}\left(x_0 + \int_0^t e^{-\tau A}\left(\sum_{i=1}^{m} u_i(\tau) b_i + c\, d\tau\right)\right), \quad t \in \mathbb{R}.$$

Here we use the standard notation for the matrix exponential:

$$e^{tA} = \mathrm{Id} + tA + \frac{t^2}{2!}A^2 + \cdots + \frac{t^n}{n!}A^n + \cdots .$$

Cauchy's formula is verified by differentiation. In view of uniqueness, it gives the solution to the Cauchy problem.

Linear system (3.1) is a particular case of an *affine in control system*:

$$\dot{x} = x \circ \left(f_0 + \sum_{i=1}^{m} u_i f_i\right), \tag{3.2}$$

in order to obtain (3.1) from (3.2), one should just take

$$f_0(x) = Ax + c, \qquad f_i(x) = b_i, \quad i = 1, \dots, m. \tag{3.3}$$

Let us show that Cauchy's formula is actually a special case of the general variations formula. We restrict ourselves with the case $c = 0$.

The variations formula for (3.2) takes the form

$$\begin{aligned}
&\overrightarrow{\exp} \int_0^t \left(f_0 + \sum_{i=1}^m u_i(\tau) f_i \right) d\tau \\
&= \overrightarrow{\exp} \int_0^t \left(\left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} f_0 \, d\theta \right) \circ \sum_{i=1}^m u_i(\tau) f_i \right) d\tau \circ \overrightarrow{\exp} \int_0^t f_0 \, d\tau \\
&= \overrightarrow{\exp} \int_0^t \left(\sum_{i=1}^m u_i(\tau) e^{\tau \operatorname{ad} f_0} f_i \right) d\tau \circ e^{t f_0}.
\end{aligned} \tag{3.4}$$

We assume that $c = 0$, i.e., $f_0(x) = Ax$. Then

$$x \circ e^{t f_0} = e^{tA} x. \tag{3.5}$$

Further, since $(\operatorname{ad} f_0) f_i = [f_0, f_i] = [Ax, b] = -Ab$ then

$$\begin{aligned}
e^{\tau \operatorname{ad} f_0} f_i &= f_i + \tau (\operatorname{ad} f_0) f_i + \frac{\tau^2}{2!} (\operatorname{ad} f_0)^2 f_i + \cdots + \frac{\tau^n}{n!} (\operatorname{ad} f_0)^n f_i + \cdots \\
&= b_i - \tau A b_i + \frac{\tau^2}{2!} (-A)^2 b_i + \cdots + \frac{\tau^n}{n!} (-A)^n b_i + \cdots \\
&= e^{-\tau A} b_i.
\end{aligned}$$

In order to compute the left flow in (3.4), recall that the curve

$$x_0 \circ \overrightarrow{\exp} \int_0^t \left(\sum_{i=1}^m u_i(\tau) e^{\tau \operatorname{ad} f_0} f_i \right) d\tau = x_0 \circ \overrightarrow{\exp} \int_0^t \left(\sum_{i=1}^m u_i(\tau) e^{-\tau A} b_i \right) d\tau \tag{3.6}$$

is the solution to the Cauchy problem

$$\dot{x}(t) = \sum_{i=1}^m u_i(t) e^{-tA} b_i, \qquad x(0) = x_0,$$

thus (3.6) is equal to

$$x(t) = x_0 + \int_0^t \left(e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i \right) d\tau.$$

Taking into account (3.5), we obtain Cauchy's formula:

$$
\begin{aligned}
x(t) &= x_0 \circ \overrightarrow{\exp} \int_0^t \left(f_0 + \sum_{i=1}^m u_i(\tau) f_i \right) d\tau \\
&= \left(x_0 + \int_0^t \left(e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i \right) d\tau \right) \circ e^{t f_0} \\
&= e^{tA} \left(x_0 + \int_0^t \left(e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i \right) d\tau \right).
\end{aligned}
$$

Notice that in the general case ($c \neq 0$) Cauchy's formula can be written as follows:

$$
\begin{aligned}
x(t, x_0) &= e^{tA} x_0 + e^{tA} \int_0^t e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i \, d\tau + e^{tA} \int_0^t e^{-\tau A} c \, d\tau \\
&= e^{tA} x_0 + e^{tA} \int_0^t e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i \, d\tau + \frac{e^{tA} - \mathrm{Id}}{A} c, \qquad (3.7)
\end{aligned}
$$

where

$$
\frac{e^{tA} - \mathrm{Id}}{A} c = tc + \frac{t^2}{2!} Ac + \frac{t^3}{3!} A^2 c + \cdots + \frac{t^n}{n!} A^{n-1} c + \cdots .
$$

3.2 Controllability of linear systems

Cauchy's formula (3.7) yields that the mapping

$$
u(\cdot) \mapsto x(t, x_0, u(\cdot)),
$$

which sends a locally integrable control to the endpoint of the corresponding trajectory, is affine. Thus the attainable set $\mathcal{A}_{x_0}(t)$ of linear system (3.1) for a fixed time $t > 0$ is an affine subspace in $\mathbb{R}^n$.

Definition 3.1. A control system on a state space M is called *completely controllable* for time $t > 0$ if

$$
\mathcal{A}_{x_0}(t) = M \qquad \forall x_0 \in M.
$$

This definition means that for any pair of points $x_0, x_1 \in M$ exists an admissible control $u(\cdot)$ such that the corresponding solution $x(\cdot, x_0, u(\cdot))$ of the control system steers x_0 to x_1 in t units of time:

$$
x(0, x_0, u(\cdot)) = x_0, \quad x(t, x_0, u(\cdot)) = x_1.
$$

The study of complete controllability of linear systems is facilitated by the following observation. The affine mapping

$$
u(\cdot) \mapsto e^{tA} x_0 + \frac{e^{tA} - \mathrm{Id}}{A} c + e^{tA} \int_0^t e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i \, d\tau
$$

is surjective if and only if its linear part

$$u(\cdot) \mapsto e^{tA} \int_0^t e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i \, d\tau \tag{3.8}$$

is onto. Moreover, (3.8) is surjective iff the following mapping is:

$$u(\cdot) \mapsto \int_0^t e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i \, d\tau. \tag{3.9}$$

Theorem 3.1. *The linear system* (3.1) *is completely controllable for a time* $t > 0$ *if and only if*

$$\operatorname{span}\{A^j b_i \mid j = 0, \dots, n-1, \ i = 1, \dots, m\} = \mathbb{R}^n. \tag{3.10}$$

Proof. Necessity. Assume, by contradiction, that condition (3.10) is violated. Then there exists a covector $p \in \mathbb{R}^{n*}$, $p \neq 0$, such that

$$pA^j b_i = 0, \qquad j = 0, \dots, n-1, \ i = 1, \dots, m. \tag{3.11}$$

By Cayley theorem,

$$A^n = \sum_{j=0}^{n-1} \alpha_j A^j$$

for some real numbers $\alpha_0, \dots, \alpha_{n-1}$, thus

$$A^k = \sum_{j=0}^{n-1} \beta_j^k A^j$$

for any $k \in \mathbb{N}$ and some $\beta_j^k \in \mathbb{R}$. Now we obtain from (3.11):

$$pA^k b_i = \sum_{j=0}^{n-1} \beta_j^k pA^j b_i = 0, \qquad k = 0, 1, \dots, \ i = 1, \dots, m.$$

That is why

$$pe^{-\tau A} b_i = 0, \qquad i = 1, \dots, m,$$

and finally

$$p \int_0^t e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i \, d\tau = \int_0^t \sum_{i=1}^m u_i(\tau) pe^{-\tau A} b_i \, d\tau = 0,$$

i.e., mapping (3.9) is not surjective. The contradiction proves necessity.

Sufficiency. By contradiction, suppose that mapping (3.9) is not surjective. Then there exists a covector $p \in \mathbb{R}^{n*}$, $p \neq 0$, such that

$$p \int_0^t \sum_{i=1}^m u_i(\tau) e^{-\tau A} b_i \, d\tau = 0 \qquad \forall u(\cdot) = (u_1(\cdot), \dots, u_m(\cdot)). \tag{3.12}$$

Choose a control of the form:

$$u(\tau) = (0, \dots, 0, v_s(\tau), 0, \dots, 0),$$

where the only nonzero i-th component is

$$v_s(\tau) = \begin{cases} 1, & 0 \le \tau \le s, \\ 0, & \tau > s. \end{cases}$$

Then equality (3.12) gives

$$p \int_0^s e^{-\tau A} b_i \, d\tau = 0, \qquad s \in \mathbb{R}, \ i = 1, \dots, m,$$

thus

$$pe^{-sA} b_i = 0, \qquad s \in \mathbb{R}, \ i = 1, \dots, m.$$

We differentiate this equality repeatedly at $s = 0$ and obtain

$$pA^k b_i = 0, \qquad k = 0, 1, \dots, \ i = 1, \dots, m,$$

a contradiction with (3.10). Sufficiency follows. □

Chapter 4

State linearizability of nonlinear systems

The aim of this chapter is to characterize nonlinear systems

$$\dot{q} = f_0(q) + \sum_{i=1}^{m} u_i f_i(q), \qquad u = (u_1, \ldots, u_m) \in \mathbb{R}^m, \ q \in M \tag{4.1}$$

that are equivalent, locally or globally, to controllable linear systems. That is, we seek conditions on vector fields $f_0, f_1, \ldots, f_m$ that guarantee existence of a diffeomorphism (global $\Phi : M \to \mathbb{R}^n$ or local $\Phi : O_{q_0} \subset M \to O_0 \subset \mathbb{R}^n$) which transforms nonlinear system (4.1) into a controllable linear one (3.1).

4.1 Local linearizability

We start with the local problem. A natural language for conditions of local linearizability is provided by Lie brackets, which are invariant under diffeomorphisms:

$$\Phi_*[V, W] = [\Phi_* V, \Phi_* W], \qquad V, W \in \operatorname{Vec} M.$$

The controllability condition (3.10) can easily be rewritten in terms of Lie brackets: since

$$(-A)^j b_i = (\operatorname{ad} f_0)^j f_i = [\underbrace{f_0, [\ldots [f_0,}_{j \text{ times}} f_i] \ldots]]$$

for vector fields (3.3), then the controllability test for linear systems (3.10) reads

$$\operatorname{span}\{x_0 \circ (\operatorname{ad} f_0)^j f_i \mid j = 0, \ldots, n-1, \ i = 1, \ldots, m\} = T_{x_0}\mathbb{R}^n.$$

Further, one can see that the following equality is satisfied for linear vector fields (3.3):

$$[(\operatorname{ad} f_0)^{j_1} f_{i_1}, (\operatorname{ad} f_0)^{j_2} f_{i_2}] = [(-A)^{j_1} b_{i_1}, (-A)^{j_2} b_{i_2}] = 0,$$
$$0 \le j_1, \ j_2, \quad 1 \le i_1, \ i_2 \le m.$$

It turns out that the two conditions found above give a precise local characterization of controllable linear systems.

Theorem 4.1. *Let M be a smooth n-dimensional manifold, and let $f_0, f_1, \dots, f_m \in \operatorname{Vec} M$. There exists a diffeomorphism*

$$\Phi : O_{q_0} \to O_0$$

of a neighborhood $O_{q_0} \subset M$ of a point $q_0 \in M$ to a neighborhood $O_0 \subset \mathbb{R}^n$ of the origin $0 \in \mathbb{R}^n$ such that

$$\begin{aligned}(\Phi_* f_0)(x) &= Ax + c, \qquad x \in O_0,\\ (\Phi_* f_i)(x) &= b_i, \qquad x \in O_0, \quad i = 1, \dots, m,\end{aligned}$$

for some $n \times n$ matrix A and $c, b_1, \dots, b_m \in \mathbb{R}^n$ that satisfy the controllability condition (3.10) if and only if the following conditions hold:

$$\operatorname{span}\{q_0 \circ (\operatorname{ad} f_0)^j f_i \mid j = 0, \dots, n-1, \ i = 1, \dots, m\} = T_{q_0} M. \tag{4.2}$$

$$q \circ [(\operatorname{ad} f_0)^{j_1} f_{i_1}, (\operatorname{ad} f_0)^{j_2} f_{i_2}] = 0, \tag{4.3}$$

$$q \in O_{q_0}, \quad 0 \le j_1, \ j_2 \le n, \quad 1 \le i_1, \ i_2 \le m.$$

Remark. In other words, the diffeomorphism Φ from the theorem transforms a nonlinear system (4.1) to a linear one (3.1).

Before proving the theorem, we consider the following proposition, which we will need later.

Lemma 4.1. *Let M be a smooth n-dimensional manifold, and let $Y_1, \dots, Y_k \in \operatorname{Vec} M$. There exists a diffeomorphism*

$$\Phi : O_0 \to O_{q_0}$$

of a neighborhood $O_0 \subset \mathbb{R}^n$ to a neighborhood $O_{q_0} \subset M$, $q_0 \in M$, such that

$$\Phi_* \left(\frac{\partial}{\partial x_i} \right) = Y_i, \qquad i = 1, \dots, k,$$

if and only if the vector fields $Y_1, \dots, Y_k$ commute:

$$[Y_i, Y_j] \equiv 0, \qquad i, \ j = 1, \dots, k,$$

and are linearly independent:

$$\dim \operatorname{span}(q_0 \circ Y_1, \dots, q_0 \circ Y_k) = k.$$

Proof. Necessity is obvious since Lie bracket and linear independence are invariant with respect to diffeomorphisms.

Sufficiency. Choose $Y_{k+1}, \dots, Y_n \in \operatorname{Vec} M$ that complete $Y_1, \dots, Y_k$ to a basis:

$$\operatorname{span}(q \circ Y_1, \dots, q \circ Y_n) = T_q M, \qquad q \in O_{q_0}.$$

The mapping
$$\Phi(s_1, \dots, s_n) = q_0 \circ e^{s_n Y_n} \circ \dots \circ e^{s_1 Y_1}$$
is defined on a sufficiently small neighborhood of the origin in $\mathbb{R}^n$. We have
$$\left.\frac{\partial}{\partial s_i}\right|_{s=0} \Phi(s) \stackrel{\text{def}}{=} \Phi_* \left(\left.\frac{\partial}{\partial s_i}\right|_{s=0} \right) = \left.\frac{\partial}{\partial \varepsilon}\right|_{\varepsilon=0} q_0 \circ e^{\varepsilon Y_i} = q_0 \circ Y_i.$$
Hence $\Phi_*|_{s=0}$ is surjective and Φ is a diffeomorphism of a neighborhood of 0 in $\mathbb{R}^n$ and a neighborhood of q_0 in M, according to the implicit function theorem.

Now we prove that Φ rectifies the vector fields $Y_1, \dots, Y_k$. First of all, notice that since these vector fields commute, then their flows also commute, thus
$$e^{s_k Y_k} \circ \dots \circ e^{s_1 Y_1} = e^{\sum_{i=1}^k s_i Y_i}$$
and
$$\Phi(s_1, \dots, s_n) = q_0 \circ e^{s_n Y_n} \circ \dots \circ e^{s_{k+1} Y_{k+1}} \circ e^{\sum_{i=1}^k s_i Y_i}.$$
Then for $i = 1, \dots, k$
$$\begin{aligned}
\Phi_* \left(\frac{\partial}{\partial s_i} \right)\Bigg|_{\Phi(s)} &= \left.\frac{\partial}{\partial \varepsilon}\right|_{\varepsilon=0} \Phi(s_1, \dots, s_i + \varepsilon, \dots, s_n) \\
&= \left.\frac{\partial}{\partial \varepsilon}\right|_{\varepsilon=0} q_0 \circ e^{s_n Y_n} \circ \dots \circ e^{s_{k+1} Y_{k+1}} \circ e^{\sum_{j=1}^k s_j Y_j} \circ e^{\varepsilon Y_i} \\
&= q_0 \circ e^{s_n Y_n} \circ \dots \circ e^{s_{k+1} Y_{k+1}} \circ e^{\sum_{j=1}^k s_j Y_j} \circ \left.\frac{\partial}{\partial \varepsilon}\right|_{\varepsilon=0} e^{\varepsilon Y_i} \\
&= \Phi(s) \circ Y_i.
\end{aligned}$$

□

Now we can prove Theorem 4.1 on local equivalence of nonlinear systems with linear ones.

Proof. Necessity is obvious since Lie brackets are invariant with respect to diffeomorphisms, and for linear systems conditions (4.2), (4.3) hold.

Sufficiency. Select a basis of the space $T_{q_0}M$ among vectors of the form $(\operatorname{ad} f_0)^j f_i(q_0)$:
$$\begin{aligned}
&Y_\alpha = (\operatorname{ad} f_0)^{j_\alpha} f_{i_\alpha}, \qquad \alpha = 1, \dots, n, \quad 0 \le j_\alpha \le n-1, \quad 1 \le i_\alpha \le m, \\
&\operatorname{span}(q_0 \circ Y_1, \dots, q_0 \circ Y_n) = T_{q_0}M.
\end{aligned}$$
By Lemma 4.1, there exists a rectifying diffeomorphism:
$$\Phi : O_{q_0} \to O_0, \qquad \Phi_* Y_\alpha = \frac{\partial}{\partial x_\alpha}, \qquad \alpha = 1, \dots, n.$$
We show that Φ is the required diffeomorphism.

(1) First we check that the vector fields $\Phi_* f_i$, $i = 1, \ldots, m$ are constant. That is, we show that in the decomposition

$$\Phi_* f_i = \sum_{\alpha=1}^{n} \beta_\alpha^i(x) \frac{\partial}{\partial x_\alpha}. \qquad i = 1, \ldots, m$$

the functions $\beta_\alpha^i(x)$ are constant. We have

$$[\frac{\partial}{\partial x_\alpha}, \Phi_* f_i] = \sum_{\alpha=1}^{n} \frac{\partial \beta_\alpha^i}{\partial x_j} \frac{\partial}{\partial x_\alpha}, \tag{4.4}$$

on the other hand

$$[\frac{\partial}{\partial x_\alpha}, \Phi_* f_i] = [\Phi_* Y_\alpha, \Phi_* f_i] = \Phi_*[Y_\alpha, f_i] = \Phi_*[(\mathrm{ad}\, f_0)^{j_\alpha} f_{i_\alpha}, f_i] = 0 \tag{4.5}$$

by hypothesis (4.3). Now we compare (4.4) and (4.5) and obtain

$$\frac{\partial \beta_\alpha^i}{\partial x_j} \frac{\partial}{\partial x_\alpha} \equiv 0 \quad \Rightarrow \quad \beta_\alpha^i = \mathrm{const}, \quad i = 1, \ldots m, \quad \alpha = 1, \ldots, n,$$

i.e., $\Phi_* f_i$, $i = 1, \ldots, m$, are constant vector fields b_i, $i = 1, \ldots, m$.

(2) Now we show that the vector field $\Phi_* f_0$ is linear. We prove that in the decomposition

$$\Phi_* f_0 = \sum_{i=1}^{n} \beta_i(x) \frac{\partial}{\partial x_i}$$

all functions $\beta_i(x)$, $i = 1, \ldots, n$ are linear. Indeed,

$$\begin{aligned}
\sum_{\alpha=1}^{n} \frac{\partial^2 \beta_i}{\partial x_\alpha \partial x_\beta} \frac{\partial}{\partial x_i} &= [\frac{\partial}{\partial x_\alpha}, [\frac{\partial}{\partial x_\beta}, \Phi_* f_0]] \\
&= [\Phi_* Y_\alpha, [\Phi_* Y_\beta, \Phi_* f_0]] = \Phi_*[Y_\alpha, [Y_\beta, f_0]] \\
&= \Phi_*[(\mathrm{ad}\, f_0)^{j_\alpha} f_{i_\alpha}, [(\mathrm{ad}\, f_0)^{j_\beta} f_{i_\beta}, f_0]] \\
&= -\Phi_*[(\mathrm{ad}\, f_0)^{j_\alpha} f_{i_\alpha}, [f_0, (\mathrm{ad}\, f_0)^{j_\beta} f_{i_\beta}]] \\
&= -\Phi_*[(\mathrm{ad}\, f_0)^{j_\alpha} f_{i_\alpha}, (\mathrm{ad}\, f_0)^{j_\beta+1} f_{i_\beta}] \\
&= 0, \qquad \alpha,\ \beta = 1, \ldots, n,
\end{aligned}$$

by hypothesis (4.3). Thus

$$\frac{\partial^2 \beta_i}{\partial x_\alpha \partial x_\beta} \frac{\partial}{\partial x_i} \equiv 0, \qquad i,\ \alpha,\ \beta = 1, \ldots, n,$$

i.e., $\Phi_* f_0$ is a linear vector field $Ax + c$.

For the linear system $\dot{x} = Ax + c + \sum_{i=1}^{m} u_i b_i$, hypothesis (4.2) implies the controllability condition (3.10) □

4.2 Global linearizability

Now we prove the following statement on global equivalence.

Theorem 4.2. *Let M be a smooth connected n-dimensional manifold, and let $f_0, f_1, \ldots, f_m \in \operatorname{Vec} M$. There exists a diffeomorphism*

$$\Phi : M \to T^k \times \mathbb{R}^{n-k}$$

of M to the product of a k-dimensional torus with $\mathbb{R}^{n-k}$ for some $k \leq n$ such that

$$\begin{aligned} (\Phi_* f_0)(x) &= Ax + c, && x \in T^k \times \mathbb{R}^{n-k}, \\ (\Phi_* f_i)(x) &= b_i, && x \in T^k \times \mathbb{R}^{n-k} \quad i = 1, \ldots, m, \end{aligned}$$

for some $n \times n$ matrix A with zero first k rows:

$$Ae_i = 0, \qquad i = 1, \ldots, k, \tag{4.6}$$

and $c, b_1, \ldots, b_m \in \mathbb{R}^n$ that satisfy the controllability condition (3.10) *if and only if the following conditions hold:*

$$(\operatorname{ad} f_0)^j f_i, \qquad j = 0, 1, \ldots, n-1, \quad i = 1, \ldots, m,$$
$$\textit{are complete vector fields}, \tag{4.7}$$
$$\operatorname{span}\{q \circ (\operatorname{ad} f_0)^j f_i \mid j = 0, \ldots, n-1, \ i = 1, \ldots, m\} = T_q M. \tag{4.8}$$
$$q \circ [(\operatorname{ad} f_0)^{j_1} f_{i_1}, (\operatorname{ad} f_0)^{j_2} f_{i_2}] = 0,$$
$$q \in M, \quad 0 \leq j_1, \ j_2 \leq n, \quad 1 \leq i_1, \ i_2 \leq m. \tag{4.9}$$

Remarks. (1) If M is additionally supposed simply connected, then it is diffeomorphic to $\mathbb{R}^n$, i.e., $k = 0$.

(2) If, on the contrary, M is compact, i.e., diffeomorphic to T^n and $m < n$, then there are no globally linearizable controllable systems on M. Indeed, then $A = 0$, and the controllability condition (3.10) is violated.

Proof. Sufficiency. Fix a point $q_0 \in M$ and find a basis in $T_{q_0} M$ of vectors of the form

$$\begin{aligned} &Y_\alpha = (\operatorname{ad} f_0)^{j_\alpha} f_{i_\alpha}, \qquad \alpha = 1, \ldots, n, \\ &\operatorname{span}(q_0 \circ Y_1, \ldots, q_0 \circ Y_n) = T_{q_0} M. \end{aligned}$$

(1) First we show that the vector fields $Y_1, \ldots, Y_n$ are linearly independent everywhere in M. The set

$$O = \{q \in M \mid \operatorname{span}(q \circ Y_1, \ldots, q \circ Y_n) = T_q M\}$$

is obviously open. We show that it is closed. In this set we have a decomposition

$$q \circ (\operatorname{ad} f_0)^j f_i = q \circ \sum_{\alpha=1}^{n} a_\alpha^{ij} Y_\alpha, \qquad q \in O, \quad j = 0, \dots, n-1, \quad i = 1, \dots, m, \tag{4.10}$$

for some functions $a_\alpha^{ij} \in C^\infty(O)$. We prove that actually all a_α^{ij} are constant. We have

$$\begin{aligned}
0 &= [Y_\beta, \sum_{\alpha=1}^{n} a_\alpha^{ij} Y_\alpha] \\
&\quad \text{(by Leibniz rule } [X, aY] = (Xa)Y + a[X,Y]) \\
&= \sum_{\alpha=1}^{n} a_\alpha^{ij} [Y_\beta, Y_\alpha] + \sum_{\alpha=1}^{n} (Y_\beta a_\alpha^{ij}) Y_\alpha \\
&= \sum_{\alpha=1}^{n} (Y_\beta a_\alpha^{ij}) Y_\alpha, \qquad \beta = 1, \dots, n, \quad j = 0, \dots, n-1, \quad i = 1, \dots, m,
\end{aligned}$$

thus

$$Y_\beta a_\alpha^{ij} = 0 \quad \Rightarrow \quad a_\alpha^{ij}\big|_O = \text{const},$$
$$\alpha = 1, \dots, n, \quad j = 0, \dots, n-1, \quad i = 1, \dots, m.$$

That is why equality (4.10) holds in the closure $\overline{O}$. Thus the vector fields $Y_1, \dots, Y_n$ are linearly independent in $\overline{O}$ (if this is not the case, then the whole family $(\operatorname{ad} f_0)^j f_i$, $j = 0, \dots, n-1$, $i = 1, \dots, m$, is not linearly independent in $\overline{O}$). Hence the set O is closed. Since it is simultaneously open and M is connected,

$$O = M,$$

i.e., the vector fields $Y_1, \dots, Y_n$ are linearly independent in M.

(2) We define the "inverse" Ψ of the required diffeomorphism as follows:

$$\begin{aligned}
\Psi(x_1, \dots, x_n) &= q_0 \circ e^{x_1 Y_1} \circ \cdots \circ e^{x_n Y_n} \\
&\quad \text{(since the vector fields } Y_\alpha \text{ commute)} \\
&= q_0 \circ e^{\sum_{\alpha=1}^{n} x_\alpha Y_\alpha}, \qquad x = (x_1, \dots, x_n) \in \mathbb{R}^n.
\end{aligned}$$

(3) We show that the (obviously smooth) mapping $\Psi : \mathbb{R}^n \to M$ is regular, i.e., its differential is surjective. Indeed,

$$\begin{aligned}
\frac{\partial \Psi}{\partial x_\alpha}(x) &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \Psi(x_1, \dots, x_\alpha + \varepsilon, \dots, x_n) \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} q_0 \circ e^{\sum_{\beta=1}^{n} x_\beta Y_\beta + \varepsilon Y_\alpha} \\
&= q_0 \circ e^{\sum_{\beta=1}^{n} x_\beta Y_\beta} \circ Y_\alpha \\
&= \Psi(x) \circ Y_\alpha, \qquad \alpha = 1, \dots, n,
\end{aligned}$$

thus
$$\Psi_{*x}(\mathbb{R}^n) = T_{\Psi(x)}M.$$
The mapping Ψ is regular, thus a local diffeomorphism. In particular, $\Psi(\mathbb{R}^n)$ is open.

(4) We prove that $\Psi(\mathbb{R}^n)$ is closed. Take any point $q \in \overline{\Psi(\mathbb{R}^n)}$. Since the vector fields $Y_1, \dots, Y_n$ are linearly independent, the image of the mapping
$$(y_1, \dots, y_n) \mapsto q \circ e^{\sum_{\alpha=1}^n y_\alpha Y_\alpha}, \qquad y = (y_1, \dots, y_n) \in \mathbb{R}^n$$
contains a neighborhood of the point q. Thus there exists $y \in \mathbb{R}^n$ such that
$$q \circ e^{\sum_{\alpha=1}^n y_\alpha Y_\alpha} \in \Psi(\mathbb{R}^n),$$
i.e.,
$$q \circ e^{\sum_{\alpha=1}^n y_\alpha Y_\alpha} = q_0 \circ e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha}$$
for some $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Then
$$\begin{aligned} q &= q_0 \circ e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha} \circ e^{-\sum_{\alpha=1}^n y_\alpha Y_\alpha} = q_0 \circ e^{\sum_{\alpha=1}^n (x_\alpha - y_\alpha) Y_\alpha} \\ &= \Psi(x - y). \end{aligned}$$
In other words, $q \in \Psi(\mathbb{R}^n)$.

That is why the set $\Psi(\mathbb{R}^n)$ is closed. Since it is open and M is connected,
$$\Psi(\mathbb{R}^n) = M.$$

(5) It is easy to see that the preimage
$$\Psi^{-1}(q_0) = \{x \in \mathbb{R}^n \mid \Psi(x) = q_0\}$$
is a subgroup of the Abelian group $\mathbb{R}^n$. Indeed, let
$$\Psi(x) = q_0 \circ e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha} = \Psi(y) = q_0 \circ e^{\sum_{\alpha=1}^n y_\alpha Y_\alpha} = q_0,$$
then
$$\Psi(x + y) = q_0 \circ e^{\sum_{\alpha=1}^n (x_\alpha + y_\alpha) Y_\alpha} = q_0 \circ e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha} \circ e^{\sum_{\alpha=1}^n y_\alpha Y_\alpha} = q_0.$$
Analogously, if
$$\Psi(x) = q_0 \circ e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha} = q_0,$$
then
$$\Psi(-x) = q_0 \circ e^{-\sum_{\alpha=1}^n x_\alpha Y_\alpha} = q_0.$$
Finally,
$$\Psi(0) = q_0.$$

(6) Moreover, $\Psi^{-1}(q_0)$ is a discrete subgroup G_0 of $\mathbb{R}^n$, i.e., there are no nonzero elements of $\Psi^{-1}(q_0)$ in some neighborhood of the origin in $\mathbb{R}^n$, since Ψ is a local diffeomorphism.

(7) The mapping Ψ is well-defined on the quotient $\mathbb{R}^n/G_0$. Indeed, let $y \in G_0$. Then

$$\begin{aligned}\Psi(x+y) &= q_0 \circ e^{\sum_{\alpha=1}^n (x_\alpha + y_\alpha) Y_\alpha} = q_0 \circ e^{\sum_{\alpha=1}^n y_\alpha Y_\alpha} \circ e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha} \\ &= q_0 \circ e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha} = \Psi(x).\end{aligned}$$

So

$$\Psi : \mathbb{R}^n/G_0 \to M. \tag{4.11}$$

(8) The mapping (4.11) is one-to-one: if

$$\Psi(x) = \Psi(y), \qquad x,\ y \in \mathbb{R}^n,$$

then

$$q_0 \circ e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha} = q_0 \circ e^{\sum_{\alpha=1}^n y_\alpha Y_\alpha},$$

thus

$$q_0 \circ e^{\sum_{\alpha=1}^n (x_\alpha - y_\alpha) Y_\alpha} = q_0,$$

i.e., $x - y \in G_0$.

(9) That is why mapping (4.11) is a diffeomorphism. By Lemma 4.2 (see below), the discrete subgroup G_0 of $\mathbb{R}^n$ is a lattice:

$$G_0 = \left\{ \sum_{i=1}^k n_i e_i \mid n_i \in \mathbb{Z} \right\},$$

thus the quotient is a cylinder:

$$\mathbb{R}^n/G_0 = T^k \times \mathbb{R}^{n-k}.$$

Hence we constructed a diffeomorphism

$$\Phi = \Psi^{-1} : M \to T^k \times \mathbb{R}^{n-k}.$$

Equalities (4.8) and (4.9) follow exactly as in Theorem 4.1.

The vector field $\Phi_* f_0 = Ax + c$ is well-defined on the quotient $T^k \times \mathbb{R}^{n-k}$, that is why equalities (4.6) hold.

Necessity. For a linear system on a cylinder $T^k \times \mathbb{R}^{n-k}$, conditions (4.7) and (4.9) obviously hold. If a linear system is controllable on the cylinder, then it is also controllable on $\mathbb{R}^n$, thus the controllability condition (4.8) is also satisfied. $\square$

Lemma 4.2. *Let Γ be a discrete subgroup in $\mathbb{R}^n$. Then it is a lattice, i.e., there exist linearly independent vectors $e_1, \dots, e_k \in \mathbb{R}^n$ such that*

$$\Gamma = \left\{ \sum_{i=1}^{k} n_i e_i \mid n_i \in \mathbb{Z} \right\}.$$

Proof. We prove by induction on dimension n of the ambient group $\mathbb{R}^n$.

(1) Let $n = 1$. Since the subgroup $\Gamma \subset \mathbb{R}$ is discrete, it contains an element $e_1 \neq 0$ closest to the origin $0 \in \mathbb{R}$. By the group property, all multiples $\pm e_1 \pm e_1 \pm \cdots \pm e_1 = \pm n e_1$, $n = 0, 1, 2, \dots$, are also in Γ. We prove that Γ contains no other elements.

By contradiction, assume that there is an element $x \in \Gamma$ such that $ne_1 < x < (n+1)e_1$, $n \in \mathbb{Z}$. Then the element $y = x - ne_1 \in \Gamma$ is in the interval $(0, e_1) \subset \mathbb{R}$. So $y \neq 0$ is closer to the origin than e_1, a contradiction. Thus $\Gamma = \mathbb{Z}e_1 = \{ne_1 \mid n \in \mathbb{Z}\}$, q.e.d.

(2) We prove the inductive step: let the statement of the lemma be proved for some $n - 1 \in \mathbb{N}$, we prove it for n.

Choose an element $e_1 \in \Gamma$, $e_1 \neq 0$, closest to the origin $0 \in \mathbb{R}^n$. Denote by l the line $\mathbb{R}e_1$, and by Γ_1 the lattice $\mathbb{Z}e_1 \subset \Gamma$. We suppose that $\Gamma \neq \Gamma_1$ (otherwise everything is proved).

Now we show that there is an element $e_2 \in \Gamma \setminus \Gamma_1$ closest to l:

$$\operatorname{dist}(e_2, l) = \min\{\operatorname{dist}(x, l) \mid x \in \Gamma \setminus l\}. \tag{4.12}$$

Take any segment $I = [ne_1, (n+1)e_1] \subset l$, and denote by $\pi : \mathbb{R}^n \to l$ the orthogonal projection from $\mathbb{R}^n$ to l along the orthogonal complement to l in $\mathbb{R}^n$. Since the segment I is compact and the subgroup Γ is discrete, the n-dimensional strip $\pi^{-1}(I)$ contains an element $e_2 \in \Gamma \setminus l$ closest to I:

$$\operatorname{dist}(e_2, I) = \min\{\operatorname{dist}(x, I) \mid x \in (\Gamma \setminus l) \cap \pi^{-1}(I)\}.$$

Then the element e_2 is the required one: it satisfies equality (4.12) since any element that satisfies (4.12) can be translated to the strip $\pi^{-1}(I)$ by elements of the lattice Γ_1.

That is why a sufficiently small neighborhood of l is free of elements of $\Gamma \setminus \Gamma_1$. Thus the quotient group Γ/Γ_1 is a discrete subgroup in $\mathbb{R}^n / l = \mathbb{R}^{n-1}$. By the inductive hypothesis, Γ/Γ_1 is a lattice, hence Γ is also a lattice. □

Chapter 5

The Orbit Theorem and its applications

5.1 Formulation of the Orbit Theorem

Let $\mathcal{F} \subset \operatorname{Vec} M$ be any set of smooth vector fields. In order to simplify notations, we assume that all fields from $\mathcal{F}$ are complete. Actually, all further definitions and results have clear generalizations to the case of noncomplete fields; we leave them to the reader.

We return to the study of attainable sets: we study the structure of the attainable sets of $\mathcal{F}$ by piecewise constant controls

$$\mathcal{A}_{q_0} = \{q_0 \circ e^{t_1 f_1} \circ \cdots \circ e^{t_k f_k} \mid t_i \geq 0,\ f_i \in \mathcal{F},\ k \in \mathbb{N}\}, \qquad q_0 \in M.$$

But first we consider a greater set — the *orbit* of the family $\mathcal{F}$ through a point:

$$\mathcal{O}_{q_0} = \{q_0 \circ e^{t_1 f_1} \circ \cdots \circ e^{t_k f_k} \mid t_i \in \mathbb{R},\ f_i \in \mathcal{F},\ k \in \mathbb{N}\}, \qquad q_0 \in M.$$

In an orbit $\mathcal{O}_{q_0}$, it is allowed to move along vector fields f_i both forward and backwards, while in an attainable set $\mathcal{A}_{q_0}$ only the forward motion is possible.

Although, if the family $\mathcal{F}$ is *symmetric* $\mathcal{F} = -\mathcal{F}$ (i.e., $f \in \mathcal{F} \ \Rightarrow -f \in \mathcal{F}$), then attainable sets coincide with orbits: $\mathcal{O}_{q_0} = \mathcal{A}_{q_0}$, $q_0 \in M$.

In general, orbits have more simple structure that attainable sets. It is described in the following fundamental proposition.

Theorem 5.1 (Orbit Theorem, Nagano–Sussmann). *Let $\mathcal{F} \subset \operatorname{Vec} M$ and $q_0 \in M$. Then:*

(1) *$\mathcal{O}_{q_0}$ is a connected immersed submanifold of M.*

(2) $T_q \mathcal{O}_{q_0} = \operatorname{span}\{q \circ (\operatorname{Ad} P) f \mid P \in \mathcal{P},\ f \in \mathcal{F}\}$, $q \in \mathcal{O}_{q_0}$.

Here we denote by $\mathcal{P}$ the group of diffeomorphisms of M generated by flows in $\mathcal{F}$:

$$\mathcal{P} = \{e^{t_1 f_1} \circ \cdots \circ e^{t_k f_k} \mid t_i \in \mathbb{R},\ f_i \in \mathcal{F},\ k \in \mathbb{N}\} \subset \operatorname{Diff} M.$$

5.2 Immersed submanifolds

Now we discuss an important notion that appears in the Orbit Theorem.

Definition 5.1. A subset W of a smooth n-dimensional manifold is called an *immersed k-dimensional submanifold* of M, $k \leq n$, if there exists a one-to-one immersion

$$\Phi : N \to M, \qquad \operatorname{Ker} \Phi_{*x} = 0 \quad \forall\, x \in N$$

of a k-dimensional smooth manifold N such that

$$W = \Phi(N).$$

Remark. An immersed submanifold W of M can also be defined as a manifold contained in M such that the inclusion mapping

$$i : W \to M, \qquad i : q \mapsto q$$

is an immersion.

Sufficiently small neighborhoods O_x in an immersed submanifold W of M are submanifolds of M, but the whole W is not necessarily a submanifold of M in the sense of Definition 1.1. In general, the own topology of W can be stronger than the topology induced on W by the topology of M.

Example 5.1. Let $\Phi : \mathbb{R} \to \mathbb{R}^2$ be a one-to-one immersion of the line into the plane such that $\lim_{t\to+\infty} \Phi(t) = \Phi(0) = 0 \in \mathbb{R}^2$, Then $W = \Phi(\mathbb{R})$ is an immersed one-dimensional submanifold of $\mathbb{R}^2$. The own topology of W (inherited from $\mathbb{R}$) is stronger than the topology induced by $\mathbb{R}^2$. The intervals $\Phi(-\varepsilon, \varepsilon)$, $\varepsilon > 0$ small enough, are open in the first topology, but not open in the second one.

The notion of immersed submanifold appears inevitably in the description of orbits of families of vector fields. Already the orbit of one vector field (i.e., its trajectory) is an immersed submanifold, but may fail to be a submanifold in the sense of Definition 1.1.

Example 5.2. Oscillator with 2 degrees of freedom is described by the equations:

$$\ddot{x} + \alpha^2 x = 0, \qquad x \in \mathbb{R},$$
$$\ddot{y} + \beta^2 y = 0, \qquad y \in \mathbb{R}.$$

In the complex variables

$$z = x - i\dot{x}/\alpha, \qquad w = y - i\dot{y}/\beta$$

these equations read

$$\begin{aligned} \dot{z} &= i\alpha z, \qquad z \in \mathbb{C}, \\ \dot{w} &= i\beta w, \qquad w \in \mathbb{C}, \end{aligned} \tag{5.1}$$

and their solutions have the form

$$z(t) = e^{i\alpha t} z(0),$$
$$w(t) = e^{i\beta t} w(0).$$

Any solution $(z(t), w(t))$ to equations (5.1) belongs to an invariant torus

$$T^2 = \{(z, w) \in \mathbb{C}^2 \mid |z| = \text{const}, \ |w| = \text{const}\}.$$

Any such torus is parametrized by arguments of z, w modulo 2π, thus it is a group: $T^2 \simeq \mathbb{R}^2/(2\pi\mathbb{Z})^2$.

We introduce a new parameter $\tau = \alpha t$, then trajectories (z, w) become images of the line $\{(\tau, (\beta/\alpha)\tau) \mid \tau \in \mathbb{R}\}$ under the immersion

$$(\tau, (\beta/\alpha)\tau) \mapsto (\tau + 2\pi\mathbb{Z}, (\beta/\alpha)\tau + 2\pi\mathbb{Z}) \in \mathbb{R}^2/(2\pi\mathbb{Z})^2,$$

thus immersed submanifolds of the torus.

If the ratio β/α is irrational, then trajectories are everywhere dense in the torus: they form the irrational winding of the torus. In this case, trajectories, i.e., orbits of a vector field, are not submanifolds, but just immersed submanifolds.

Remark. Immersed submanifolds inherit many local properties of submanifolds. In particular, the tangent space to an immersed submanifold $W = \operatorname{Im}\Phi \subset M$, Φ an immersion, is given by

$$T_{\Phi(x)} W = \operatorname{Im}\Phi_{*x}.$$

Further, it is easy to prove the following property of a vector field $V \in \operatorname{Vec} M$:

$$V(q) \in T_q W \quad \forall\, q \in W \quad \Rightarrow \quad q \circ e^{tV} \in W \quad \forall\, q \in W,$$

for all t close enough to 0.

5.3 Corollaries of the Orbit Theorem

Before proving the Orbit Theorem, we obtain several its corollaries.

Let $\mathcal{O}_{q_0}$ be an orbit of a family $\mathcal{F} \subset \operatorname{Vec} M$.

First of all, if $f \in \mathcal{F}$, then $f(q) \in T_q\mathcal{O}_{q_0}$ for all $q \in \mathcal{O}_{q_0}$. Indeed, the trajectory $q \circ e^{tf}$ belongs to the orbit $\mathcal{O}_{q_0}$, thus its velocity vector $f(q)$ is in the tangent space $T_q\mathcal{O}_{q_0}$.

Further, if $f_1, f_2 \in \mathcal{F}$, then $[f_1, f_2](q) \in T_q\mathcal{O}_{q_0}$ for all $q \in \mathcal{O}_{q_0}$. This follows since the vector $[f_1, f_2](q)$ is tangent to the trajectory

$$t \mapsto q \circ e^{tf_1} \circ e^{tf_2} \circ e^{-tf_1} \circ e^{-tf_2} \in \mathcal{O}_{q_0}.$$

Given three vector fields $f_1, f_2, f_3 \in \mathcal{F}$, we have $[f_1, [f_2, f_3]](q) \in T_q\mathcal{O}_{q_0}$, $q \in \mathcal{O}_{q_0}$. Indeed, it follows that $[f_2, f_3](q) \in T_q\mathcal{O}_{q_0}$, $q \in \mathcal{O}_{q_0}$, then all trajectories of the field $[f_2, f_3]$ starting in the immersed submanifold $\mathcal{O}_{q_0}$ do not leave it. Then we repeat the argument of the previous items.

We can go on and consider Lie brackets of arbitrarily high order

$$[f_1, [\dots [f_{k-1}, f_k] \dots]](q)$$

as tangent vectors to $\mathcal{O}_{q_0}$ if $f_i \in \mathcal{F}$. These considerations can be summarized in terms of the Lie algebra of vector fields generated by $\mathcal{F}$:

$$\operatorname{Lie} \mathcal{F} = \operatorname{span}\{[f_1, [\dots [f_{k-1}, f_k] \dots]] \mid f_i \in \mathcal{F},\ k \in \mathbb{N}\} \subset \operatorname{Vec} M,$$

and its evaluation at a point $q \in M$:

$$\operatorname{Lie}_q \mathcal{F} = \{q \circ V \mid V \in \operatorname{Lie} \mathcal{F}\} \subset T_q M.$$

We obtain the following statement.

Corollary 5.1.

$$\operatorname{Lie}_q \mathcal{F} \subset T_q \mathcal{O}_{q_0} \tag{5.2}$$

for all $q \in \mathcal{O}_{q_0}$.

Remark. We show soon that in many important cases inclusion (5.2) turns into equality. In the general case, we have the following estimate:

$$\dim \operatorname{Lie}_q \mathcal{F} \leq \dim \mathcal{O}_{q_0}, \qquad q \in \mathcal{O}_{q_0}.$$

Another important corollary of the Orbit Theorem is the following proposition often used in control theory.

Theorem 5.2 (Rashevsky–Chow). *Let M be a connected smooth manifold, and let $\mathcal{F} \subset \operatorname{Vec} M$. If the family $\mathcal{F}$ is completely nonholonomic:*

$$\operatorname{Lie}_q \mathcal{F} = T_q M \qquad \forall\ q \in M, \tag{5.3}$$

then

$$\mathcal{O}_{q_0} = M \quad \forall\ q_0 \in M. \tag{5.4}$$

Definition 5.2. A family $\mathcal{F} \subset \operatorname{Vec} M$ that satisfies property (5.3) is called *completely nonholonomic* or *bracket-generating*.

Proof. By Corollary 5.1, equality (5.3) means that any orbit $\mathcal{O}_{q_0}$ is an open set in M.

Further, consider the following equivalence relation in M:

$$q_1 \sim q_2 \ \Leftrightarrow q_2 \in \mathcal{O}_{q_1}, \qquad q_1,\ q_2 \in M. \tag{5.5}$$

The manifold M is the union of (naturally disjoint) equivalence classes. Each class is an open subset of M and M is connected. Hence there is only one nonempty class. That is, M is a single orbit $\mathcal{O}_{q_0}$. □

For symmetric families attainable sets coincide with orbits, thus we have the following statement.

Corollary 5.2. *Symmetric bracket-generating families are completely controllable on connected manifolds.*

5.4 Proof of the Orbit Theorem

Introduce the notation:

$$(\operatorname{Ad}\mathcal{P})\mathcal{F} \stackrel{\text{def}}{=} \{(\operatorname{Ad}P)f \mid P \in \mathcal{P},\ f \in \mathcal{F}\} \subset \operatorname{Vec} M.$$

Consider the following subspace of $T_q M$:

$$\Pi_q \stackrel{\text{def}}{=} \operatorname{span}\{q \circ (\operatorname{Ad}\mathcal{P})\mathcal{F}\}.$$

This space is a candidate for the tangent space $T_q \mathcal{O}_{q_0}$.

Lemma 5.1. $\dim \Pi_q = \dim \Pi_{q_0}$ *for all* $q \in \mathcal{O}_{q_0}$, $q_0 \in M$.

Proof. If $q \in \mathcal{O}_{q_0}$, then $q = q_0 \circ Q$ for some diffeomorphism $Q \in \mathcal{P}$.

Take an arbitrary element $q_0 \circ (\operatorname{Ad}P)f$ in Π_{q_0}, $P \in \mathcal{P}$, $f \in \mathcal{F}$. Then

$$\begin{aligned} Q_*(q_0 \circ (\operatorname{Ad}P)f) &= q_0 \circ (\operatorname{Ad}P)f \circ Q = q_0 \circ P \circ f \circ P^{-1} \circ Q \\ &= (q_0 \circ Q) \circ (Q^{-1} \circ P \circ f \circ P^{-1} \circ Q) \\ &= q \circ \operatorname{Ad}(Q^{-1} \circ P)f \in \Pi_q \end{aligned}$$

since $Q^{-1} \circ P \in \mathcal{P}$.

We have $Q_*\Pi_{q_0} \subset \Pi_q$, thus $\dim \Pi_{q_0} \leq \dim \Pi_q$. But q_0 and q can be switched, that is why $\dim \Pi_q \leq \dim \Pi_{q_0}$. Finally, $\dim \Pi_q = \dim \Pi_{q_0}$. □

Now we prove the Orbit Theorem.

Proof. The manifold M is divided into disjoint equivalence classes of relation (5.5) — orbits $\mathcal{O}_q$. We introduce a new "strong" topology on M in which all orbits are connected components.

For any point $q \in M$, denote $m = \dim \Pi_q$ and pick elements $V_1, \ldots, V_m \in (\operatorname{Ad}\mathcal{P})\mathcal{F}$ such that

$$\operatorname{span}(V_1(q), \ldots, V_m(q)) = \Pi_q. \tag{5.6}$$

Introduce a mapping:

$$G_q \ :\ (t_1, \ldots, t_m) \mapsto q \circ e^{t_1 V_1} \circ \cdots \circ e^{t_m V_m}, \qquad t_i \in \mathbb{R}.$$

We have

$$\left.\frac{\partial\, G_q}{\partial\, t_i}\right|_0 = V_i(q),$$

thus in a sufficiently small neighborhood O_0 of the origin $0 \in \mathbb{R}^m$

$$\frac{\partial\, G_q}{\partial\, t_1} \wedge \cdots \wedge \left.\frac{\partial\, G_q}{\partial\, t_m}\right|_{O_0} \neq 0,$$

i.e., $G_q|_{O_0}$ is an immersion.

The sets of the form $G_q(O_0)$, $q \in M$, are candidates for elements of a topology base on M. We prove several properties of these sets.

(1) Since the mappings G_q are regular, the sets $G_q(O_0)$ are m-dimensional submanifolds of M, may be, for smaller neighborhoods O_0.

(2) We show that $G_q(O_0) \subset \mathcal{O}_q$. Any element of base (5.6) has the form $V_i = (\operatorname{Ad} P_i) f_i$, $P_i \in \mathcal{P}$, $f_i \in \mathcal{F}$. Then

$$e^{tV_i} = e^{t(\operatorname{Ad} P_i) f_i} = e^{t P_i \circ f_i \circ P_i^{-1}} = P_i \circ e^{tf_i} \circ P_i^{-1} \in \mathcal{P},$$

thus

$$G_q(t) = q \circ e^{tV_i} \in \mathcal{O}_q, \qquad t \in O_0.$$

(3) We show that $G_{*t}(T_t\mathbb{R}^m) = \Pi_{G(t)}$, $t \in O_0$. Since $\operatorname{rank} G_{*t}|_{O_0} = m$ and $\dim \Pi_{G(t)}\big|_{O_0} = m$, it remains ro prove that $\left.\frac{\partial G_q}{\partial t_i}\right|_t \in \Pi_{G_q(t)}$ for $t \in O_0$. We have

$$\begin{aligned}
\frac{\partial}{\partial t_i} G_q(t) &= \frac{\partial}{\partial t_i} q \circ e^{t_1 V_1} \circ \cdots \circ e^{t_m V_m} \\
&= q \circ e^{t_1 V_1} \circ \cdots \circ e^{t_i V_i} \circ V_i \circ e^{t_{i+1} V_{i+1}} \circ \cdots \circ e^{t_m V_m} \\
&= q \circ e^{t_1 V_1} \circ \cdots \circ e^{t_i V_i} \circ e^{t_{i+1} V_{i+1}} \circ \cdots \circ e^{t_m V_m} \\
&\quad \circ e^{-t_m V_m} \circ \cdots \circ e^{-t_{i+1} V_{i+1}} \circ V_i \circ e^{t_{i+1} V_{i+1}} \circ \cdots \circ e^{t_m V_m}
\end{aligned}$$

(introduce the notation $Q = e^{t_{i+1} V_{i+1}} \circ \cdots \circ e^{t_m V_m} \in \mathcal{P}$)

$$= G_q(t) \circ Q^{-1} \circ V_i \circ Q = G_q(t) \circ \operatorname{Ad} Q^{-1} V_i \in \Pi_{G_q(t)}.$$

(4) We prove that sets of the form $G_q(O_0)$, $q \in M$, form a topology base in M. It is enough to prove that any nonempty intersection $G_q(O_0) \cap G_{\tilde{q}}(\widetilde{O}_0)$ contains a set of the form $G_{\hat{q}}(\widehat{O}_0)$.

Let a point $\hat{q}$ belong to $G_q(O_0)$. Then $\dim \Pi_{\hat{q}} = \dim \Pi_q = m$. Consider the mapping

$$\begin{aligned}
&G_{\hat{q}} \ : \ (t_1, \ldots, t_m) \mapsto \hat{q} \circ e^{t_1 \widehat{V}_1} \circ \cdots \circ e^{t_m \widehat{V}_m}, \\
&\operatorname{span}(\hat{q} \circ \widehat{V}_1, \ldots, \hat{q} \circ \widehat{V}_m) = \Pi_{\hat{q}}.
\end{aligned}$$

It is enough to prove that for small enough $(t_1, \ldots, t_m)$

$$G_{\hat{q}}(t_1, \ldots, t_m) \in G_q(O_0),$$

then we can replace $G_q(O_0)$ by $G_{\tilde{q}}(\widetilde{O}_0)$. We do this step by step. Consider the curve $t_1 \mapsto \widehat{q} \circ e^{t_1 \widehat{V}_1}$. By property (3) above, $\widehat{V}_1(q') \in \Pi_{q'}$ for $q' \in G_q(O_0)$ and sufficiently close to $\widehat{q}$. Since $G_q(O_0)$ is a submanifold of M and $\Pi_q = T_q G_q(O_0)$, the curve $\widehat{q} \circ e^{t_1 \widehat{V}_1}$ belongs to $G_q(O_0)$ for sufficiently small $|t_1|$.

We repeat this argument and show that

$$(\widehat{q} \circ e^{t_1 \widehat{V}_1}) \circ e^{t_2 \widehat{V}_2} \in G_q(O_0)$$

for small $|t_1|$, $|t_2|$. We continue this procedure and obtain the inclusion

$$(\widehat{q} \circ e^{t_1 \widehat{V}_1} \circ \cdots \circ e^{t_{m-1} \widehat{V}_{m-1}}) \circ e^{t_m \widehat{V}_m} \in G_q(O_0)$$

for $(t_1, \ldots, t_m)$ sufficiently close to $0 \in \mathbb{R}^m$.

Property (4) follows, and the sets $G_q(O_0)$, $q \in M$, form a topology base on M. We denote by $M^{\mathcal{F}}$ the topological space obtained, i.e., the set M endowed with the "strong" topology just introduced.

(5) For any $q_0 \in M$, the orbit $\mathcal{O}_{q_0}$ is connected, open, and closed in the "strong" topology.

Connectedness: all mappings $t \mapsto q \circ e^{tf}$, $f \in \mathcal{F}$, are continuous in the "strong" topology, thus any point $q \in \mathcal{O}_{q_0}$ can be connected with q_0 by a path continuous in $M^{\mathcal{F}}$.

Openness: for any $q \in \mathcal{O}_{q_0}$, a set of the form $G_q(O_0) \subset \mathcal{O}_{q_0}$ is a neighborhood of the point q in $M^{\mathcal{F}}$.

Closedness: any orbit is a complement to a union of open sets (orbits), thus it is closed.

So each orbit $\mathcal{O}_{q_0}$ is a connected component of the topological space $M^{\mathcal{F}}$.

(6) A smooth structure on each orbit $\mathcal{O}_{q_0}$ is defined by choosing $G_q(O_0)$ to be coordinate neighborhoods and G_q^{-1} coordinate mappings. Since $G_q|_{O_0}$ are immersions, then each orbit $\mathcal{O}_{q_0}$ is an immersed submanifold of M. Notice that dimension of these submanifolds may vary for different q_0.

(7) By property (3) above, $T_q \mathcal{O}_{q_0} = \Pi_q$, $q \in \mathcal{O}_{q_0}$.

The Orbit Theorem is proved. $\square$

The description of the tangent space of an orbit given by the Orbit Theorem:

$$T_q \mathcal{O}_{q_0} = q \circ (\operatorname{Ad} \mathcal{P}) \mathcal{F},$$

is rather implicit since the structure of the group $\mathcal{P}$ is quite complex. However, we already obtained the lower estimate

$$\operatorname{Lie}_q \mathcal{F} \subset \operatorname{span}(q \circ (\operatorname{Ad} \mathcal{P}) \mathcal{F}). \tag{5.7}$$

from the Orbit Theorem. Notice that this inclusion can easily be proved directly. We make use of the asymptotic expansion of the field $\operatorname{Ad} e^{tf} \widehat{f} = e^{t \operatorname{ad} f} \widehat{f}$. Take an arbitrary element $\operatorname{ad} f_1 \circ \cdots \circ \operatorname{ad} f_k \widehat{f} \in \operatorname{Lie} \mathcal{F}$, $f_i, \widehat{f} \in \mathcal{F}$. We have $\operatorname{Ad}(e^{t_1 f_1} \circ$

$\cdots \circ e^{t_k f_k})\widehat{f} \in (\operatorname{Ad}\mathcal{P})\mathcal{F}$, thus

$$
\begin{aligned}
q \circ & \left.\frac{\partial^k}{\partial t_1 \cdots \partial t_k}\right|_0 \operatorname{Ad}(e^{t_1 f_1} \circ \cdots \circ e^{t_k f_k})\widehat{f} \\
&= q \circ \left.\frac{\partial^k}{\partial t_1 \cdots \partial t_k}\right|_0 (e^{t_1 \operatorname{ad} f_1} \circ \cdots \circ e^{t_k \operatorname{ad} f_k})\widehat{f} \\
&= q \circ \operatorname{ad} f_1 \circ \cdots \circ \operatorname{ad} f_k \ \widehat{f} \in \operatorname{span}(q \circ (\operatorname{Ad}\mathcal{P})\mathcal{F}).
\end{aligned}
$$

Now we consider a situation where inclusion (5.7) is strict.

Example 5.3. Let $M = \mathbb{R}^2$, $\mathcal{F} = \left\{\frac{\partial}{\partial x^1},\ a(x^1)\frac{\partial}{\partial x^2}\right\}$, where the function $a \in C^\infty(\mathbb{R})$, $a \not\equiv 0$, has a compact support.

It is easy to see that the orbit $\mathcal{O}_x$ through any point $x \in \mathbb{R}^2$ is the whole plane $\mathbb{R}^2$. Indeed, the family $\mathcal{F} \cup (-\mathcal{F})$ is completely controllable in the plane. Given an initial point $x_0 = (x_0^1, x_0^2)$ and a terminal point $x_1 = (x_1^1, x_1^2)$, we can steer x_0 to x_1: first we go from x_0 by a field $\pm\frac{\partial}{\partial x^1}$ to a point $(\tilde{x}^1, x_0^2)$ with $a(\tilde{x}^1) \neq 0$, then we go by a field $\pm a(x^1)\frac{\partial}{\partial x^2}$ to a point $(\tilde{x}^1, x_1^2)$, and finally we reach (x_1^1, x_1^2) along $\pm\frac{\partial}{\partial x^1}$.

On the other hand, we have

$$
\dim \operatorname{Lie}_{(x^1,x^2)}(\mathcal{F}) = \begin{cases} 1, & x^1 \notin \operatorname{supp} a, \\ 2, & a(x^1) \neq 0. \end{cases}
$$

That is, $x \circ (\operatorname{Ad}\mathcal{P})\mathcal{F} = T_x\mathbb{R}^2 \neq \operatorname{Lie}_x \mathcal{F}$ if $x^1 \notin \operatorname{supp} a$.

Although, such example is essentially non-analytic. In the analytic case, inclusion (5.7) turns into equality. We prove this statement in the next section.

5.5 Analytic case

The set $\operatorname{Vec} M$ is not just a Lie algebra (i.e., a vector space close under the operation of Lie bracket), but also a *module* over $C^\infty(M)$: any vector field $V \in \operatorname{Vec} M$ can be multiplied by a function $a \in C^\infty(M)$, and the resulting vector field $aV \in \operatorname{Vec} M$. If vector fields are considered as derivations of $C^\infty(M)$, then the product of a function a and a vector field V is the vector field

$$
(aV)b = a \cdot (Vb), \qquad b \in C^\infty(M).
$$

In local coordinates, each component of V at a point $q \in M$ is multiplied by $a(q)$.

Exercise 5.1. Let $X, Y \in \operatorname{Vec} M$, $a \in C^\infty(M)$, $P \in \operatorname{Diff} M$. Prove the equalities:

$$
\begin{aligned}
&(\operatorname{ad} X)(aY) = (Xa)Y + a(\operatorname{ad} X)Y, \\
&(\operatorname{Ad} P)(aX) = (Pa)\ \operatorname{Ad} P\, X.
\end{aligned}
$$

A submodule $\mathcal{V} \subset \operatorname{Vec} M$ is called *finitely generated* over $C^\infty(M)$ if it has a global basis of vector fields:

$$\exists\ V_1, \dots, V_k \in \operatorname{Vec} M \quad \text{such that} \quad \mathcal{V} = \left\{ \sum_{i=1}^k a_i V_i \mid a_i \in C^\infty(M) \right\}.$$

Lemma 5.2. *Let $\mathcal{V} \subset \operatorname{Vec} M$ be a finitely generated submodule over $C^\infty(M)$. Assume that*

$$(\operatorname{ad} X)\mathcal{V} = \{(\operatorname{ad} X)V \mid V \in \mathcal{V}\} \subset \mathcal{V}$$

for a vector field $X \in \operatorname{Vec} M$. Then

$$\operatorname{Ad} e^{tX} \mathcal{V} = \mathcal{V}.$$

Proof. Let $V_1, \dots, V_k$ be a basis of $\mathcal{V}$. By the hypothesis of the lemma,

$$[X, V_i] = \sum_{j=1}^k a_{ij} V_j \tag{5.8}$$

for some functions $a_{ij} \in C^\infty(M)$. We have to prove that the vector fields

$$V_i(t) = (\operatorname{Ad} e^{tX})V_i = e^{t \operatorname{ad} X} V_i, \qquad t \in \mathbb{R},$$

can be expressed as linear combinations of the fields V_i with coefficients from $C^\infty(M)$.

We define an ODE for $V_i(t)$:

$$\begin{aligned} \dot{V}_i(t) &= e^{t \operatorname{ad} X}[X, V_i] = e^{t \operatorname{ad} X} \sum_{j=1}^k a_{ij} V_j \\ &= \sum_{j=1}^k \left(e^{tX} a_{ij}\right) V_j(t). \end{aligned}$$

For a fixed $q \in M$, define the $k \times k$ matrix:

$$A(t) = (a_{ij}(t)), \qquad a_{ij}(t) = e^{tX} a_{ij}, \quad i,\ j = 1, \dots, k.$$

Then we have a linear system of ODEs:

$$\dot{V}_i(t) = \sum_{j=1}^k a_{ij}(t) V_j(t). \tag{5.9}$$

Find a fundamental matrix Γ of this system:

$$\dot{\Gamma} = A(t)\Gamma, \qquad \Gamma(0) = \operatorname{Id}.$$

Since $A(t)$ smoothly depends on q, then Γ depends smoothly on q as well:

$$\Gamma(t) = (\gamma_{ij}(t)), \qquad \gamma_{ij}(t) \in C^\infty(M), \quad i,\ j = 1, \dots, k, \quad t \in \mathbb{R}.$$

Now solutions of the linear system (5.9) can be written as follows:

$$V_i(t) = \sum_{j=1}^{k} \gamma_{ij}(t) V_j(0).$$

But $V_i(0) = V_i$ are the generators of the module, and the required decomposition of $V_i(t)$ along the generators is obtained. $\square$

A submodule $\mathcal{V} \subset \operatorname{Vec} M$ is called *locally finitely generated* over $C^\infty(M)$ if any point $q \in M$ has a neighborhood $O \subset M$ in which the restriction $\mathcal{F}|_O$ is finitely generated over $C^\infty(O)$, i.e., has a basis of vector fields.

Theorem 5.3. *Let $\mathcal{F} \subset \operatorname{Vec} M$. Suppose that the module* $\operatorname{Lie} \mathcal{F}$ *is locally finitely generated over $C^\infty(M)$. Then*

$$T_q \mathcal{O}_{q_0} = \operatorname{Lie}_q \mathcal{F}, \qquad q \in \mathcal{O}_{q_0} \tag{5.10}$$

for any orbit $\mathcal{O}_{q_0}$, $q_0 \in M$, of the family $\mathcal{F}$.

We prove this theorem later, but now obtain from it the following consequence.

Corollary 5.3. *If M and $\mathcal{F}$ are real analytic, then equality* (5.10) *holds.*

Proof. In the analytic case, $\operatorname{Lie} \mathcal{F}$ is locally finitely generated. Indeed, any module generated by analytic vector fields is locally finitely generated. This is Nötherian property of the ring of germs of analytic functions. $\square$

Now we prove Theorem 5.3.

Proof. By the Orbit Theorem,

$$T_q \mathcal{O}_{q_0} = \operatorname{span} \left\{ q \circ \operatorname{Ad}\left(e^{t_1 f_1} \circ \cdots \circ e^{t_k f_k}\right) \widehat{f} \mid f_i,\ \widehat{f} \in \mathcal{F},\ t_k \in \mathbb{R},\ k \in \mathbb{N} \right\}. \tag{5.11}$$

By definition of the Lie algebra $\operatorname{Lie} \mathcal{F}$,

$$(\operatorname{ad} f) \operatorname{Lie} \mathcal{F} \subset \operatorname{Lie} \mathcal{F} \qquad \forall\ f \in \mathcal{F}.$$

Apply Lemma 5.2 for the locally finitely generated $C^\infty(M)$-module $\mathcal{V} = \operatorname{Lie} \mathcal{F}$. We obtain

$$\left(\operatorname{Ad} e^{tf}\right) \operatorname{Lie} \mathcal{F} \subset \operatorname{Lie} \mathcal{F} \qquad \forall\ f \in \mathcal{F}.$$

That is why

$$\operatorname{Ad}\left(e^{t_1 f_1} \circ \cdots \circ e^{t_k f_k}\right) \widehat{f} = \operatorname{Ad} e^{t_1 f_1} \circ \cdots \circ \operatorname{Ad} e^{t_k f_k} \widehat{f} \in \operatorname{Lie} \mathcal{F}$$

for any $f_i,\ \widehat{f} \in \mathcal{F}$, $t_k \in \mathbb{R}$. In view of equality (5.11),

$$T_q \mathcal{O}_{q_0} \subset \operatorname{Lie}_q \mathcal{F}.$$

But the reverse inclusion (5.7) was already obtained. Thus $T_q \mathcal{O}_{q_0} = \operatorname{Lie}_q \mathcal{F}$.

Another proof of the theorem can be obtained via local convergence of the exponential series in the analytic case. $\square$

5.6 Frobenius Theorem

We apply the Orbit Theorem to obtain the classical Frobenius Theorem as a corollary.

Definition 5.3. A *distribution* $\Delta \subset TM$ on a smooth manifold M is a family of linear subspaces $\Delta_q \subset T_qM$ smoothly depending on a point $q \in M$. Dimension of the subspaces Δ_q, $q \in M$, is assumed constant.

Geometrically, at each point $q \in M$ there is attached a space $\Delta_q \subset T_qM$, i.e., we have a field of tangent subspaces on M.

Definition 5.4. A distribution Δ on a manifold M is called *integrable* if for any point $q \in M$ there exists an immersed submanifold $N_q \subset M$, $q \in N_q$, such that

$$T_{q'}N_q = \Delta_{q'} \qquad \forall\, q' \in N_q.$$

The submanifold N_q is called an *integral manifold* of the distribution Δ through the point q.

In other words, integrability of a distribution $\Delta \subset TM$ means that through any point $q \in M$ we can draw a submanifold N_q whose tangent spaces are elements of the distribution Δ.

Remark. If $\dim \Delta_q = 1$, then Δ is integrable by Theorem 1.2 on existence and uniqueness of solutions of ODEs. Indeed, in a neighborhood of any point in M, we can find a base of the distribution Δ, i.e., a vector field $V \in \operatorname{Vec} M$ such that $\Delta_q = \operatorname{span}(V(q))$, $q \in M$. Then trajectories of the ODE $\dot{q} = V(q)$ are one-dimensional submanifolds with tangent spaces Δ_q.

But in the general case ($\dim \Delta_q > 1$), a distribution Δ may be nonintegrable. Indeed, consider the family of vector fields tangent to Δ:

$$\overline{\Delta} = \{V \in \operatorname{Vec} M \mid V(q) \in \Delta_q \quad \forall\, q \in M\}.$$

Assume that the distribution Δ is integrable. Any vector field from the family $\overline{\Delta}$ is tangent to integral manifolds N_q, thus the orbit $\mathcal{O}_q$ of the family $\overline{\Delta}$ restricted to a small enough neighborhood of q is contained in the integral manifold N_q. Moreover, since $\dim \mathcal{O}_q \geq \dim \Delta_q = \dim N_q$, then locally $\mathcal{O}_q = N_q$: we can go in N_q in any direction along vector fields of the family $\overline{\Delta}$. By the Orbit Theorem, $T_q\mathcal{O}_q \supset \operatorname{Lie}_q \overline{\Delta}$, that is why

$$\operatorname{Lie}_q \overline{\Delta} = \Delta_q.$$

This means that

$$[V_1, V_2] \in \overline{\Delta} \quad \forall\, V_1,\ V_2 \in \overline{\Delta}. \tag{5.12}$$

Let $\dim \Delta_q = k$. In a neighborhood O_{q_0} of a point $q_0 \in M$ we can find a base of the distribution Δ:

$$\Delta_q = \operatorname{span}(f_1(q), \ldots, f_k(q)) \quad \forall\, q \in O_{q_0}.$$

Then condition (5.12) reads as *Frobenius condition*:

$$[f_i, f_j] = \sum_{l=1}^{k} c_{ij}^l f_l, \quad c_{ij}^l \in C^\infty(O_{q_0}). \tag{5.13}$$

We have shown that integrability of a distribution Δ implies Frobenius condition for its base.

Conversely, if condition (5.13) holds in a neighborhood of any point $q_0 \in M$, then $\mathrm{Lie}(\overline{\Delta}) = \overline{\Delta}$. Thus $\mathrm{Lie}(\overline{\Delta})$ is a locally finitely generated module over $C^\infty(M)$. By Theorem 5.3,

$$T_q \mathcal{O}_{q_0} = \mathrm{Lie}_q \overline{\Delta}, \qquad q \in \mathcal{O}_{q_0}.$$

So

$$T_q \mathcal{O}_{q_0} = \Delta_q, \qquad q \in \mathcal{O}_{q_0},$$

i.e., the orbit $\mathcal{O}_{q_0}$ is an integral manifold of Δ through q_0. We proved the following proposition.

Theorem 5.4 (Frobenius). *A distribution $\Delta \subset TM$ is integrable if and only if Frobenius condition* (5.13) *holds for any base of Δ in a neighborhood of any point $q_0 \in M$.*

Remarks. (1) In view of the Leibniz rule

$$[f, ag] = (fa)g + a[f, g], \qquad f,\ g \in \mathrm{Vec}\, M, \quad a \in C^\infty(M),$$

Frobenius condition is independent on the choice of a base $f_1, \ldots, f_k$: if it holds in one base, then it also holds in any other base.

(2) One can prove analogously the generalized Frobenius theorem for "distributions with variable $\dim \Delta_q$". If we do not demand that $\dim \Delta_q$ is constant, then Frobenius condition implies integrability; but dimension of integrable manifolds becomes, in general, different, although it stays constant along orbits of $\overline{\Delta}$. This is a generalization of phase portraits of vector fields. Although, notice once more that in general distributions with $\dim \Delta_q > 1$ are nonintegrable.

5.7 State equivalence of control systems

In this section we consider one more application of the Orbit Theorem — to the problem of equivalence of control systems (or families of vector fields).

Let U be an arbitrary index set. Consider two families of vector fields on smooth manifolds M and N parametrized by the same set U:

$$f_U = \{f_u \mid u \in U\} \subset \mathrm{Vec}\, M,$$
$$g_U = \{g_u \mid u \in U\} \subset \mathrm{Vec}\, N.$$

Take any pair of points $x_0 \in M$, $y_0 \in N$, and assume that the families f_U, g_U are bracket-generating:

$$\mathrm{Lie}_{x_0} f_U = T_{x_0} M, \qquad \mathrm{Lie}_{y_0} g_U = T_{y_0} N.$$

Definition 5.5. Families f_U and g_U are called *locally state equivalent* if there exists a local diffeomorphism

$$\begin{aligned} \Phi &: O_{x_0} \subset M \to O_{y_0} \subset N, \\ \Phi &: x_0 \mapsto y_0, \end{aligned}$$

that transforms one family to another:

$$\Phi_* f_u = g_u \qquad \forall\ u \in U.$$

Notation: $(f_U, x_0) \simeq (g_U, y_0)$.

Remark. Here we consider only smooth transformations of state $x \mapsto y$, while the controls u do not change. That is why this kind of equivalence is called state equivalence. We already studied state equivalence of nonlinear and linear systems, both local and global, see Chapter 4.

Now, we first try to find necessary conditions for local equivalence of systems f_U and g_U. Assume that

$$(f_U, x_0) \simeq (g_U, y_0).$$

By invariance of Lie bracket, we get

$$\Phi_*[f_{u_1}, f_{u_2}] = [\Phi_* f_{u_1}, \Phi_* f_{u_2}] = [g_{u_1}, g_{u_2}], \qquad u_1,\ u_2 \in U,$$

i.e., relations between Lie brackets of vector fields of the equivalent families f_U and g_U must be preserved. We collect all relations between these Lie brackets at one point: define the systems of tangent vectors

$$\begin{aligned} \xi_{u_1 \dots u_k} &= [f_{u_1}, [\dots, f_{u_k}] \dots](x_0) \in T_{x_0} M, \\ \eta_{u_1 \dots u_k} &= [g_{u_1}, [\dots, g_{u_k}] \dots](y_0) \in T_{y_0} N. \end{aligned}$$

Then we have

$$\Phi_*\big|_{x_0} \xi_{u_1 \dots u_k} = \eta_{u_1 \dots u_k} \quad u_1, \dots, u_k \in U,\ k \in \mathbb{N}.$$

Now we can state a necessary condition for local equivalence of families f_U and g_U in terms of the linear isomorphism

$$\Phi_*\big|_{x_0} = A\ :\ T_{x_0} M \leftrightarrow T_{y_0} N.$$

If $f_U \simeq g_U$, then there exists a linear isomorphism

$$A\ :\ T_{x_0} M \leftrightarrow T_{y_0} N$$

that maps the configuration of vectors $\{\xi_{u_1 \dots u_k}\}$ to the configuration $\{\eta_{u_1 \dots u_k}\}$. It turns out that in the analytic case this condition is sufficient. I.e., in the analytic case the combinations of partial derivatives of vector fields f_u, $u \in U$, that enter $\{\xi_{u_1 \dots u_k}\}$, form a complete system of state invariants of a family f_U.

Theorem 5.5. *Let f_U and g_U be real analytic and bracket-generating families of vector fields on real analytic manifolds M and N respectively. Let $x_0 \in M$, $y_0 \in N$. Then $(f_U, x_0) \simeq (g_U, y_0)$ if and only if there exists a linear isomorphism*

$$A \ : \ T_{x_0}M \leftrightarrow T_{y_0}N$$

such that

$$A\{\xi_{u_1 \dots u_k}\} = \{\eta_{u_1 \dots u_k}\} \qquad \forall \ u_1, \dots, u_k \in U, \quad k \in \mathbb{N}. \tag{5.14}$$

Remark. If in addition M, N are simply connected and the fields f_U, g_U are complete, then we have the global equivalence.

Before proving Theorem 5.5, we reformulate condition (5.14) and provide a method to check it.

Let a family f_U be bracket-generating:

$$\operatorname{span}\{\xi_{u_1 \dots u_k} \mid u_1, \dots, u_k \in U, \ k \in \mathbb{N}\} = T_{x_0}M.$$

We can choose a basis:

$$\operatorname{span}(\xi_{\bar{\alpha}_1}, \dots, \xi_{\bar{\alpha}_n}) = T_{x_0}M, \qquad \bar{\alpha}_i = (u_{1i}, \dots, u_{ki}), \quad i = 1, \dots, n, \tag{5.15}$$

and express all vectors in the configuration ξ through the base vectors:

$$\xi_{u_1 \dots u_k} = \sum_{i=1}^{n} c^i_{u_1 \dots u_k} \xi_{\bar{\alpha}_i}. \tag{5.16}$$

If there exists a linear isomorphism $A \ : \ T_{x_0}M \leftrightarrow T_{y_0}N$ with (5.14), then the vectors

$$\eta_{\bar{\alpha}_i}, \qquad i = 1, \dots, n,$$

should form a basis of $T_{y_0}N$:

$$\operatorname{span}(\eta_{\bar{\alpha}_1}, \dots, \eta_{\bar{\alpha}_n}) = T_{y_0}N, \tag{5.17}$$

and all vectors of the configuration η should be expressed through the base vectors with the same coefficients as the configuration ξ, see (5.16):

$$\eta_{u_1 \dots u_k} = \sum_{i=1}^{n} c^i_{u_1 \dots u_k} \eta_{\bar{\alpha}_i}. \tag{5.18}$$

It is easy to see the converse implication: if we can choose bases in $T_{x_0}M$ and $T_{y_0}N$ from the configurations ξ and η as in (5.15) and (5.17) such that decompositions (5.16) and (5.18) with the same coefficients $c^i_{u_1 \dots u_k}$ hold, then there exists a linear isomorphism A with (5.14). Indeed, we define then the isomorphism on the bases:

$$A \ : \ \xi_{\bar{\alpha}_i} \mapsto \eta_{\bar{\alpha}_1}, \qquad i = 1, \dots, n.$$

We can obtain one more reformulation via the following agreement. Configurations $\{\xi_{u_1\dots u_k}\}$ and $\{\eta_{u_1\dots u_k}\}$ are called equivalent if the sets of relations $K(f_U)$ and $K(g_U)$ between elements of these configurations coincide: $K(f_U) = K(g_U)$. We denote here by $K(f_U)$ the set of all systems of coefficients such that the corresponding linear combinations vanish:

$$K(f_U) = \left\{ (b_{u_1\dots u_k}) \mid \sum_{u_1\dots u_k} b_{u_1\dots u_k} \xi_{u_1\dots u_k} = 0 \right\}.$$

Then Theorem 5.5 can be expressed in the following form.

Nagano Principle. *All local information about bracket-generating families of analytic vector fields is contained in Lie brackets.*

Notice, although, that the configuration $\xi_{u_1\dots u_k}$ or the system of relations $K(f_U)$ are, in general, immense and cannot be easily characterized. Thus Nagano Principle cannot usually be applied directly to describe properties of control systems, but it is an important guiding principle.

Now we prove Theorem 5.5.

Proof. Necessity was already shown. We prove sufficiency by reduction to the Orbit Theorem. For this we construct an auxiliary system on the Cartesian product

$$M \times N = \{(x, y) \mid x \in M, \ y \in N\}.$$

For vector fields $f \in \operatorname{Vec} M$, $g \in \operatorname{Vec} N$, define their direct product $f \times g \in \operatorname{Vec}(M \times N)$ as the derivation

$$(f \times g) a\big|_{(x,y)} = (f a_y^1)\big|_x + (g a_x^2)\big|_y , \qquad a \in C^\infty(M \times N), \tag{5.19}$$

where the families of functions $a_y^1 \in C^\infty(M)$, $a_x^2 \in C^\infty(N)$ are defined as follows:

$$a_y^1 \ : \ x \mapsto a(x, y), \quad a_x^2 \ : \ y \mapsto a(x, y), \qquad x \in M, \ y \in N.$$

So projection of $f \times g$ to M is f, and projection to N is g. Finally, we define the direct product of systems f_U and g_U as

$$f_U \times g_U = \{f_u \times g_u \mid u \in U\} \subset \operatorname{Vec}(M \times N).$$

We suppose that there exists a linear isomorphism $A \ : \ T_{x_0} M \leftrightarrow T_{y_0} N$ that maps the configuration ξ to η as in (5.14), and construct the local equivalence $f_U \simeq g_U$.

In view of definition (5.19), Lie bracket in the family $f_U \times g_U$ is computed as

$$[f_{u_1} \times g_{u_1}, f_{u_2} \times g_{u_2}] = [f_{u_1}, f_{u_2}] \times [g_{u_1}, g_{u_2}], \qquad u_1, \ u_2 \in U,$$

thus

$$\begin{aligned}
&[f_{u_1} \times g_{u_1}, [\dots, f_{u_k} \times g_{u_k}] \dots](x_0, y_0) \\
&\quad = [f_{u_1}, [\dots, f_{u_k}] \dots](x_0) \times [g_{u_1}, [\dots, g_{u_k}] \dots](y_0) \\
&\quad = \xi_{u_1\dots u_k} \times \eta_{u_1\dots u_k} = \xi_{u_1\dots u_k} \times A\xi_{u_1\dots u_k}, \qquad u_1, \dots, u_k \in U, \ k \in \mathbb{N}.
\end{aligned}$$

That is why

$$\dim \operatorname{Lie}_{(x_0,y_0)}(f_U \times g_U) = n,$$

where $n = \dim M$. By the analytic version of the Orbit Theorem (Corollary 5.3) for the family $f_U \times g_U \subset \operatorname{Vec}(M \times N)$, the orbit $\mathcal{O}$ of $f_U \times g_U$ through the point (x_0, y_0) is an n-dimensional immersed submanifold (thus, locally a submanifold) of $M \times N$. The tangent space of the orbit is

$$\begin{aligned} T_{(x_0,y_0)}\mathcal{O} &= \operatorname{span}(\xi_{u_1\dots u_k} \times A\xi_{u_1\dots u_k}) \\ &= \operatorname{span}\{v \times Av \mid v \in T_{x_0}\} \subset T_{(x_0,y_0)}M \times N = T_{x_0}M \times T_{y_0}N, \end{aligned}$$

i.e., the graph of the linear isomorphism A. Consider the canonical projections onto the factors:

$$\begin{aligned} \pi_1 &: M \times N \to M, \qquad \pi_1(x,y) = x, \\ \pi_2 &: M \times N \to N, \qquad \pi_2(x,y) = y. \end{aligned}$$

The restrictions $\pi_1|_{\mathcal{O}}$, $\pi_2|_{\mathcal{O}}$ are local diffeomorphisms since the differentials

$$\begin{aligned} \pi_{1*}|_{(x_0,y_0)} &: (v, Av) \mapsto v, \qquad v \in T_{x_0}M, \\ \pi_{2*}|_{(x_0,y_0)} &: (v, Av) \mapsto Av, \qquad v \in T_{x_0}M, \end{aligned}$$

are one-to-one.

Now $\Phi = \pi_2 \circ (\pi_1|_{\mathcal{O}})^{-1}$ is a local diffeomorphism from M to N with the graph $\mathcal{O}$, and

$$\Phi_* = \pi_{2*} \circ (\pi_1|_{\mathcal{O}})_*^{-1} : f_u \mapsto g_u, \qquad u \in U.$$

Consequently, $(f_U, x_0) \simeq (g_U, y_0)$. $\square$

Chapter 6

Rotations of the rigid body

In this chapter we consider rotations of a *rigid body* around a fixed point. That is, we study motions of a body in the three-dimensional space such that:

- distances between all points in the body remain fixed (rigidity), and
- there is a point in the body that stays immovable during motion (fixed point).

We consider both free motions (in the absence of external forces) and controlled motions (when external forces are applied in order to bring the body to a desired state).

Such system is a very simplified model of a satellite in the space rotating around its center of mass.

For details about ODEs describing rotations of the rigid body, see [4].

6.1 State space

The state of the rigid body is determined by its position and velocity.

We fix an orthonormal frame attached to the body at the fixed point (the moving frame), and an orthonormal frame attached to the ambient space at the fixed point of the body (the fixed frame). The set of positions of the rigid body is the set of all orthonormal frames in the three-dimensional space with positive orientation. This set can be identified with SO(3), the group of linear orthogonal orientation-preserving transformations of $\mathbb{R}^3$, or, equivalently, with the group of 3×3 orthogonal unimodular matrices:

$$\begin{aligned} \mathrm{SO}(3) &= \{Q : \mathbb{R}^3 \to \mathbb{R}^3 \mid (Qx, Qy) = (x, y),\ \det Q = 1\} \\ &= \{Q : \mathbb{R}^3 \to \mathbb{R}^3 \mid QQ^* = \mathrm{Id},\ \det Q = 1\}. \end{aligned}$$

The mapping $Q : \mathbb{R}^3 \to \mathbb{R}^3$ transforms the coordinate representation of a point in the moving frame to the coordinate representation of this point in the fixed frame.

Remark. We denote above the standard inner product in $\mathbb{R}^3$ by $(\cdot,\cdot)$. If a pair of vectors $x, y \in \mathbb{R}^3$ have coordinates $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$ in some orthonormal frame, then $(x, y) = x_1 y_1 + x_2 y_2 + x_3 y_3$.

Notice that the set of positions of the rigid body $\mathrm{SO}(3)$ is not a linear space, but a nontrivial smooth manifold.

Now we describe velocities of the rigid body. Let $Q_t \in \mathrm{SO}(3)$ be position of the body at a moment of time t. Since the operators $Q_t : \mathbb{R}^3 \to \mathbb{R}^3$ are orthogonal, then

$$(Q_t x, Q_t y) = (x, y), \qquad x,\ y \in \mathbb{R}^3, \quad t \in \mathbb{R}.$$

We differentiate this equality w.r.t. t and obtain

$$(\dot{Q}_t x, Q_t y) + (Q_t x, \dot{Q}_t y) = 0. \tag{6.1}$$

The matrix

$$\Omega_t = Q_t^{-1} \dot{Q}_t$$

is called the *angular velocity* of the body in the moving frame. Since

$$\dot{Q}_t = Q_t \Omega_t,$$

then equality (6.1) reads

$$(Q_t \Omega_t x, Q_t y) + (Q_t x, Q_t \Omega_t y) = 0,$$

whence by orthogonality

$$(\Omega_t x, y) + (x, \Omega_t y) = 0,$$

i.e.,

$$\Omega_t^* = -\Omega_t,$$

the matrix Ω_t is antisymmetric. So velocities of the rigid body have the form

$$\dot{Q}_t = Q_t \Omega_t, \qquad \Omega_t^* = -\Omega_t.$$

In other words, we found the tangent space

$$T_Q \,\mathrm{SO}(3) = \{ Q\Omega \mid \Omega^* = -\Omega \}, \qquad Q \in \mathrm{SO}(3).$$

The space of antisymmetric 3×3 matrices is denoted by $\mathrm{so}(3)$, it is the tangent space to $\mathrm{SO}(3)$ at the identity:

$$\mathrm{so}(3) = \{ \Omega : \mathbb{R}^3 \to \mathbb{R}^3 \mid \Omega^* = -\Omega \} = T_{\mathrm{Id}} \,\mathrm{SO}(3).$$

The space $\mathrm{so}(3)$ is the Lie algebra of the Lie group $\mathrm{SO}(3)$.

To each antisymmetric matrix $\Omega \in \mathrm{so}(3)$, we associate a vector $\omega \in \mathbb{R}^3$:

$$\Omega \sim \omega, \qquad \Omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}, \quad \omega = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}. \tag{6.2}$$

Then the action of the operator Ω on a vector $x \in \mathbb{R}^3$ can be represented via the cross product in $\mathbb{R}^3$:

$$\Omega x = \omega \times x, \qquad x \in \mathbb{R}^3.$$

Let x be a point in the rigid body. Then its position in the ambient space $\mathbb{R}^3$ is $Q_t x$. Further, velocity of this point is

$$\dot{Q}_t x = Q_t \Omega_t x = Q_t(\omega_t \times x).$$

ω_t is the vector of angular velocity of the point x in the moving frame: if we fix the moving frame Q_t at one moment of time t, then the instantaneous velocity of the point x at the moment of time t in the moving frame is $Q_t^{-1}\dot{Q}_t x = \Omega_t x = \omega_t \times x$, i.e., the point x rotates around the line through ω_t with the angular velocity $\|\omega_t\|$.

Introduce the following scalar product of matrices $\Omega = (\Omega_{ij}) \in \mathrm{so}(3)$:

$$\langle \Omega^1, \Omega^2 \rangle = -\frac{1}{2}\operatorname{tr}(\Omega^1\Omega^2) = \frac{1}{2}\sum_{i,j=1}^{3} \Omega^1_{ij}\Omega^2_{ij} = \sum_{i<j} \Omega^1_{ij}\Omega^2_{ij}.$$

This product is compatible with identification of 3×3 antisymmetric matrices and 3-dimensional vectors (6.2):

$$\begin{aligned} &\langle \Omega^1, \Omega^2 \rangle = (\omega^1, \omega^2), \\ &\Omega^i \sim \omega^i, \qquad \Omega^i \in \mathrm{so}(3), \quad \omega^i \in \mathbb{R}^3, \quad i = 1, 2. \end{aligned}$$

Moreover, this product is invariant in the following sense:

$$\langle (\operatorname{Ad} Q)\Omega^1, (\operatorname{Ad} Q)\Omega^2 \rangle = \langle \Omega^1, \Omega^2 \rangle, \qquad Q \in \mathrm{SO}(3), \quad \Omega^1,\ \Omega^2 \in \mathrm{so}(3), \tag{6.3}$$

i.e., $\operatorname{Ad} Q : \mathrm{so}(3) \to \mathrm{so}(3)$ is an orthogonal transformation w.r.t. $\langle \cdot, \cdot \rangle$. Indeed:

$$\operatorname{tr}((\operatorname{Ad} Q)\Omega^1(\operatorname{Ad} Q)\Omega^2) = \operatorname{tr}(Q\Omega^1Q^{-1}Q\Omega^2Q^{-1}) = \operatorname{tr}(Q\Omega^1\Omega^2Q^{-1}) = \operatorname{tr}(\Omega^1\Omega^2)$$

by invariance of trace.

Now we derive the infinitesimal version of invariance (6.3). Take an arbitrary $\Omega \in \mathrm{so}(3)$ and consider a smooth curve $Q_t \in \mathrm{SO}(3)$ that starts from identity with the velocity Ω:

$$\dot{Q}_0 = \Omega, \qquad Q_0 = \operatorname{Id}.$$

Then

$$\left.\frac{d}{dt}\right|_0 \operatorname{Ad} Q_t = \operatorname{ad} \Omega,$$

and differentiation of (6.3) w.r.t. t at $t = 0$ yields the equality:

$$\langle (\operatorname{ad} \Omega)\Omega^1, \Omega^2 \rangle + \langle \Omega^1, (\operatorname{ad} \Omega)\Omega^2 \rangle = 0, \qquad \Omega,\ \Omega^1,\ \Omega^2 \in \mathrm{so}(3), \tag{6.4}$$

i.e., $\operatorname{ad} \Omega : \mathrm{so}(3) \to \mathrm{so}(3)$ is antisymmetric w.r.t. $\langle \cdot, \cdot \rangle$.

Matrix commutator in $\mathrm{so}(3)$ corresponds to vector product in $\mathbb{R}^3$, thus equality (6.4) can be rewritten in terms of cross product:

$$(\omega \times \omega^1, \omega^2) + (\omega^1, \omega \times \omega^2) = 0, \qquad \omega,\ \omega^1, \omega^2 \in \mathbb{R}^3.$$

6.2 Euler equations

We derive equations of motion of the rigid body from the least action principle.

Let the distribution of mass in the rigid body have density $\rho(x)$, where $\rho : \mathbb{R}^3 \to \mathbb{R}_+$ is an integrable nonnegative function with compact support. Let $Q_t \in \mathrm{SO}(3)$ be position and $\Omega_t \in \mathrm{so}(3)$ angular velocity of the body so that

$$\dot{Q}_t = Q_t \Omega_t. \tag{6.5}$$

Take a point x in the body. Then position of this point in the ambient space is $Q_t x$, and velocity of this point is $\dot{Q}_t x$. Distribution of the kinetic energy in the body has density $\frac{1}{2}\rho(x)(\dot{Q}_t x, \dot{Q}_t x)$, thus the total kinetic energy of the body at a moment of time t is

$$j(\Omega_t) = \frac{1}{2}\int_{\mathbb{R}^3} \rho(x)(Q_t\Omega_t x, Q_t\Omega_t x)\, dx = \frac{1}{2}\int_{\mathbb{R}^3} \rho(x)(\Omega_t x, \Omega_t x)\, dx,$$

i.e., a quadratic form $j = j(\Omega_t)$ on the space $\mathrm{so}(3)$. The corresponding bilinear form can be written as

$$\int_{\mathbb{R}^3} \rho(x)(\Omega^1 x, \Omega^2 x)\, dx = \langle A\Omega^1, \Omega^2\rangle, \qquad \Omega^1,\ \Omega^2 \in \mathrm{so}(3)$$

for some linear symmetric positive definite operator

$$A : \mathrm{so}(3) \to \mathrm{so}(3), \qquad A = A^* > 0,$$

called inertia tensor of the rigid body. Finally, the functional of action has the form

$$J(\Omega_\cdot) = \int_0^{t_1} j(\Omega_t)\, dt = \frac{1}{2}\int_0^{t_1} \langle A\Omega_t, \Omega_t\rangle\, dt,$$

where 0 and t_1 are the initial and terminal moments of motion.

Let Q_0 and Q_{t_1} be the initial and terminal positions of the moving body. By the least action principle, the motion Q_t, $t \in [0, t_1]$, of the body should be an extremal of the following problem:

$$\begin{aligned} &J(\Omega_\cdot) \to \min, \\ &\dot{Q}_t = Q_t \Omega_t, \qquad Q_0,\ Q_{t_1} \text{ fixed}. \end{aligned} \tag{6.6}$$

We find these extremals.

Let Ω_t be angular velocity along the reference trajectory Q_t, then

$$Q_0^{-1} \circ Q_{t_1} = \overrightarrow{\exp} \int_0^{t_1} \Omega_t\, dt.$$

Consider an arbitrary small perturbation of the angular velocity:

$$\Omega_t + \varepsilon U_t + O(\varepsilon^2), \qquad \varepsilon \to 0.$$

In order that such perturbation was admissible, the starting point and endpoint of the corresponding trajectory should not depend on ε:

$$Q_0^{-1} \circ Q_{t_1} = \overrightarrow{\exp} \int_0^{t_1} \left(\Omega_t + \varepsilon U_t + O(\varepsilon^2)\right) \, dt,$$

thus

$$0 = \left.\frac{\partial}{\partial \varepsilon}\right|_{\varepsilon=0} Q_0^{-1} \circ Q_{t_1} = \left.\frac{\partial}{\partial \varepsilon}\right|_{\varepsilon=0} \overrightarrow{\exp} \int_0^{t_1} \left(\Omega_t + \varepsilon U_t + O(\varepsilon^2)\right) \, dt. \tag{6.7}$$

By formula (2.31) of derivative of a flow w.r.t. parameter, the right- hand side above is equal to

$$\begin{aligned} &\int_0^{t_1} \mathrm{Ad}\left(\overrightarrow{\exp} \int_0^t \Omega_\tau \, d\tau\right) U_t \, dt \circ \overrightarrow{\exp} \int_0^{t_1} \Omega_t \, dt \\ &= \int_0^{t_1} \mathrm{Ad}\left(Q_0^{-1} \circ Q_t\right) U_t \, dt \circ Q_0^{-1} \circ Q_{t_1} \\ &= Q_0^{-1} \int_0^{t_1} \mathrm{Ad}\, Q_t U_t \, dt \circ Q_{t_1}. \end{aligned}$$

Taking into account (6.7), we obtain

$$\int_0^{t_1} \mathrm{Ad}\, Q_t U_t \, dt = 0.$$

Denote

$$V_t = \int_0^t \mathrm{Ad}\, Q_\tau U_\tau \, d\tau, \tag{6.8}$$

then admissibility condition of a variation U_t takes the form

$$V_0 = V_{t_1} = 0. \tag{6.9}$$

Now we find extremals of problem (6.6).

$$0 = \left.\frac{\partial}{\partial \varepsilon}\right|_{\varepsilon=0} J(\Omega_{\cdot\varepsilon}) = \int_0^{t_1} \langle A\Omega_t, U_t \rangle \, dt$$

by (6.3)

$$= \int_0^{t_1} \langle (\mathrm{Ad}\, Q_t) A\Omega_t, (\mathrm{Ad}\, Q_t) U_t \rangle \, dt$$

by (6.8)

$$= \int_0^{t_1} \langle (\mathrm{Ad}\, Q_t) A\Omega_t, \dot{V}_t \rangle \, dt$$

integrating by parts with the admissibility condition (6.9)

$$= -\int_0^{t_1} \left\langle \frac{d}{d\,t}(\operatorname{Ad} Q_t) A\Omega_t, V_t \right\rangle dt.$$

So the previous integral vanishes for any admissible operator V_t, thus

$$\frac{d}{d\,t}(\operatorname{Ad} Q_t) A\Omega_t = 0, \qquad t \in [0, t_1].$$

Hence

$$\operatorname{Ad} Q_t([\Omega_t, A\Omega_t] + A\dot{\Omega}_t) = 0, \qquad t \in [0, t_1],$$

that is why

$$A\dot{\Omega}_t = [A\Omega_t, \Omega_t], \qquad t \in [0, t_1]. \tag{6.10}$$

Introduce the operator

$$M_t = A\Omega_t,$$

called kinetic momentum of the body, and denote

$$B = A^{-1}.$$

We combine equations (6.10), (6.5) and come to *Euler equations* of rotations of a free rigid body:

$$\begin{cases} \dot{M}_t = [M_t, BM_t], & M_t \in \operatorname{so}(3), \\ \dot{Q}_t = Q_t B M_t, & Q_t \in \operatorname{SO}(3). \end{cases}$$

Remark. The presented way to derive Euler equations can be applied to the curves on the group $\operatorname{SO}(n)$ of orthogonal orientation-preserving $n \times n$ matrices with an arbitrary $n > 0$. Then we come to equations of rotations of a generalized n-dimensional rigid body.

Now we rewrite Euler equations via isomorphism (6.2) of $\operatorname{so}(3)$ and $\mathbb{R}^3$, which is essentially 3-dimensional and does not generalize to higher dimensions. Recall that for an antisymmetric matrix

$$M = \begin{pmatrix} 0 & -\mu_3 & \mu_2 \\ \mu_3 & 0 & -\mu_1 \\ -\mu_2 & \mu_1 & 0 \end{pmatrix} \in \operatorname{so}(3),$$

the corresponding vector $\mu \in \mathbb{R}^3$ is

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \qquad M \sim \mu.$$

Now Euler equations read as follows:

$$\begin{cases} \dot{\mu}_t = \mu_t \times \beta\mu_t, & \mu_t \in \mathbb{R}^3, \\ \dot{Q}_t = Q_t \widehat{\beta\mu_t}, & Q_t \in \operatorname{SO}(3), \end{cases}$$

where $\beta : \mathbb{R}^3 \to \mathbb{R}^3$ and $\widehat{\beta} : \mathbb{R}^3 \to \mathrm{so}(3)$ are the operators corresponding to $B : \mathrm{so}(3) \to \mathrm{so}(3)$ via the isomorphism $\mathrm{so}(3) \leftrightarrow \mathbb{R}^3$ (6.2).

Eigenvectors of the symmetric positive definite operator $\beta : \mathbb{R}^3 \to \mathbb{R}^3$ are called principal axes of inertia of the rigid body. In the sequel we assume that the rigid body is asymmetric, i.e., the operator β has 3 distinct eigenvalues $\lambda_1, \lambda_2, \lambda_3$. We order the eigenvalues of β:

$$\lambda_1 > \lambda_2 > \lambda_3,$$

and choose an orthonormal frame e_1, e_2, e_3 of the corresponding eigenvectors, i.e., principal axes of inertia. In the basis e_1, e_2, e_3, the operator β is diagonal:

$$\beta \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} \lambda_1\mu_1 \\ \lambda_2\mu_2 \\ \lambda_3\mu_3 \end{pmatrix},$$

and the equation $\dot{\mu}_t = \mu_t \times \beta\mu_t$ reads as follows:

$$\begin{cases} \dot{\mu}_1 = (\lambda_3 - \lambda_2)\mu_2\mu_3, \\ \dot{\mu}_2 = (\lambda_1 - \lambda_3)\mu_1\mu_3, \\ \dot{\mu}_3 = (\lambda_2 - \lambda_1)\mu_1\mu_2. \end{cases} \tag{6.11}$$

6.3 Phase portrait

Now we draw the phase portrait of the first of Euler equations:

$$\dot{\mu}_t = \mu_t \times \beta\mu_t, \qquad \mu_t \in \mathbb{R}^3. \tag{6.12}$$

This equation has two integrals: energy

$$(\mu_t, \mu_t) = \mathrm{const}$$

and moment of momentum

$$(\mu_t, \beta\mu_t) = \mathrm{const}\,.$$

Indeed:

$$\begin{aligned} \frac{d}{dt}(\mu_t, \mu_t) &= 2(\mu_t \times \beta\mu_t, \mu_t) = -2(\beta\mu_t, \mu_t \times \mu_t) = 0, \\ \frac{d}{dt}(\mu_t, \beta\mu_t) &= (\mu_t \times \beta\mu_t, \beta\mu_t) + (\mu_t, \beta(\mu_t \times \beta\mu_t)) = 2(\mu_t \times \beta\mu_t, \beta\mu_t) \\ &= -2(\mu_t, \beta\mu_t \times \beta\mu_t) = 0 \end{aligned}$$

by the invariance property (6.4) and symmetry of β.

So all trajectories μ_t of equation (6.12) satisfy the restrictions

$$\begin{cases} \mu_1^2 + \mu_2^2 + \mu_3^2 = \mathrm{const}, \\ \lambda_1\mu_1^2 + \lambda_2\mu_2^2 + \lambda_3\mu_3^2 = \mathrm{const}, \end{cases} \tag{6.13}$$

i.e., belong to intersection of spheres with ellipsoids. Moreover, since ODE (6.12) is homogeneous, we draw its trajectories on one sphere — the unit sphere

$$\mu_1^2 + \mu_2^2 + \mu_3^2 = 1, \tag{6.14}$$

and all other trajectories are obtained by homotheties.

First of all, intersections of the unit sphere with the principal axes of inertia, i.e., the points

$$\pm e_1, \ \pm e_2, \ \pm e_3$$

are equilibria, and there are no other equilibria, see equations (6.11).

Further, the equilibria $\pm e_1$, $\pm e_3$ corresponding to the maximal and minimal eigenvalues λ_1, λ_3 are stable, more precisely, they are centers, and the equilibria $\pm e_2$ corresponding to λ_2 are unstable — saddles. This is obvious from the geometry of intersections of the unit sphere with ellipsoids

$$\lambda_1 \mu_1^2 + \lambda_2 \mu_2^2 + \lambda_3 \mu_3^2 = C.$$

Indeed, for $C < \lambda_3$ the ellipsoids are inside the sphere and do not intersect it. For $C = \lambda_3$, the ellipsoid touches the unit sphere from inside at the points $\pm e_3$. Further, for $C > \lambda_3$ and close to λ_3, the ellipsoids intersect the unit sphere by 2 closed curves surrounding e_3 and $-e_3$ respectively. The behavior of intersections is similar in the neighborhood of $C = \lambda_1$. If $C > \lambda_1$, then the ellipsoids are big enough and do not intersect the unit sphere; for $C = \lambda_1$, the small semiaxis of the ellipsoid becomes equal to radius of the sphere, so the ellipsoid touches the sphere from outside at $\pm e_1$; and for $C < \lambda_1$ and close to λ_1 the intersection consists of 2 closed curves surrounding $\pm e_1$. If $C = \lambda_2$, then the ellipsoid touches the sphere at the endpoints of the medium semiaxes $\pm e_2$, and in the neighborhood of each point e_2, $-e_2$, the intersection consists of four separatrix branches tending to this point. Equations for the separatrices are derived from the system

$$\begin{cases} \mu_1^2 + \mu_2^2 + \mu_3^2 = 1, \\ \lambda_1 \mu_1^2 + \lambda_2 \mu_2^2 + \lambda_3 \mu_3^2 = \lambda_2. \end{cases}$$

We multiply the first equation by λ_2 and subtract it from the second equation:

$$(\lambda_1 - \lambda_2)\mu_1^2 - (\lambda_2 - \lambda_3)\mu_3^2 = 0.$$

Thus the separatrices belong to intersection of the unit sphere with two planes

$$\Pi_\pm \stackrel{\text{def}}{=} \{(\mu_1, \mu_2, \mu_3) \in \mathbb{R}^3 \mid \sqrt{\lambda_1 - \lambda_2}\, \mu_1 = \pm\sqrt{\lambda_2 - \lambda_3}\, \mu_3\},$$

thus they are arcs of great circles.

It turns out that separatrices and equilibria are the only trajectories belonging to a 2-dimensional plane. Moreover, all other trajectories satisfy the following condition:

$$\mu \notin \Pi_\pm, \mu \notin \mathbb{R}e_i \quad \Rightarrow \quad \mu \wedge \dot{\mu} \wedge \ddot{\mu} \neq 0. \tag{6.15}$$

Indeed, take any trajectory μ_t on the unit sphere. All trajectories homothetic to the chosen one form a cone of the form

$$C(\mu_1^2 + \mu_2^2 + \mu_3^2) = \lambda_1 \mu_1^2 + \lambda_2 \mu_2^2 + \lambda_3 \mu_3^2, \qquad \lambda_3 \leq C \leq \lambda_1. \tag{6.16}$$

But a quadratic cone in $\mathbb{R}^3$ is either degenerate or elliptic. The conditions $\mu \notin \Pi_\pm$, $\mu \notin \mathbb{R}e_i$ mean that $C \neq \lambda_i$, $i = 1, 2, 3$, i.e., cone (6.16) is elliptic. Now inequality (6.15) follows from the next two facts. First, $\mu \wedge \dot{\mu} \neq 0$, i.e., the trajectory μ_t is not tangent to the generator of the cone. Second, the section of an elliptic cone by a plane not containing the generator of the cone is an ellipse — a strictly convex curve.

In view of ODE (6.12), the convexity condition (6.15) for the cone generated by the trajectory is rewritten as follows:

$$\mu \notin \Pi_\pm, \mu \notin \mathbb{R}e_i \Rightarrow \mu \wedge (\mu \times \beta\mu) \wedge ((\mu \times \beta\mu) \times \beta\mu + \mu \times \beta(\mu \times \beta\mu)) \neq 0. \tag{6.17}$$

The planar separatrix curves in the phase portrait are regular curves on the sphere, hence

$$\mu \in \Pi_\pm, \mu \notin \mathbb{R}e_2 \Rightarrow \mu \wedge \dot{\mu} \neq 0,$$

or, by ODE (6.12),

$$\mu \in \Pi_\pm, \mu \notin \mathbb{R}e_2 \Rightarrow \mu \wedge (\mu \times \beta\mu) \neq 0. \tag{6.18}$$

6.4 Controlled rigid body: orbits

Assume that we can control rotations of the rigid body by applying a torque along a line that is fixed in the body. We can change the direction of torque to the opposite one in any moment of time.

Then the control system for the angular velocity is written as

$$\dot{\mu}_t = \mu_t \times \beta\mu_t \pm l, \qquad \mu_t \in \mathbb{R}^3, \tag{6.19}$$

and the whole control system for the controlled rigid body is

$$\begin{cases} \dot{\mu}_t = \mu_t \times \beta\mu_t \pm l, & \mu_t \in \mathbb{R}^3, \\ \dot{Q}_t = Q_t \widehat{\beta\mu_t}, & Q_t \in \mathrm{SO}(3), \end{cases} \tag{6.20}$$

where $l \neq 0$ is a fixed vector along the chosen line.

Now we describe orbits and attainable sets of the 6-dimensional control system (6.20). But before that we study orbits of the 3-dimensional system (6.19).

6.4.1 Orbits of the 3-dimensional system

System (6.19) is analytic, thus dimension of the orbit through a point $\mu \in \mathbb{R}^3$ coincides with dimension of the space

$$\mathrm{Lie}_\mu(\mu \times \beta\mu \pm l) = \mathrm{Lie}_\mu(\mu \times \beta\mu, l).$$

Denote the vector fields:

$$f(\mu) = \mu \times \beta\mu, \qquad g(\mu) \equiv l,$$

and compute several Lie brackets:

$$\begin{aligned}
&[g, f](\mu) = \frac{d\,f}{d\,\mu} g(\mu) - \frac{d\,g}{d\,\mu} f(\mu) = l \times \beta\mu + \mu \times \beta l, \\
&[g, [g, f]](\mu) = l \times \beta l + l \times \beta l = 2l \times \beta l, \\
&\frac{1}{2}[[g, [g, f]], [g, f]](\mu) = l \times \beta(l \times \beta l) + (l \times \beta l) \times \beta l.
\end{aligned}$$

We apply (6.17) with $l = \mu$ and obtain that three constant vector fields g, $[g, f]$, $[[g, [g, f]], [g, f]]$ are linearly independent:

$$\begin{aligned}
g(\mu) \wedge \frac{1}{2}[g, f](\mu) \wedge \frac{1}{2}[[g, [g, f]], [g, f]](\mu) \\
= l \wedge l \times \beta l \wedge ((l \times \beta l) \times \beta l + l \times \beta(l \times \beta l)) \neq 0
\end{aligned}$$

if $l \notin \Pi_\pm$, $l \notin \mathbb{R}e_i$.

We obtain the following statement for the generic

Case 1. $l \notin \Pi_\pm$, $l \notin \mathbb{R}e_i$.

Proposition 6.1. *Assume that $l \notin \Pi_\pm$, $l \notin \mathbb{R}e_i$. Then $\mathrm{Lie}_\mu(f, g) = \mathbb{R}^3$ for any $\mu \in \mathbb{R}^3$. System (6.19) has one 3-dimensional orbit, $\mathbb{R}^3$.*

Now consider special dispositions of the vector l.

Case 2. Let $l \in \Pi_+$, $l \notin \mathbb{R}e_2$. Since the plane Π_+ is invariant for the free body (6.12) and $l \in \Pi_+$, then the plane Π_+ is also invariant for the controlled body (6.19), i.e., the orbit through any point of Π_+ is contained in Π_+. On the other hand, implication (6.18) yields

$$l \wedge (l \times \beta l) \neq 0.$$

But the vectors $l = g(\mu)$ and $l \times \beta l = \frac{1}{2}[g, [g, f]](\mu)$ form a basis of the plane Π_+, thus Π_+ is in the orbit through any point $\mu \in \Pi_+$. Consequently, the plane Π_+ is an orbit of (6.19). If an initial point $\mu_0 \notin \Pi_+$, then the trajectory μ_t of (6.19) through μ_0 is not flat, thus

$$(\mu_t \times \beta\mu_t) \wedge l \wedge (l \times \beta l) \not\equiv 0.$$

So the orbit through μ_0 is 3-dimensional. We proved the following statement.

Proposition 6.2. *Assume that $l \in \Pi_+ \setminus \mathbb{R}e_2$. Then system (6.19) has one 2-dimensional orbit, the plane Π_+, and two 3-dimensional orbits, connected components of $\mathbb{R}^3 \setminus \Pi_+$.*

The case $l \in \Pi_- \setminus \mathbb{R}e_2$ is completely analogous, and there holds a similar proposition with Π_+ replaced by Π_-.

Case 3. Now let $l \in \mathbb{R}e_1 \setminus \{0\}$, i.e., $l = ce_1$, $c \neq 0$. First of all, the line $\mathbb{R}e_1$ is an orbit. Indeed, if $\mu \in \mathbb{R}e_1$, then $f(\mu) = 0$, and $g(\mu) = l$ is also tangent to the line $\mathbb{R}e_1$.

To find other orbits, we construct an integral of the control system (6.19) from two integrals (6.13) of the free body. Since $g(\mu) = l = ce_1$, we seek for a linear combination of the integrals in (6.13) that does not depend on μ_1. We multiply the first integral by λ_1, subtract from it the second integral and obtain an integral for the controlled rigid body:

$$(\lambda_1 - \lambda_2)\mu_2^2 + (\lambda_1 - \lambda_3)\mu_3^2 = C. \tag{6.21}$$

Since $\lambda_1 > \lambda_2 > \lambda_3$, this is an elliptic cylinder in $\mathbb{R}^3$.

So each orbit of (6.19) is contained in a cylinder (6.21). On the other hand, the orbit through any point $\mu_0 \in \mathbb{R}^3 \setminus \mathbb{R}e_1$ must be at least 2-dimensional. Indeed, if $\mu_0 \notin \mathbb{R}e_2, \mathbb{R}e_3$, then the free body has trajectories not tangent to the field g; and if $\mu_0 \in \mathbb{R}e_2$ or $\mathbb{R}e_3$, this can be achieved by a small translation of μ_0 along the field g. Thus all orbits outside of the line $\mathbb{R}e_1$ are elliptic cylinders (6.21).

Proposition 6.3. *Let $l \in \mathbb{R}e_1 \setminus \{0\}$. Then all orbits of system* (6.19) *have the form* (6.21)*: there is one* 1*-dimensional orbit — the line $\mathbb{R}e_1$ ($C = 0$), and an infinite number of* 2*-dimensional orbits — elliptic cylinders* (6.21) *with $C > 0$.*

The case $l \in \mathbb{R}e_3 \setminus \{0\}$ is completely analogous to the previous one.

Proposition 6.4. *Let $l \in \mathbb{R}e_3 \setminus \{0\}$. Then system* (6.19) *has one* 1*-dimensional orbit — the line $\mathbb{R}e_3$, and an infinite number of* 2*-dimensional orbits — elliptic cylinders*

$$(\lambda_1 - \lambda_3)\mu_1^2 + (\lambda_2 - \lambda_3)\mu_2^2 = C, \qquad C > 0.$$

Case 4. Finally, consider the most complicated case: let $l \in \mathbb{R}e_2 \setminus \{0\}$. As above, we obtain an integral of control system (6.19):

$$(\lambda_1 - \lambda_2)\mu_1^2 - (\lambda_2 - \lambda_3)\mu_3^2 = C. \tag{6.22}$$

If $C \neq 0$, this equation determines a hyperbolic cylinder. By an argument similar to that used in Case 3, we obtain the following description of orbits.

Proposition 6.5. *Let $l \in \mathbb{R}e_2 \setminus \{0\}$. Then there is one* 1*-dimensional orbit — the line $\mathbb{R}e_2$, and an infinite number of* 2*-dimensional orbits of the following form:*

(1) *connected components of hyperbolic cylinders* (6.22) *for $C \neq 0$;*

(2) *half-planes — connected components of the set $(\Pi_+ \cup \Pi_-) \setminus \mathbb{R}e_2$.*

So we considered all possible dispositions of the vector $l \in \mathbb{R}^3 \setminus \{0\}$, and in all cases described orbits of the 3-dimensional system (6.19). Now we study orbits of the full 6-dimensional system (6.20).

6.4.2 Orbits of the 6-dimensional system

The vector fields in the right-hand side of the 6-dimensional system (6.20) are

$$f(Q,\mu) = \begin{pmatrix} Q\widehat{\beta\mu} \\ \mu \times \beta\mu \end{pmatrix}, \qquad g(Q,\mu) = \begin{pmatrix} 0 \\ l \end{pmatrix}, \qquad (Q,\mu) \in \mathrm{SO}(3) \times \mathbb{R}^3.$$

Notice the commutation rule for vector fields of the form that appear in our problem:

$$f_i(Q,\mu) = \begin{pmatrix} Q\widehat{\beta w_i(\mu)} \\ v_i(\mu) \end{pmatrix} \in \mathrm{Vec}(\mathrm{SO}(3) \times \mathbb{R}^3),$$

$$[f_1,f_2](Q,\mu) = \begin{pmatrix} Q[\widehat{\beta w_1}, \widehat{\beta w_2}]_{\mathrm{so}(3)} + Q\widehat{\beta\left(\dfrac{\partial\, w_2}{\partial\, \mu}v_1 - \dfrac{\partial\, w_1}{\partial\, \mu}v_2\right)} \\ \\ \dfrac{\partial\, v_2}{\partial\, \mu}v_1 - \dfrac{\partial\, v_1}{\partial\, \mu}v_2 \end{pmatrix}.$$

We compute first the same Lie brackets as in the 3-dimensional case:

$$[g,f] = \begin{pmatrix} Q\widehat{\beta l} \\ l \times \beta\mu + \mu \times \beta l \end{pmatrix},$$

$$\frac{1}{2}[g,[g,f]] = \begin{pmatrix} 0 \\ l \times \beta l \end{pmatrix},$$

$$\frac{1}{2}[[g,[g,f]],[g,f]] = \begin{pmatrix} 0 \\ l \times \beta(l \times \beta l) + (l \times \beta l) \times \beta l \end{pmatrix}.$$

Further, for any vector field $X \in \mathrm{Vec}(\mathrm{SO}(3) \times \mathbb{R}^3)$ of the form

$$X = \begin{pmatrix} 0 \\ x \end{pmatrix}, \qquad x \text{ — a constant vector field on } \mathbb{R}^3, \tag{6.23}$$

we have

$$[X,f] = \begin{pmatrix} Q\widehat{\beta x} \\ * \end{pmatrix}. \tag{6.24}$$

To study the orbit of the 6-dimensional system (6.20) through a point $(Q,\mu) \in \mathrm{SO}(3) \times \mathbb{R}^3$, we follow the different cases for the 3-dimensional system (6.19) in Subsec. 6.4.1.

Case 1. $l \notin \Pi_{\pm}$, $l \notin \mathbb{R}e_i$. We can choose 3 linearly independent vector fields in $\mathrm{Lie}(f,g)$ of the form (6.23):

$$X_1 = g, \qquad X_2 = \frac{1}{2}[g,[g,f]], \qquad X_3 = \frac{1}{2}[[g,[g,f],[g,f]].$$

By the commutation rule (6.24), we have 6 linearly independent vectors in $\mathrm{Lie}_{(Q,\mu)}(f,g)$:

$$X_1 \wedge X_2 \wedge X_3 \wedge [X_1,f] \wedge [X_2,f] \wedge [X_3,f] \neq 0.$$

Thus the orbit through (Q, μ) is 6-dimensional.

Case 2. $l \in \Pi_{\pm} \setminus \mathbb{R}e_2$.

Case 2.1. $\mu \notin \Pi_{\pm}$. First of all, $\mathrm{Lie}(f, g)$ contains 2 linearly independent vector fields of the form (6.23):

$$X_1 = g, \qquad X_2 = \frac{1}{2}[g, [g, f]].$$

Since the trajectory of the free body in $\mathbb{R}^3$ through μ is not flat, we can assume that the vector $v = \mu \times \beta\mu$ is linearly independent of l and $l \times \beta l$. Now our aim is to show that $\mathrm{Lie}(f, g)$ contains 2 vector fields of the form

$$Y_1 = \begin{pmatrix} QM_1 \\ v_1 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} QM_2 \\ v_2 \end{pmatrix}, \qquad M_1 \wedge M_2 \neq 0, \tag{6.25}$$

where the vector fields v_1 and v_2 vanish at the point μ. If this is the case, then $\mathrm{Lie}_{(Q,\mu)}(f, g)$ contains 6 linearly independent vectors:

$$\begin{aligned}
&X_1(Q, \mu), \quad X_2(Q, \mu), \quad f(Q, \mu), \\
&Y_1(Q, \mu) = \begin{pmatrix} QM_1 \\ 0 \end{pmatrix}, \quad Y_2(Q, \mu) = \begin{pmatrix} QM_2 \\ 0 \end{pmatrix}, \\
&[Y_1, Y_2](Q, \mu) = \begin{pmatrix} Q[M_1, M_2] \\ 0 \end{pmatrix},
\end{aligned}$$

and the orbit through the point (Q, μ) is 6-dimensional.

Now we construct 2 vector fields of the form (6.25) in $\mathrm{Lie}(f, g)$. Taking appropriate linear combinations with the fields X_1, X_2, we project the second component of the fields $[g, f]$ and $\frac{1}{2}[f, [g, [g, f]]$ to the line $\mathbb{R}v$, thus we obtain the vector fields

$$\begin{pmatrix} Q\widehat{\beta l} \\ k_1 v \end{pmatrix}, \quad \begin{pmatrix} Q\widehat{\beta(l \times \beta l)} \\ k_2 v \end{pmatrix} \in \mathrm{Lie}(f, g). \tag{6.26}$$

If both k_1 and k_2 vanish at μ, these vector fields can be taken as Y_1, Y_2 in (6.25). And if k_1 or k_2 does not vanish at μ, we construct such vector fields Y_1, Y_2 taking appropriate linear combinations of fields (6.26) and f with the fields g, $[g, [g, f]]$.

So in Case 2.1 the orbit is 6-dimensional.

Case 2.2. $\mu \in \Pi_{\pm}$. There are 5 linearly independent vectors in $\mathrm{Lie}_{(Q,\mu)}(f, g)$:

$$X_1 = g, \quad X_2 = \frac{1}{2}[g, [g, f]], \quad [X_1, f], \quad [X_2, f], \quad [[X_1, f], [X_2, f]].$$

Since the orbit in $\mathbb{R}^3$ is 2-dimensional, the orbit in $\mathrm{SO}(3) \times \mathbb{R}^3$ is 5-dimensional.

Case 3. $l \in \mathbb{R}e_1 \setminus \{0\}$.

Case 3.1. $\mu \notin \mathbb{R}e_1$. The argument is similar to that of Case 2.1. We can assume that the vectors l and $v = \mu \times \beta\mu$ are linearly independent. The orbit in $\mathbb{R}^3$ is 2-dimensional and the vectors l, v span the tangent space to this orbit, thus we can find vector fields in $\operatorname{Lie}(f, g)$ of the form:

$$Y_1 = [g, f] - C_1 g - C_2 f = \begin{pmatrix} Q\widehat{\beta l} + C_3 Q\widehat{\beta\mu} \\ 0 \end{pmatrix},$$

$$Y_2 = [Y_1, f] = \begin{pmatrix} Q[\widehat{\beta l}, \widehat{\beta\mu}] + C_4 Q\widehat{\beta\mu} \\ 0 \end{pmatrix}$$

for some real functions C_i, $i = 1, \dots, 4$. Then we have 5 linearly independent vectors in $\operatorname{Lie}_{(Q,\mu)}(f, g)$:

$$g, \quad f, \quad Y_1, \quad Y_2, \quad [Y_1, Y_2].$$

So the orbit of the 6-dimensional system (6.20) is 5-dimensional (it cannot have dimension 6 since the 3-dimensional system (6.19) has a 2-dimensional orbit).

Case 3.2. $\mu \in \mathbb{R}e_1$. The vectors

$$f(Q, \mu) = \begin{pmatrix} Q\widehat{\beta\mu} \\ 0 \end{pmatrix}, \qquad [g, f](Q, \mu) = \begin{pmatrix} Q\widehat{\beta l} \\ 0 \end{pmatrix},$$

are linearly dependent, thus $\dim \operatorname{Lie}_{(Q,\mu)}(f, g) = \dim \operatorname{span}(f, g)|_{(Q,\mu)} = 2$. So the orbit is 2-dimensional.

The cases $l \in \mathbb{R}e_i \setminus \{0\}$, $i = 1, 2$, are similar to Case 3.

We completed the study of orbits of the controlled rigid body (6.20) and now summarize it.

Proposition 6.6. *Let (Q, μ) be a point in $\mathrm{SO}(3) \times \mathbb{R}^3$. If the orbit $\mathcal{O}$ of the 3-dimensional system (6.19) through the point μ is 3- or 2-dimensional, then the orbit of the 6-dimensional system (6.20) through the point (Q, μ) is $\mathrm{SO}(3) \times \mathcal{O}$, i.e., respectively 6- or 5-dimensional. If $\dim \mathcal{O} = 1$, then the 6-dimensional system has a 2-dimensional orbit.*

We will describe attainable sets of this system in Section 8.4 after acquiring some general facts on attainable sets.

Chapter 7

Control of configurations

In this chapter we consider systems which can be controlled by the change of their configuration, i.e., of relative position of parts of the systems. A falling cat exhibits a well-known example of such a control. If a cat is left free over ground (e.g. if it falls from a tree or is thrown down by a child), then the cat starts to rotate its tail and bend its body, and finally falls to the ground exactly on its paws, regardless of its initial orientation over the ground. Such a behavior cannot be demonstrated by a mechanical system less skillful in turning and bending its parts (e.g. a dog or just a rigid body), so the crucial point in the falling cat phenomenon seems to be control by the change of configuration. We present a simple model of systems controlled in such a way, and study orbits in several simplest examples.

7.1 Model

A system of mass points, i.e., a mass distribution in $\mathbb{R}^n$, is described by a nonnegative measure μ in $\mathbb{R}^n$. We restrict ourselves by measures with compact support. For example, a system of points $x_1, \dots, x_k \in \mathbb{R}^n$ with masses $\mu_1, \dots, \mu_k > 0$ is modeled by the atomic measure $\mu = \sum_{i=1}^k \mu_i \delta_{x_i}$, where δ_{x_i} is the Dirac function concentrated at x_i. One can consider points x_i free or restricted by constraints in $\mathbb{R}^n$. More generally, mass can be distributed along segments or surfaces of various dimensions. So the state space M of a system to be considered is a reasonable class of measures in $\mathbb{R}^n$.

A controller is supposed to sit in the construction and change its configuration. The system is conservative, i.e., impulse and angular momentum are conserved.

Mathematically, conservation laws of a system come from Nöther theorem due to symmetries of the system. Kinetic energy of our system is

$$L = \frac{1}{2} \int |\dot{x}|^2 \, d\mu, \tag{7.1}$$

in particular, for an atomic measure $\mu = \sum_{i=1}^k \mu_i \delta_{x_i}$,

$$L = \frac{1}{2} \sum_{i=1}^k \mu_i |\dot{x}_i|^2.$$

By Nöther theorem (see e.g. [4]), if the flow of a vector field $V \in \operatorname{Vec} \mathbb{R}^n$ preserves a Lagrangian L, then the system has an integral of the form

$$\frac{\partial L}{\partial \dot{x}} V(x) = \text{const} .$$

In our case, Lagrangian (7.1) is invariant w.r.t. isometries of the Euclidean space, i.e., translations and rotations in $\mathbb{R}^n$.

Translations in $\mathbb{R}^n$ are generated by constant vector fields:

$$V(x) = a \in \mathbb{R}^n,$$

and our system is subject to the conservation laws

$$\int \langle \dot{x}, a \rangle \, d\mu = \text{const} \qquad \forall \; a \in \mathbb{R}^n,$$

that is,

$$\int \dot{x} \, d\mu = \text{const},$$

i.e., the center of mass of the system moves with a constant velocity (the total impulse is preserved). We choose the inertial frame of reference in which the center of mass is fixed:

$$\int \dot{x} \, d\mu = 0.$$

For an atomic measure $\mu = \sum_{i=1}^k \mu_i \delta_{x_i}$, this equality takes the form

$$\sum_{i=1}^k \mu_i x_i = \text{const},$$

which is reduced by a change of coordinates in $\mathbb{R}^n$ to

$$\sum_{i=1}^k \mu_i x_i = 0.$$

Now we pass to rotations in $\mathbb{R}^n$. Let a vector field

$$V(x) = Ax, \qquad x \in \mathbb{R}^n,$$

preserve the Euclidean structure in $\mathbb{R}^n$, i.e., its flow

$$e^{tV(x)} = e^{tA} x$$

preserve the scalar product:

$$\langle e^{tA}x, e^{tA}y\rangle = \langle x, y\rangle, \qquad x,\ y \in \mathbb{R}^n.$$

Differentiation of this equality at $t = 0$ yields

$$\langle Ax, y\rangle + \langle x, Ay\rangle = 0, \qquad x,\ y \in \mathbb{R}^n,$$

i.e., the matrix A is skew-symmetric:

$$A^* = -A.$$

Conversely, if the previous equality holds, then

$$\left(e^{tA}\right)^* = e^{tA^*} = e^{-tA} = \left(e^{tA}\right)^{-1},$$

i.e., the matrix e^{tA} is orthogonal. We proved that the flow e^{tA} preserves the Euclidean structure in $\mathbb{R}^n$ if and only if $A^* = -A$. Similarly to the 3-dimensional case considered in Sec. 6.1, the group of orientation-preserving linear orthogonal transformations of the Euclidean space $\mathbb{R}^n$ is denoted by $\mathrm{SO}(n)$, and the corresponding Lie algebra of skew-symmetric transformations in $\mathbb{R}^n$ is denoted by $\mathrm{so}(n)$. In these notations,

$$e^{tA} \in \mathrm{SO}(n) \Leftrightarrow\ A \in \mathrm{so}(n).$$

Return to derivation of conservation laws for our system of mass points. The Lagrangian $L = \frac{1}{2}\int |\dot{x}|^2\, d\mu$ is invariant w.r.t. rotations in $\mathbb{R}^n$, so Nöther theorem gives integrals of the form

$$\frac{\partial L}{\partial \dot{x}} V(x) = \int \langle \dot{x}, Ax\rangle\, d\mu = \text{const}, \qquad A \in \mathrm{so}(n).$$

For an atomic measure $\mu = \sum_{i=1}^k \mu_i \delta_{x_i}$, we obtain

$$\sum_{i=1}^k \mu_i \langle \dot{x}_i, Ax_i\rangle = \text{const}, \qquad A \in \mathrm{so}(n), \tag{7.2}$$

and we restrict ourselves by the simplest case where the constant in the right-hand side is just zero.

Summing up, we have the following conservation laws for a system of points $x_1, \dots, x_k \in \mathbb{R}^n$ with masses $\mu_1, \dots, \mu_k$:

$$\sum_{i=1}^k \mu_i x_i = 0, \tag{7.3}$$

$$\sum_{i=1}^k \mu_i \langle \dot{x}_i, Ax_i\rangle = 0, \qquad A \in \mathrm{so}(n). \tag{7.4}$$

The state space is a subset

$$M \subset \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k},$$

and admissible paths are piecewise smooth curves in M that satisfy constraints (7.3), (7.4). The first equality (7.3) determines a submanifold in M; in fact, this equality can obviously be resolved w.r.t. any variable x_i, and one can get rid of this constraint by decreasing dimension of M. The second equality (7.4) is a linear constraint for velocities $\dot{x}_i$, it determines a distribution on M. So the admissibility conditions (7.3), (7.4) define a linear in control, thus symmetric, control system on M. Notice that a more general condition (7.2) determines an "affine distribution", and control system (7.3), (7.2) is affine in controls, thus, in general, not symmetric.

We consider only the symmetric case (7.3), (7.4). Then orbits coincide with attainable sets. We compute orbits in the following simple situations:

(1) Two free points: $k = 2$,

(2) Three free points: $k = 3$,

(3) A broken line with 3 links in $\mathbb{R}^2$.

7.2 Two free points

We have $k = 2$, and the first admissibility condition (7.3) reads

$$\mu_1 x_1 + \mu_2 x_2 = 0, \qquad x_1, \ x_2 \in \mathbb{R}^n.$$

We eliminate the second point:

$$x_1 = x, \quad x_2 = -\frac{\mu_1}{\mu_2} x_1,$$

and exclude collisions of the points:

$$x \neq 0.$$

So the state space of the system is

$$M = \mathbb{R}^n \setminus \{0\}.$$

The second admissibility condition (7.4)

$$\mu_1 \langle \dot{x}_1, A x_1 \rangle + \mu_2 \langle \dot{x}_2, A x_2 \rangle = 0, \qquad A \in \mathrm{so}(n),$$

is rewritten as

$$\left(\mu_1 + \mu_1^2/\mu_2\right) \langle \dot{x}, A x \rangle = 0, \qquad A \in \mathrm{so}(n),$$

i.e.,

$$\langle \dot{x}, A x \rangle = 0, \qquad A \in \mathrm{so}(n). \tag{7.5}$$

This equation can easily be analyzed via the following proposition.

Exercise 7.1. If $A \in \mathrm{so}(n)$, then

$$\langle Ax, x \rangle = 0, \qquad x \in \mathbb{R}^n.$$

Moreover, for any vector $x \in \mathbb{R}^n \setminus \{0\}$, the space

$$\{Ax \mid A \in \mathrm{so}(n)\}$$

coincides with the whole orthogonal complement

$$x^\perp = \{y \in \mathbb{R}^n \mid \langle y, x \rangle = 0\}.$$

So restriction (7.5) means that

$$\dot{x} \wedge x = 0,$$

i.e., velocity of an admissible curve is proportional to the state vector. The distribution determined by this condition is one-dimensional, thus integrable. So admissible curves have the form

$$x(t) = \alpha(t)x(0), \qquad \alpha(t) > 0.$$

The orbit and admissible set through any point $x \in \mathbb{R}^n \setminus \{0\}$ is

$$\mathcal{O}_x = \mathbb{R}_+ x = \{\alpha x \mid \alpha > 0\}.$$

The points x_1, x_2 can move only along a fixed line in $\mathbb{R}^n$, and orientation of the system cannot be changed. In order to have a more sophisticated behavior, one should consider more complex systems.

7.3 Three free points

Now $k = 3$, and we eliminate the third point via the first admissibility condition (7.3):

$$x = \mu_1 x_1, \qquad y = \mu_2 x_2,$$
$$x_3 = -\frac{1}{\mu_3}(x + y).$$

In order to exclude the singular configurations where the points x_1, x_2, x_3 are collinear, we assume that the vectors x, y are linearly independent. So the state space is

$$M = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid x \wedge y \neq 0\}.$$

Introduce the notation

$$\rho_i = \frac{1}{\mu_i}, \qquad i = 1, 2, 3.$$

Then the second admissibility condition (7.4) takes the form:

$$\langle \dot{x}, A((\rho_1 + \rho_3)x + \rho_3 y) \rangle + \langle \dot{y}, A((\rho_2 + \rho_3)y + \rho_3 x) \rangle = 0, \qquad A \in \mathrm{so}(n).$$

It turns out then that admissible velocities $\dot{x}$, $\dot{y}$ should belong to the plane $\mathrm{span}(x, y)$. This follows by contradiction from the following proposition.

Lemma 7.1. *Let vectors $v, w, \xi, \eta \in \mathbb{R}^n$ satisfy the conditions*

$$v \wedge w \neq 0, \qquad \operatorname{span}(v, w, \xi, \eta) \neq \operatorname{span}(v, w).$$

Then there exists $A \in \mathrm{so}(n)$ such that

$$\langle Av, \xi \rangle + \langle Aw, \eta \rangle \neq 0.$$

Proof. First of all, we may assume that

$$\langle v, w \rangle = 0. \tag{7.6}$$

Indeed, choose a vector $\widehat{w} \in \operatorname{span}(v, w)$ such that $\langle v, \widehat{w} \rangle = 0$. Then $w = \widehat{w} + \alpha v$ and

$$\langle Av, \xi \rangle + \langle Aw, \eta \rangle = \langle Av, \xi + \alpha\eta \rangle + \langle A\widehat{w}, \eta \rangle,$$

thus we can replace w by $\widehat{w}$.

Second, we can renormalize vectors v, w and assume that

$$|v| = |w| = 1. \tag{7.7}$$

Now let $\xi \notin \operatorname{span}(v, w)$, we can assume this since the hypotheses of the lemma are symmetric w.r.t. ξ, η. Then

$$\xi = \alpha v + \beta w + l$$

for some vector

$$l \perp \operatorname{span}(v, w).$$

Choose an operator $A \in \mathrm{so}(n)$ such that

$$\begin{aligned} &Aw = 0, \\ &A : \operatorname{span}(v, l) \to \operatorname{span}(v, l) \text{ is invertible.} \end{aligned}$$

Then

$$\langle Av, \xi \rangle + \langle Aw, \eta \rangle = \langle Av, l \rangle \neq 0,$$

i.e., the operator A is the required one. $\square$

This lemma means that for any pair of initial points $(x, y) \in M$, all admissible curves x_t and y_t are contained in the plane $\operatorname{span}(x, y) \subset \mathbb{R}^n$. So we can reduce our system to such a plane and thus assume that $x, y \in \mathbb{R}^2$.

Thus we obtain the following system:

$$\begin{aligned} &\langle \dot{x}, A((\rho_1 + \rho_3)x + \rho_3 y) \rangle + \langle \dot{y}, A((\rho_2 + \rho_3)y + \rho_3 x) \rangle = 0, \quad A \in \mathrm{so}(2), \\ &(x, y) \in M = \{(v, w) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid v \wedge w \neq 0\}. \end{aligned} \tag{7.8}$$

Consequently,

$$A = \mathrm{const} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

i.e., equality (7.8) defines one linear equation on velocities, thus a rank 3 distribution on a 4-dimensional manifold M. Using Exercise 7.1, it is easy to see that this distribution is spanned by the following 3 linear vector fields:

$$\begin{aligned}
V_1 &= ((\rho_1+\rho_3)x+\rho_3 y)\frac{\partial}{\partial x} = \begin{pmatrix} (\rho_1+\rho_3)x+\rho_3 y \\ 0 \end{pmatrix} = B_1\begin{pmatrix} x \\ y \end{pmatrix}, \\
V_2 &= ((\rho_2+\rho_3)y+\rho_3 x)\frac{\partial}{\partial y} = \begin{pmatrix} 0 \\ \rho_3 x+(\rho_2+\rho_3)y \end{pmatrix} = B_2\begin{pmatrix} x \\ y \end{pmatrix}, \\
V_3 &= x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y} = \begin{pmatrix} x \\ y \end{pmatrix} = \mathrm{Id}\begin{pmatrix} x \\ y \end{pmatrix},
\end{aligned}$$

where

$$B_1 = \begin{pmatrix} \rho_1+\rho_3 & \rho_3 \\ 0 & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & 0 \\ \rho_3 & \rho_2+\rho_3 \end{pmatrix}, \quad \mathrm{Id} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In order to simplify notations, we write here 4-dimensional vectors as 2-dimensional columns: e.g.,

$$V_1 = \begin{pmatrix} (\rho_1+\rho_3)x+\rho_3 y \\ 0 \end{pmatrix} = \begin{pmatrix} (\rho_1+\rho_3)x_1+\rho_3 y_1 \\ (\rho_1+\rho_3)x_2+\rho_3 y_2 \\ 0 \\ 0 \end{pmatrix},$$

where

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

The rank 3 distribution in question can have only orbits of dimensions 3 or 4. In order to find out, which of these possibilities are realized, compute the Lie bracket:

$$\begin{aligned}
[V_1, V_2] &= [B_1, B_2]\begin{pmatrix} x \\ y \end{pmatrix}, \\
[B_1, B_2] &= \rho_3\begin{pmatrix} \rho_3 & \rho_2+\rho_3 \\ -(\rho_1+\rho_3) & -\rho_3 \end{pmatrix}.
\end{aligned}$$

It is easy to check that

$$V_1 \wedge V_2 \wedge V_3 \wedge [V_1, V_2] \neq 0 \;\Leftrightarrow\; B_1 \wedge B_2 \wedge \mathrm{Id} \wedge [B_1, B_2] \neq 0.$$

We write 2×2 matrices as vectors in the standard basis of $\mathrm{gl}(2)$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

then

$$\begin{aligned}
\det(\mathrm{Id}, B_1, B_2, [B_1, B_2]) &= \begin{vmatrix} 1 & \rho_1+\rho_3 & 0 & \rho_3 \\ 0 & \rho_3 & 0 & \rho_2+\rho_3 \\ 0 & 0 & \rho_3 & -(\rho_1+\rho_3) \\ 1 & 0 & \rho_2+\rho_3 & -\rho_3 \end{vmatrix} \\
&= 2\rho_3(\rho_1\rho_2+\rho_1\rho_3+\rho_2\rho_3) > 0.
\end{aligned}$$

Consequently, the fields $V_1, V_2, V_3, [V_1, V_2]$ are linearly independent everywhere on M, i.e., the control system has only 4-dimensional orbits. So the orbits coincide with connected components of the state space. The manifold M is decomposed into 2 connected components corresponding to positive or negative orientation of the frame (x, y):

$$M = M_+ \cup M_-,$$
$$M_\pm = \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid \det(x, y) \gtrless 0\}.$$

So the system on M has 2 orbits, thus 2 attainable sets: M_+ and M_-. Given any pair of linearly independent vectors $(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2$, we can reach any other nonsingular configuration $(\tilde{x}, \tilde{y}) \in \mathbb{R}^2 \times \mathbb{R}^2$ with $\tilde{x}, \tilde{y} \in \operatorname{span}(x, y)$ and the frame $(\tilde{x}, \tilde{y})$ oriented in the same way as (x, y).

Returning to the initial problem for 3 points $x_1, x_2, x_3 \in \mathbb{R}^n$: the 2-dimensional linear plane of the triangle (x_1, x_2, x_3) should be preserved, as well as orientation and center of mass of the triangle. Except this, the triangle (x_1, x_2, x_3) can be rotated, deformed or dilated as we wish.

Configurations of 3 points that define distinct 2-dimensional planes (or define distinct orientations in the same 2-dimensional plane) are not mutually reachable: attainable sets from these configurations do not intersect one with another. Although, if two configurations define 2-dimensional planes having a common line, then intersection of *closures* of attainable sets from these configurations is nonempty: it consists of collinear triples lying in the common line. Theoretically, one can imagine a motion that steers one configuration into another: first the 3 points are made collinear in the initial 2-dimensional plane, and then this collinear configuration is steered to the final one in the terminal 2-dimensional plane.

7.4 Broken line

Consider a system of 4 mass points placed at vertices of a broken line of 3 segments in a 2-dimensional plane. We study the most symmetric case, where all masses are equal to 1 and lengths of all segments are also equal to 1.

The holonomic constraints for the points

$$x_0,\ x_1,\ x_2,\ x_3 \in \mathbb{R}^2 = \mathbb{C}$$

have the form

$$\sum_{j=0}^{3} x_j = 0, \qquad |x_j - x_{j-1}| = 1, \qquad j = 1, 2, 3. \tag{7.9}$$

Thus

$$x_j - x_{j-1} = e^{i\theta_j}, \qquad \theta_j \in S^1, \quad j = 1, 2, 3.$$

Position of the system is determined by the 3-tuple of angles $(\theta_1, \theta_2, \theta_3)$, so the state space is the 3-dimensional torus:

$$M = S^1 \times S^1 \times S^1 = T^3 = \{(\theta_1, \theta_2, \theta_3) \mid \theta_j \in S^1,\ j = 1, 2, 3\}.$$

The nonholonomic constraints on velocities reduce to the equality

$$\sum_{j=0}^{3} \langle i x_j, \dot{x}_j \rangle = 0.$$

In order to express this equality in terms of the coordinates θ_j, denote first

$$y_j = x_j - x_{j-1}, \qquad j = 1, 2, 3.$$

Taking into account the condition $\sum_{j=0}^{3} x_j = 0$, we obtain:

$$\begin{aligned} x_0 &= -\frac{3y_1}{4} - \frac{y_2}{2} - \frac{y_3}{4}, \\ x_1 &= \frac{y_1}{4} - \frac{y_2}{2} - \frac{y_3}{4}, \\ x_2 &= \frac{y_1}{4} + \frac{y_2}{2} - \frac{y_3}{4}, \\ x_3 &= \frac{y_1}{4} + \frac{y_2}{2} + \frac{3y_3}{4}. \end{aligned}$$

Now compute the differential form:

$$\begin{aligned} \omega = \sum_{j=0}^{3} \langle i x_j, d x_j \rangle = {} & \langle i \left((3/4) y_1 + (1/2) y_2 + (1/4) y_3\right), d y_1 \rangle \\ & + \langle i \left((1/2) y_1 + y_2 + (1/2) y_3\right), d y_2 \rangle \\ & + \langle i \left((1/4) y_1 + (1/2) y_2 + (3/4) y_3\right), d y_3 \rangle . \end{aligned}$$

Since $\langle i y_j, d y_k \rangle = \langle e^{i\theta_j}, e^{i\theta_k} d\theta_k \rangle = \cos(\theta_j - \theta_k) d\theta_k$, we have

$$\begin{aligned} \omega = {} & \left((3/4) + (1/2)\cos(\theta_2 - \theta_1) + (1/4)\cos(\theta_3 - \theta_1)\right) d\theta_1 \\ & + \left((1/2)\cos(\theta_1 - \theta_2) + 1 + (1/2)\cos(\theta_3 - \theta_2)\right) d\theta_2 \\ & + \left((1/4)\cos(\theta_1 - \theta_3) + (1/2)\cos(\theta_2 - \theta_3) + 3/4\right) d\theta_3 . \end{aligned}$$

Consequently, the system under consideration is the rank 2 distribution $\Delta = \operatorname{Ker} \omega$ on the 3-dimensional manifold $M = T^3$. The orbits can be 2- or 3-dimensional. To distinguish these cases, we can proceed as before: find a vector field basis and compute Lie brackets. But now we study integrability of Δ in a dual way, via techniques of differential forms.

Assume that the distribution Δ has a 2-dimensional integral manifold $N \subset M$. Then

$$\omega|_N = 0,$$

consequently,

$$0 = d\left(\omega|_N\right) = (d\omega)|_N ,$$

thus

$$0 = d\omega_q|_{\Delta_q} = d\omega_q|_{\operatorname{Ker} \omega_q}, \qquad q \in N.$$

In terms of exterior product of differential forms,

$$(\omega \wedge d\omega)_q = 0, \qquad q \in N.$$

We compute the differential and exterior product:

$$d\omega = \sin(\theta_2 - \theta_1) d\theta_1 \wedge d\theta_2 + \sin(\theta_3 - \theta_2) d\theta_2 \wedge d\theta_3 + \frac{1}{2} \sin(\theta_3 - \theta_1) d\theta_1 \wedge \theta_3,$$
$$\omega \wedge d\omega = \frac{1}{2} (\sin(\theta_2 - \theta_1) + \sin(\theta_3 - \theta_2)) d\theta_1 \wedge d\theta_2 \wedge d\theta_3.$$

Thus $\omega \wedge d\omega = 0$ if and only if

$$\sin(\theta_2 - \theta_1) + \sin(\theta_3 - \theta_2) = 0,$$

i.e.,

$$\theta_3 = \theta_1 \tag{7.10}$$

or

$$(\theta_1 - \theta_2) + (\theta_3 - \theta_2) = \pi. \tag{7.11}$$

Configurations (7.10) and (7.11) are hard to control: if neither of these equalities is satisfied, then $\omega \wedge d\omega \neq 0$, i.e., the system has 3-dimensional orbits through such points. If we choose basis vector fields X_1, X_2 of the distribution Δ, then already the first bracket $[X_1, X_2]$ is linearly independent of X_1, X_2 at points where both equalities (7.10), (7.11) are violated.

Now it remains to study integrability of Δ at points of surfaces (7.10), (7.11). Here $[X_1, X_2](q) \in \Delta_q$, but we may obtain nonintegrability of Δ via brackets of higher order.

Consider first the two-dimensional surface

$$P = \{\theta_3 = \theta_1\}.$$

If the orbit through a point $q \in P$ is two-dimensional, then the distribution Δ should be tangent to P in the neighborhood of q. But it is easy to see that Δ is everywhere transversal to P: e.g.,

$$T_q P \ni \left. \frac{\partial}{\partial \theta_2} \right|_q \notin \Delta_q, \qquad q \in P.$$

So the system has 3-dimensional orbits through any point of P.

In the same way one can see that the orbits through points of the second surface (7.11) are 3-dimensional as well.

The state space M is connected, thus there is the only orbit (and attainable set) — the whole manifold M. The system is completely controllable.

Chapter 8

Attainable sets

Let M be a smooth manifold and $\mathcal{F} \subset \operatorname{Vec} M$ a *bracket-generating* family of vector fields on M:

$$\operatorname{Lie}_q \mathcal{F} = T_q M \qquad \forall\, q \in M. \tag{8.1}$$

Such a system is also called *full-rank*.

If a family $\mathcal{F} \subset \operatorname{Vec} M$ is not bracket-generating, and M and $\mathcal{F}$ are real analytic, we can pass from $\mathcal{F}$ to a bracket-generating family $\mathcal{F}|_{\mathcal{O}}$, where $\mathcal{O}$ is an orbit of $\mathcal{F}$ (see the analytic version of the Orbit Theorem, Corollary 5.3). Thus in the analytic case requirement (8.1) is not restrictive in essence.

8.1 Attainable sets of full-rank systems

The following proposition describes important properties of attainable sets $\mathcal{A}_{q_0}$, $q_0 \in M$, for arbitrary nonnegative time.

Theorem 8.1 (Krener). *Let $\mathcal{F} \subset \operatorname{Vec} M$ be a bracket-generating system. Then $\mathcal{A}_{q_0} \subset \overline{\operatorname{int} \mathcal{A}_{q_0}}$ for any $q_0 \in M$.*

Remark. In particular, attainable sets for arbitrary time have nonempty interior:

$$\operatorname{int} \mathcal{A}_{q_0} \neq \emptyset.$$

Attainable sets may be:

- open sets,
- manifolds with smooth boundary,
- manifolds with boundary having singularities (corner or cuspidal points).

One can easily construct control systems (e.g. in the plane) that realize these possibilities.

On the other hand, Krener's theorem prohibits an attainable set $\mathcal{A}_{q_0}$ of a bracket-generating family to be:

- a lower-dimensional subset of M,
- a set where boundary points are isolated from interior points.

Now we prove Krener's theorem.

Proof. Fix an arbitrary point $q_0 \in M$ and take a point $q' \in \mathcal{A}_{q_0}$. We show that

$$q' \in \overline{\operatorname{int} \mathcal{A}_{q_0}}. \tag{8.2}$$

(1) There exists a vector field $f_1 \in \mathcal{F}$ such that $f_1(q') \neq 0$, otherwise $\operatorname{Lie}_{q'}(\mathcal{F}) = 0$ and $\dim M = 0$. The curve

$$s_1 \mapsto q' \circ e^{s_1 f_1}, \qquad s_1 \in (0, \varepsilon) \tag{8.3}$$

is a 1-dimensional submanifold of M for small enough $\varepsilon > 0$.

If $\dim M = 1$, then $q' \circ e^{s_1 f_1} \in \operatorname{int} \mathcal{A}_{q_0}$ for sufficiently small $s_1 > 0$, and inclusion (8.2) follows.

(2) Assume that $\dim M > 1$. Then arbitrarily close to q' we can find a point q_1 on curve (8.3) and a field $f_2 \in \mathcal{F}$ such that $f_2(q_1)$ is not tangent to manifold (8.3):

$$\begin{aligned} &q_1 = q' \circ e^{t_1^1 f_1}, \qquad t_1^1 \text{ sufficiently small}, \\ &(q_1 \circ f_1) \wedge (q_1 \circ f_2) \neq 0, \end{aligned}$$

otherwise $\dim \operatorname{Lie}_q \mathcal{F} = 1$ for q on curve (8.3) with small s_1. Then the mapping

$$(s_1, s_2) \mapsto q_1 \circ e^{s_1 f_1} \circ e^{s_2 f_2}, \qquad s_1 > 0, \ s_2 > 0, \tag{8.4}$$

is an immersion in a small neighborhood $O(0) \subset \mathbb{R}^2_{s_1, s_2}$, thus its image is a 2-dimensional submanifold of M.

If $\dim M = 2$, inclusion (8.2) is proved.

(3) Assume that $\dim M > 2$. We can find a vector $f_3(q)$, $f_3 \in \mathcal{F}$, not tangent to surface (8.4) sufficiently close to q': there exist $t_2^1, t_2^2 > 0$ and $f_3 \in \mathcal{F}$ such that the vector field f_3 is not tangent to surface (8.4) at a point $q_2 = q' \circ e^{t_2^1 f_1} \circ e^{t_2^2 f_2}$. Otherwise the family $\mathcal{F}$ is not bracket-generating.

The mapping

$$(s_1, s_2, s_3) \mapsto q_2 \circ e^{s_1 f_1} \circ e^{s_2 f_2} \circ e^{s_3 f_3}, \qquad s_i > 0,$$

is an immersion in a small neighborhood $O(0) \subset \mathbb{R}^3_{s_1, s_2, s_3}$, thus its image is a smooth 3-dimensional submanifold of M.

If $\dim M = 3$, inclusion (8.2) follows. Otherwise we continue this procedure.

(4) For $\dim M = n$, inductively, we find a point

$$(t_{n-1}^1, t_{n-1}^2, \dots, t_{n-1}^{n-1}) \in \mathbb{R}^{n-1}, \qquad t_{n-1}^i > 0$$

and fields $f_1, \dots, f_n \in \mathcal{F}$ such that such that the mapping

$$(s_1, \dots, s_n) \mapsto q_{n-1} \circ e^{s_1 f_1} \circ \dots \circ e^{s_n f_n}, \qquad s_i > 0,$$
$$q_{n-1} = q' \circ e^{t^1_{n-1} f_1} \circ e^{t^2_{n-1} f_2} \circ \dots \circ e^{t^{n-1}_{n-1} f_{n-1}},$$

is an immersion. The image of this immersion is an n-dimensional submanifold of M, thus an open set. This open set is contained in $\mathcal{A}_{q_0}$ and can be chosen as close to the point q' as we wish. Inclusion (8.2) is proved, and the theorem follows. □

We obtain the following proposition from Krener's theorem.

Corollary 8.1. *Let $\mathcal{F} \subset \operatorname{Vec} M$ be a bracket-generating system. If $\overline{\mathcal{A}_{q_0}(\mathcal{F})} = M$ for some $q_0 \in M$, then $\mathcal{A}_{q_0}(\mathcal{F}) = M$.*

Proof. Take an arbitrary point $q \in M$. We show that $q \in \mathcal{A}_{q_0}(\mathcal{F})$.

Consider the system

$$-\mathcal{F} = \{-V \mid V \in \mathcal{F}\} \subset \operatorname{Vec} M.$$

This system is bracket-generating, thus by Theorem 8.1

$$\mathcal{A}_q(-\mathcal{F}) \subset \overline{\operatorname{int} \mathcal{A}_q(-\mathcal{F})} \qquad \forall q \in M.$$

Take any point $\widehat{q} \in \operatorname{int} \mathcal{A}_q(-\mathcal{F})$ and a neighborhood of this point $O_{\widehat{q}} \subset \mathcal{A}_q(-\mathcal{F})$. Since $\mathcal{A}_{q_0}(\mathcal{F})$ is dense in M, then

$$\mathcal{A}_{q_0}(\mathcal{F}) \cap O_{\widehat{q}} \neq \emptyset.$$

That is why $\mathcal{A}_{q_0}(\mathcal{F}) \cap \mathcal{A}_q(-\mathcal{F}) \neq \emptyset$, i.e., there exists a point

$$q' \in \mathcal{A}_{q_0}(\mathcal{F}) \cap \mathcal{A}_q(-\mathcal{F})$$

Thus the point q' can be represented as follows:

$$q' = q_0 \circ e^{t_1 f_1} \circ \dots \circ e^{t_k f_k}, \qquad f_i \in \mathcal{F},\ t_i > 0,$$
$$q' = q \circ e^{-s_1 g_1} \circ \dots \circ e^{-s_l g_l}, \qquad g_i \in \mathcal{F},\ s_i > 0.$$

We multiply both decompositions from the right by $e^{s_l g_l} \circ \dots \circ e^{s_1 g_1}$ and obtain $q = q_0 \circ e^{t_1 f_1} \circ \dots \circ e^{t_k f_k} \circ e^{s_l g_l} \circ \dots \circ e^{s_1 g_1}$, i.e., $q \in \mathcal{A}_{q_0}(\mathcal{F})$. □

The sense of the previous proposition is that in the study of controllability, we can replace the attainable set of a bracket-generating system by its closure. In the following section we try to add new vector fields to a system so that the closure of its attainable set do not change.

8.2 Compatible systems and relaxations

Definition 8.1. A vector field $f \in \operatorname{Vec} M$ is called *compatible* with a system $\mathcal{F} \subset \operatorname{Vec} M$ if

$$\mathcal{A}_q(\mathcal{F} \cup f) \subset \overline{\mathcal{A}_q(\mathcal{F})} \qquad \forall q \in M.$$

Easy compatibility condition is given by the following statement.

Proposition 8.1. *Let $\mathcal{F} \subset \operatorname{Vec} M$. For any vector fields $f_1, f_2 \in \mathcal{F}$, and any functions $a_1, a_2 \in C^\infty(M)$, $a_1, a_2 \geq 0$, the vector field $a_1 f_1 + a_2 f_2$ is compatible with $\mathcal{F}$.*

In view of Corollary 5.2, the following proposition holds.

Corollary 8.2. *If $\mathcal{F} \subset \operatorname{Vec} M$ is a bracket-generating system such that the positive convex cone generated by $\mathcal{F}$*

$$\operatorname{cone}(\mathcal{F}) = \left\{ \sum_{i=1}^{k} a_i f_i \mid f_i \in \mathcal{F},\ a_i \in C^\infty(M),\ a_i \geq 0,\ k \in \mathbb{N} \right\} \subset \operatorname{Vec} M$$

is symmetric, then $\mathcal{F}$ is controllable.

Proposition 8.1 is a corollary of the following general and strong statement.

Theorem 8.2. *Let X_τ, Y_τ, $\tau \in [0, t_1]$, be nonautonomous vector fields, which are bounded w.r.t. τ and have a compact support. Let $0 \leq \alpha(\tau) \leq 1$ be a measurable function. Then there exists a sequence of nonautonomous vector fields $Z_\tau^n \in \{X_\tau, Y_\tau\}$, i.e., $Z_\tau^n = X_\tau$ or Y_τ for any τ and n, such that the flow*

$$\overrightarrow{\exp} \int_0^t Z_\tau^n \, d\tau \to \overrightarrow{\exp} \int_0^t \left(\alpha(\tau) X_\tau + (1 - \alpha(\tau)) Y_\tau\right) d\tau, \qquad n \to \infty,$$

uniformly w.r.t. $(t, q) \in [0, t_1] \times M$ and uniformly with all derivatives w.r.t. $q \in M$.

Now Proposition 8.1 follows: in the case $a_1(x) + a_2(x) = 1$ it is a corollary of Theorem 8.2, for the case $a_1(x), a_2(x) > 0$ we generalize by multiplication of control parameters by arbitrary positive function (this does not change attainable set for all nonnegative times), and the case $a_1(x), a_2(x) \geq 0$ is obtained by passage to limit.

Theorem 8.2 follows from the next two lemmas.

Lemma 8.1. *Under conditions of Theorem 8.2, there exists a sequence of nonautonomous vector fields $Z_\tau^n \in \{X_\tau, Y_\tau\}$ such that*

$$\int_0^t Z_\tau^n \, d\tau \to \int_0^t \left(\alpha(\tau) X_\tau + (1 - \alpha(\tau)) Y_\tau\right) d\tau$$

uniformly w.r.t. $(t, q) \in [0, t_1] \times M$ and uniformly with all derivatives w.r.t. $q \in M$.

Proof. Fix an arbitrary positive integer $n \in \mathbb{N}$. We can choose a covering of the segment $[0, t_1]$ by subsets

$$[0, t_1] = \bigcup_{i=1}^{N} E_i$$

such that

$$\forall i = 1, \dots, N \ \exists X_i, Y_i \in \operatorname{Vec} M \quad \text{s.t.} \quad \|X_\tau - X_i\|_{C^n} \le \frac{1}{n}, \quad \|Y_\tau - Y_i\|_{C^n} \le \frac{1}{n}.$$

Indeed, the fields X_τ, Y_τ are bounded in C^{n+1}-topology, thus they form a precompact set in C^n-topology.

Then divide E_i into n subsets of equal measure:

$$E_i = \cup_{j=1}^{n} E_{ij}, \quad |E_{ij}| = \frac{1}{n}|E_i|, \qquad i, j = 1, \dots, n.$$

In each E_{ij} pick a subset F_{ij} so that

$$F_{ij} \subset E_{ij}, \qquad |F_{ij}| = \int_{E_{ij}} \alpha(\tau)\, d\tau.$$

Finally, define the following vector field:

$$Z_\tau^n = \begin{cases} X_\tau, & \tau \in F_{ij}, \\ Y_\tau, & \tau \in E_{ij} \setminus F_{ij}. \end{cases}$$

Then the sequence of vector fields Z_τ^n is the required one. $\square$

Now we prove the second part of Theorem 8.2.

Lemma 8.2. *Let Z_τ^n, $n = 1, 2, \dots$, and Z_τ, $\tau \in [0, t_1]$, be nonautonomous vector fields on M, bounded w.r.t. τ, and let these vector fields have a compact support. If*

$$\int_0^t Z_\tau^n \, d\tau \to \int_0^t Z_\tau \, d\tau, \qquad n \to \infty,$$

then

$$\overrightarrow{\exp} \int_0^t Z_\tau^n \, d\tau \to \overrightarrow{\exp} \int_0^t Z_\tau \, d\tau, \qquad n \to \infty,$$

the both convergences being uniform w.r.t. $(t, q) \in [0, t_1] \times M$ and uniform with all derivatives w.r.t. $q \in M$.

Proof. (1) First we prove the statement for the case $Z_\tau = 0$. Denote the flow

$$P_t^n = \overrightarrow{\exp} \int_0^t Z_\tau^n \, d\tau.$$

Then

$$P_t^n = \operatorname{Id} + \int_0^t P_\tau^n \circ Z_\tau^n \, d\tau$$

integrating by parts

$$= \mathrm{Id} + P_t^n \circ \int_0^t Z_\tau^n \, d\tau - \int_0^t \left(P_\tau^n \circ Z_\tau^n \circ \int_0^\tau Z_\theta^\tau \, d\theta \right) \, d\tau.$$

Since $\int_0^t Z_\tau^n \, d\tau \to 0$, the last two terms above tend to zero, thus

$$P_t^n \to \mathrm{Id},$$

and the statement of the lemma in the case $Z_\tau = 0$ is proved.

(2) Now we consider the general case. Decompose vector fields in the sequence as follows:

$$Z_\tau^n = Z_\tau + V_\tau^n, \qquad \int_0^t V_\tau^n \, d\tau \to 0, \qquad n \to \infty.$$

Denote $P_t^n = \overrightarrow{\exp} \int_0^t V_\tau^n \, d\tau$. From the variations formula, we have

$$\overrightarrow{\exp} \int_0^t Z_\tau^n \, d\tau = \overrightarrow{\exp} \int_0^t (V_\tau^n + Z_\tau) \, d\tau = \overrightarrow{\exp} \int_0^t \mathrm{Ad}\, P_\tau^n \, Z_\tau \, d\tau \circ P_t^n.$$

Since $P_t^n \to \mathrm{Id}$ by part (1) of this proof and thus $\mathrm{Ad}\, P_t^n \to \mathrm{Id}$, we obtain the required convergence:

$$\overrightarrow{\exp} \int_0^t Z_\tau^n \, d\tau \to \overrightarrow{\exp} \int_0^t Z_\tau \, d\tau.$$

□

So we proved Theorem 8.2 and thus Proposition 8.1.

8.3 Poisson stability

Now we return to the study of controllability.

Definition 8.2. Let $f \in \mathrm{Vec}\, M$ be a complete vector field. A point $q \in M$ is called *Poisson stable* for f if for any $t > 0$ and any neighborhood O_q of q there exists a point $q' \in O_q$ and a time $t' > t$ such that $q' \circ e^{t'f} \in O_q$.

In other words, all trajectories cannot leave a neighborhood of a Poisson stable point forever, some of them must return to this neighborhood for arbitrarily large times.

Remark. If a trajectory $q \circ e^{tf}$ is periodic, then q is Poisson stable for f.

Definition 8.3. A complete vector field $f \in \mathrm{Vec}\, M$ is *Poisson stable* if all points of M are Poisson stable for f.

The condition of Poisson stability seems to be rather restrictive, but nevertheless there are surprisingly many Poisson stable vector fields in applications, see Poincaré's theorem below.

But first we prove a consequence of Poisson stability for controllability.

Proposition 8.2. *Let $\mathcal{F} \subset \operatorname{Vec} M$ be a bracket-generating system. If a vector field $f \in \mathcal{F}$ is Poisson stable, then the field $-f$ is compatible with $\mathcal{F}$.*

Proof. Choose an arbitrary point $q_0 \in M$ and a moment of time $t > 0$. To prove the statement, we should approximate the point $q_0 \circ e^{-tf}$ by reachable points.

Since $\mathcal{F}$ is bracket-generating, we can choose an open set $W \subset \operatorname{int} \mathcal{A}_{q_0}(\mathcal{F})$ arbitrarily close to q_0. Then the set $W \circ e^{-tf}$ is close enough to $q_0 \circ e^{-tf}$.

By Poisson stability, there exists $t' > t$ such that

$$\emptyset \neq \left(W \circ e^{-tf}\right) \circ e^{t'f} \cap W \circ e^{-tf} = W \circ e^{(t'-t)f} \cap W \circ e^{-tf}.$$

But $W \circ e^{(t'-t)f} \subset \mathcal{A}_{q_0}(\mathcal{F})$, thus

$$A_{q_0}(\mathcal{F}) \cap W \circ e^{-tf} \neq \emptyset.$$

So in any neighborhood of $q_0 \circ e^{-tf}$ there are points of the attainable set $\mathcal{A}_{q_0}(\mathcal{F})$, i.e., $q_0 \circ e^{-tf} \in \overline{\mathcal{A}_{q_0}(\mathcal{F})}$. □

Theorem 8.3 (Poincaré). *Let M be a smooth manifold with a volume form* Vol*. Let a vector field $f \in \operatorname{Vec} M$ be complete and its flow e^{tf} preserve volume. Let $W \subset M$, $W \subset \overline{\operatorname{int} W}$, be a subset of finite volume, invariant for f:*

$$\operatorname{Vol}(W) < \infty, \qquad W \circ e^{tf} \subset W \quad \forall t > 0.$$

Then all points of W are Poisson stable for f.

Proof. Take any point $q \in W$ and any its neighborhood $O \subset M$ of finite volume. The set $V = W \cap O$ contains an open nonempty subset $\operatorname{int} W \cap O$, thus $\operatorname{Vol}(V) > 0$. In order to prove the theorem, we show that

$$V \circ e^{t'f} \cap V \neq \emptyset \quad \text{for some large } t'.$$

Fix any $t > 0$. Then all sets

$$V \circ e^{ntf}, \qquad n = 0, 1, 2, \dots,$$

have the same positive volume, thus they cannot be disjoint. Indeed, if

$$V \circ e^{ntf} \cap V \circ e^{mtf} = \emptyset \qquad \forall n,\ m = 0, 1, 2, \dots,$$

then $\operatorname{Vol}(W) = \infty$ since all these sets are contained in W. Consequently, there exist nonnegative integers $n > m$ such that

$$V \circ e^{ntf} \cap V \circ e^{mtf} \neq \emptyset.$$

We multiply this inequality by e^{-mtf} from the right and obtain

$$V \circ e^{(n-m)tf} \cap V \neq \emptyset.$$

Thus the point q is Poisson stable for f. Since $q \in W$ is arbitrary, the theorem follows. $\square$

A vector field that preserves volume is called *conservative*.

Remark. Recall that on $M = \mathbb{R}^n = \{(x_1, \dots, x_n)\}$, a flow e^{tf} is conservative, i.e., preserves the standard volume $\operatorname{Vol}(V) = \int_V dx_1 \dots dx_n$ iff the field $f = \sum_{i=1}^n f_i \frac{\partial}{\partial x_i}$ is divergence-free:

$$\operatorname{div}_x f = \sum_{i=1}^n \frac{\partial f_i}{\partial x_i} = 0.$$

8.4 Controlled rigid body: attainable sets

We apply preceding general results on controllability to the control system that governs rotations of the rigid body, see (6.20):

$$\begin{pmatrix} \dot{Q} \\ \dot{\mu} \end{pmatrix} = f(Q, \mu) \pm g(Q, \mu), \qquad f = \begin{pmatrix} Q\widehat{\beta\mu} \\ \mu \times \beta\mu \end{pmatrix}, \qquad g = \begin{pmatrix} 0 \\ l \end{pmatrix}, \tag{8.5}$$
$$(Q, \mu) \in \mathrm{SO}(3) \times \mathbb{R}^3.$$

The vector field $f = \frac{1}{2}(f+g) + \frac{1}{2}(f-g)$ is compatible with system (8.5).

We show now that the field f is Poisson stable on $\mathrm{SO}(3) \times \mathbb{R}^3$.

Consider first the vector field $f(Q, \mu)$ on the larger space $\mathbb{R}^9_Q \times \mathbb{R}^3_\mu$, where $\mathbb{R}^9_Q$ is the space of all 3×3 matrices. Since $\operatorname{div}_{(Q,\mu)} f = 0$, the field f is conservative on $\mathbb{R}^9_Q \times \mathbb{R}^3_\mu$.

Further, since the first component of the field f is linear in Q, it has the following right-invariant property in Q:

$$e^{tf}\begin{pmatrix} Q \\ \mu \end{pmatrix} = \begin{pmatrix} Q_t \\ \mu_t \end{pmatrix} \quad \Rightarrow \quad e^{tf}\begin{pmatrix} PQ \\ \mu \end{pmatrix} = \begin{pmatrix} PQ_t \\ \mu_t \end{pmatrix}, \tag{8.6}$$
$$Q,\ Q_t,\ P \in \mathbb{R}^9_Q, \quad \mu,\ \mu_t \in \mathbb{R}^3_\mu.$$

In view of this property, the field f has compact invariant sets in $\mathbb{R}^9_Q \times \mathbb{R}^3_\mu$ of the form

$$W = K \cdot \mathrm{SO}(3) \times \{\, (\mu, \mu) \leq C \,\}, \quad K \Subset \mathbb{R}^9_Q,\ K \subset \overline{\operatorname{int} K},\ C > 0,$$

so that

$$W \subset \overline{\operatorname{int} W}.$$

By Poincaré's theorem, the field f is Poisson stable on all such sets W, thus on $\mathbb{R}^9_Q \times \mathbb{R}^3_\mu$. In view of the invariance property (8.6), the field f is Poisson stable on $\mathrm{SO}(3) \times \mathbb{R}^3$.

Since f is compatible with (8.5), then $-f$ is also compatible. The vector fields $\pm g = (f \pm g) - f$ are compatible with (8.5) as well. The vector fields $f \pm g$ are contained in the convex hull of the compatible vector fields found: $\mathrm{span}(f, g)$, a symmetric system. Thus if system (8.5) is bracket-generating, then its attainable set coincides with the orbit, i.e., the whole state space $\mathrm{SO}(3) \times \mathbb{R}^3$. So in the bracket-generating case system (8.5) is completely controllable.

In the non-bracket-generating cases, the structure of attainable sets is more complicated. If l is a principal axis of inertia, then the orbits of system (8.5) coincide with attainable sets. If $l \in \Pi_\pm \setminus \mathbb{R}e_2$, they do not coincide. This is easy to see from the phase portrait of the vector field $f(\mu) = \mu \times \beta\mu$ in the plane $\Pi_\pm$: the line $\mathbb{R}e_2$ consists of equilibria of f, and in the half-planes $\Pi_\pm \setminus \mathbb{R}e_2$ trajectories of f are semicircles centered at the origin.

Chapter 9

Feedback and state equivalence of control systems

9.1 Feedback equivalence

Consider control systems of the form

$$\dot{q} = f(q, u), \qquad q \in M, \ u \in U. \tag{9.1}$$

We suppose that not only M, but also U is a smooth manifold. For the right-hand side, we suppose that for all fixed $u \in U$, $f(q, u)$ is a smooth vector field on M, and, moreover, the mapping

$$(u, q) \mapsto f(q, u)$$

is smooth. Admissible controls are measurable locally bounded mappings

$$t \mapsto u(t) \in U$$

(for simplicity, one can consider piecewise continuous controls). If such a control $u(t)$ is substituted to control system (9.1), one obtains a nonautonomous ODE

$$\dot{q} = f(q, u(t)), \tag{9.2}$$

with the right-hand side smooth in q and measurable, locally bounded in t. For such ODEs, there holds a standard theorem on existence and uniqueness of solutions, at least local. Solutions $q(\cdot)$ to ODEs (9.2) are Lipschitzian curves in M.

In Section 5.7 we already considered state transformations of control systems, i.e., diffeomorphisms of M. State transformations map trajectories of control systems to trajectories, with the same control. Now we introduce a new class of feedback transformations, which also map trajectories to trajectories, but possibly with a new control.

Denote the space of new control parameters by $\widehat{U}$. We assume that it is a smooth manifold.

Definition 9.1. Let $\varphi : M \times \widehat{U} \to U$ be a smooth mapping. A transformation of the form

$$f(q,u) \mapsto f(q,\varphi(q,\widehat{u})), \qquad q \in M, \quad u \in U, \quad \widehat{u} \in \widehat{U},$$

is called a *feedback transformation*.

Remark. A feedback transformation reparametrizes control u in a way depending on q.

It is easy to see that any admissible trajectory $q(\cdot)$ of the system $\dot{q} = f(q,\varphi(q,\widehat{u}))$ corresponding to a control $\widehat{u}(\cdot)$ is also admissible for the system $\dot{q} = f(q,u)$ with the control $u(\cdot) = \varphi(q(\cdot),\widehat{u}(\cdot))$, but, in general, not vice versa.

In order to consider feedback equivalence, we consider invertible feedback transformations with

$$\widehat{U} = U, \qquad \varphi|_{q\times U} \in \operatorname{Diff} U.$$

Such mappings $\varphi : M \times U \to U$ generate feedback transformations

$$f(q,u) \mapsto f(q,\varphi(q,u)).$$

The corresponding control systems

$$\dot{q} = f(q,u) \quad \text{and} \quad \dot{q} = f(q,\varphi(q,u))$$

are called *feedback equivalent*.

Our aim is to simplify control systems with state and feedback transformations.

Remark. In mathematical physics, feedback transformations are often called gauge transformations.

Consider affine in control systems

$$\dot{q} = f(q) + \sum_{i=1}^{k} u_i g_i(q), \qquad u = (u_1,\dots,u_k) \in \mathbb{R}^k, \ q \in M. \tag{9.3}$$

To such systems, it is natural to apply feedback transformations affine in control:

$$\varphi = (\varphi_1,\dots,\varphi_k) : M \times \mathbb{R}^k \to \mathbb{R}^k,$$

$$\varphi_i(q,u) = c_i(q) + \sum_{j=1}^{k} d_{ij}(q) u_j, \qquad i = 1,\dots,k. \tag{9.4}$$

Our aim is to characterize affine in control systems (9.3) which are locally equivalent to linear controllable systems w.r.t. state and feedback transformations (9.4) and to classify them w.r.t. this class of transformations.

First we consider linear controllable systems

$$\dot{x} = Ax + \sum_{i=1}^{k} u_i b_i, \qquad x \in \mathbb{R}^n, \ u = (u_1, \dots, u_k) \in \mathbb{R}^k, \tag{9.5}$$

where A is an $n \times n$ matrix and b_i, $i = 1, \dots, k$, are vectors in $\mathbb{R}^n$. We assume that the vectors $b_1, \dots, b_k$ are linearly independent:

$$\dim \operatorname{span}(b_1, \dots, b_k) = k.$$

If this is not the case, we eliminate some b_i's. We find normal forms of linear systems w.r.t. linear state and feedback transformations.

To linear systems (9.5) we apply feedback transformations which have the form (9.4) and, moreover, preserve the linear structure:

$$\begin{aligned} c_i(x) &= \langle c_i, x \rangle, \qquad && c_i \in \mathbb{R}^{n*}, \ i = 1, \dots, k, \\ d_{ij}(x) &= d_{ij} \in \mathbb{R}, && i, j = 1, \dots, k. \end{aligned} \tag{9.6}$$

Denote by $D : \operatorname{span}(b_1, \dots, b_k) \to \operatorname{span}(b_1, \dots, b_k)$ the linear operator with the matrix (d_{ij}) in the base $b_1, \dots, b_k$. Linear feedback transformations (9.4), (9.6) map the vector fields in the right-hand side of the linear system (9.5) as follows:

$$(Ax, b_1, \dots, b_k) \mapsto \left(Ax + \sum_{i=1}^{k} \langle c_i, x \rangle b_i, Db_1, \dots, Db_k \right). \tag{9.7}$$

Such mapping should be invertible, so we assume that the operator D (or, equivalently, its matrix (d_{ij})) is invertible.

Linear state transformations act on linear systems as follows:

$$(Ax, b_1, \dots, b_k) \mapsto \left(CAC^{-1}x, Cb_1, \dots, Cb_k \right), \tag{9.8}$$

where $C : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear operator. State equivalence of linear systems means that these systems have the same coordinate representation in suitably chosen bases in the state space $\mathbb{R}^n$.

9.2 Linear systems with scalar control

Consider a simple model linear control system — scalar high-order control:

$$x^{(n)} + \sum_{i=0}^{n-1} \alpha_i x^{(i)} = u, \qquad u \in \mathbb{R}, \ x \in \mathbb{R}, \tag{9.9}$$

where $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{R}$. We rewrite this system in the standard form in the variables $x_i = x^{(i-1)}$, $i = 1, \dots, n$:

$$\begin{cases} \dot{x}_1 = x_2, \\ \dots \\ \dot{x}_{n-1} = x_n, \\ \dot{x}_n = -\sum_{i=0}^{n-1} \alpha_i x_{i+1} + u, \end{cases} \qquad u \in \mathbb{R}, \ x = (x_1, \dots, x_n) \in \mathbb{R}^n. \tag{9.10}$$

It is easy to see that if we take $-\sum_{i=1}^{n-1} \alpha_i x_{i+1} + u$ as a new control, i.e., apply the feedback transformation (9.4), (9.6) with

$$k = 1, \quad c = (-\alpha_0, \dots, -\alpha_{n-1}), \quad d = 1,$$

then system (9.10) maps into the system

$$\begin{cases} \dot{x}_1 = x_2, \\ \dots \\ \dot{x}_{n-1} = x_n, \\ \dot{x}_n = u, \end{cases} \qquad u \in \mathbb{R}, \ x = (x_1, \dots, x_n) \in \mathbb{R}^n, \tag{9.11}$$

which is written in the scalar form as

$$x^{(n)} = u, \qquad u \in \mathbb{R}, \ x \in \mathbb{R}. \tag{9.12}$$

So system (9.10) is feedback equivalent to system (9.11).

It turns out that the simple systems (9.10) and (9.11) are normal forms of linear controllable systems with scalar control under state transformations and state-feedback transformations respectively.

Proposition 9.1. *Any linear controllable system with scalar control*

$$\dot{x} = Ax + ub, \qquad u \in \mathbb{R}, \ x \in \mathbb{R}^n, \tag{9.13}$$
$$\operatorname{span}(b, Ab, \dots, A^{n-1}b) = \mathbb{R}^n, \tag{9.14}$$

is state equivalent to a system of the form (9.10), *thus state-feedback equivalent to system* (9.11).

Proof. We find a basis $e_1, \dots, e_n$ in $\mathbb{R}^n$ in which system (9.13) is written in the form (9.10). Coordinates $y_1, \dots, y_n$ of a point $x \in \mathbb{R}^n$ in a basis $e_1, \dots, e_n$ are found from the decomposition

$$x = \sum_{i=1}^{n} y_i e_i.$$

In view of the desired form (9.10), the vector b should have coordinates $b = (0, \dots, 0, 1)^*$, thus the n-th basis vector is uniquely determined:

$$e_n = b.$$

Now we find the rest basis vectors $e_1, \dots, e_{n-1}$. We can rewrite our linear system (9.13) as follows:
$$\dot{x} = Ax \mod \mathbb{R}\, b,$$
then we obtain in coordinates:
$$\dot{x} = \sum_{i=1}^{n} \dot{y}_i e_i = \sum_{i=1}^{n} y_i A e_i \mod \mathbb{R}\, b,$$
thus
$$\sum_{i=1}^{n-1} \dot{y}_i e_i = \sum_{i=0}^{n-1} y_{i+1} A e_{i+1} \mod \mathbb{R}\, b.$$
The required differential equations:
$$\dot{y}_i = y_{i+1}, \qquad i = 1, \dots, n-1,$$
are fulfilled in a basis $e_1, \dots, e_n$ if and only if the following equalities hold:
$$Ae_{i+1} = e_i + \beta_i b, \qquad i = 1, \dots, n-1, \tag{9.15}$$
$$Ae_1 = \beta_0 b \tag{9.16}$$
for some numbers $\beta_0, \dots, \beta_{n-1} \in \mathbb{R}$.

So it remains to show that we can find basis vectors $e_1, \dots, e_{n-1}$ which satisfy equalities (9.15), (9.16). We rewrite equality (9.15) in the form
$$e_i = Ae_{i+1} - \beta_i b, \qquad i = 1, \dots, n-1, \tag{9.17}$$
and obtain recursively:
$$\begin{aligned}
e_n &= b, \\
e_{n-1} &= Ab - \beta_{n-1} b, \\
e_{n-2} &= A^2 b - \beta_{n-1} Ab - \beta_{n-2} b, \\
&\dots \\
e_1 &= A^{n-1} b - \beta_{n-1} A^{n-2} b - \cdots - \beta_1 b.
\end{aligned} \tag{9.18}$$
So equality (9.16) yields
$$Ae_1 = A^n b - \beta_{n-1} A^{n-1} b - \cdots - \beta_1 Ab = \beta_0 b.$$
The equality
$$A^n b = \sum_{i=0}^{n-1} \beta_i A^i b \tag{9.19}$$
is satisfied for a unique n-tuple $(\beta_0, \dots, \beta_{n-1})$ since the vectors $b, Ab, \dots, A^{n-1}b$ form a basis of $\mathbb{R}^n$ (in fact, β_i are coefficients of the characteristic polynomial of A).

With these numbers β_i, the vectors $e_1, \dots, e_n$ given by (9.18) form the required basis. Indeed, equalities (9.15), (9.16) hold by construction. The vectors $e_1, \dots, e_n$ are linearly independent by the controllability condition (9.14). □

Remark. The basis $e_1, \dots, e_n$ constructed in the previous proof is unique, thus the state transformation that maps a controllable linear system with scalar control (9.13) to the normal form (9.10) is also unique.

9.3 Linear systems with vector control parameters

Now consider general controllable linear systems:

$$\dot{x} = Ax + \sum_{i=1}^{k} u_i b_i, \qquad x \in \mathbb{R}^n, \ u = (u_1, \dots, u_k) \in \mathbb{R}^k, \tag{9.20}$$

$$\operatorname{span}\{A^j b_i \mid j = 0, \dots, n-1, \ i = 1, \dots, k\} = \mathbb{R}^n. \tag{9.21}$$

Recall that we assume vectors $b_1, \dots, b_k$ linearly independent.

In the case $k = 1$, all controllable linear systems in $\mathbb{R}^n$ are state-feedback equivalent to the normal form (9.11), thus there are no state-feedback invariants in a given dimension n.

If $k > 1$, this is not the case, and we start from description of state-feedback invariants.

Consider the following subspaces in $\mathbb{R}^n$:

$$D^m = \operatorname{span}\{A^j b_i \mid j = 0, \dots, m-1, \ i = 1, \dots, k\}, \qquad m = 1, \dots, n. \tag{9.22}$$

Invertible linear state transformations (9.8) preserve dimension of these subspaces, thus the numbers

$$\dim D^m, \qquad m = 1, \dots, n,$$

are state invariants.

Now we show that invertible linear feedback transformations (9.7) preserve the spaces D^m, $m = 1, \dots, n$. Any such transformation can be decomposed into two feedback transformations of the form:

$$(Ax, b_1, \dots, b_k) \mapsto (Ax + \sum_{i=1}^{k} \langle c_i, x \rangle b_i, b_1, \dots, b_k), \tag{9.23}$$

$$(Ax, b_1, \dots, b_k) \mapsto (Ax, Db_1, \dots, Db_k). \tag{9.24}$$

Transformations (9.24), i.e., changes of b_i, obviously preserve the spaces D^m. Consider transformations (9.23). Denote the new matrix:

$$\widehat{A}x = Ax + \sum_{i=1}^{k} \langle c_i, x \rangle b_i.$$

We have:

$$\widehat{A}^j x = A^j x \mod D^j, \qquad j = 1, \dots, n-1.$$

But $D^{m-1} \subset D^m$, $m = 2, \dots, n$, thus feedback transformations (9.23) preserve the spaces D^m, $m = 1, \dots, n$.

So the spaces D^m, $m = 1, \dots, n$, are invariant under feedback transformations, and their dimensions are state-feedback invariants.

Now we express the numbers $\dim D^m$, $m = 1, \dots, n$, through other integers — Kronecker indices. Construct the following $n \times k$ matrix:

$$\begin{pmatrix} b_1 & \cdots & b_k \\ Ab_1 & \cdots & Ab_k \\ \vdots & \vdots & \vdots \\ A^{n-1}b_1 & \cdots & A^{n-1}b_k \end{pmatrix}. \tag{9.25}$$

Replace each vector $A^j b_i$, $j = 0, \dots, n-1$, $i = 1, \dots, k$, in this matrix by a sign: cross $\times$ or circle $\circ$, by the following rule. We go in matrix (9.25) by raws, i.e., order its elements as follows:

$$b_1, \dots, b_k, Ab_1, \dots, Ab_k, \dots, A^{n-1}b_1, \dots, A^{n-1}b_k. \tag{9.26}$$

A vector $A^j b_i$ in matrix (9.25) is replaced by $\times$ if it is linearly independent of the previous vectors in chain (9.26), otherwise it is replaced by $\circ$. After this procedure we obtain a matrix of the form:

$$\Sigma = \begin{pmatrix} \times & \times & \times & \times & \cdots & \times \\ \times & \circ & \times & \times & \cdots & \circ \\ \times & \circ & \circ & \times & \cdots & \circ \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \circ & \circ & \circ & \times & \cdots & \circ \end{pmatrix}.$$

Notice that there are some restrictions on appearance of crosses and circles in matrix Σ. The total number of crosses in this matrix is n (by the controllability condition (9.21)), and the first row is filled only with crosses (since $b_1, \dots, b_k$ are linearly independent). Further, if a column of Σ contains a circle, then all elements below it are circles as well. Indeed, if a vector $A^j b_i$ in (9.25) is replaced by circle in Σ, then

$$A^j b_i \in \operatorname{span}\{A^j b_\alpha \mid \alpha < i\} + \operatorname{span}\{A^\beta b_\alpha \mid \beta < j,\ \alpha = 1, \dots, k\}.$$

Then the similar inclusions hold for all vectors $A^{j+1}b_i, \dots, A^{n-1}b_i$, i.e., below circles are only circles. So each column in the matrix Σ consists of a column of crosses over a column of circles (the column of circles can be absent).

Denote by n_1 the height of the highest column of crosses in the matrix Σ, by n_2 the height of the next highest column of crosses, ..., and by n_k the height of the lowest column of crosses in Σ. The positive integers obtained:

$$n_1 \geq n_2 \geq \cdots \geq n_k$$

are called *Kronecker indices* of the linear control system (9.20). Since the total number of crosses in matrix Σ is equal to dimension of the state space, then

$$\sum_{i=1}^{k} n_i = n = \dim \mathbb{R}^n.$$

Moreover, by the construction, we have

$$\operatorname{span}(b_1, Ab_1, \ldots, A^{n_1-1}b_1; \ldots; b_k, Ab_k, \ldots, A^{n_k-1}b_k) = \mathbb{R}^n. \tag{9.27}$$

Now we show that Kronecker indices n_i are expressed through the numbers $\dim D^i$. We have:

$$\begin{aligned}
\dim D^1 &= k = \text{number of crosses in the first row of } \Sigma,\\
\dim D^2 &= \text{number of crosses in the first 2 rows of } \Sigma,\\
&\ldots\\
\dim D^i &= \text{number of crosses in the first } i \text{ rows of } \Sigma,
\end{aligned}$$

so that

$$\Delta(i) = \dim D^i - \dim D^{i-1} = \text{ number of crosses in the } i\text{-th row of } \Sigma.$$

Permute columns in matrix Σ, so that the first column become the highest one, the second column becomes the next highest one, etc. We obtain an $n \times k$-matrix in the "block-triangular" form. This matrix rotated at the angle $\pi/2$ gives the subgraph of the function $\Delta : \{1, \ldots, n\} \to \{1, \ldots, k\}$. It is easy to see that the values of the Kronecker indices is equal to the points of jumps of the function Δ, and the number of Kronecker indices for each value is equal to the height of the corresponding jump of Δ.

So Kronecker indices are expressed through $\dim D^i$, $i = 1, \ldots, k$, thus are state-feedback invariants.

Now we show that the set of Kronecker indices n_i, $i = 1, \ldots, k$, is a complete set of state-feedback invariants of controllable linear systems (9.20).

We show first that any linear controllable system (9.20) can be written, in a suitable basis in $\mathbb{R}^n$:

$$e_1^1, \ldots, e_{n_1}^1; \ldots, ; e_1^k, \ldots, e_{n_k}^k \tag{9.28}$$

in the following canonical form:

$$\begin{cases}
\dot y_1^1 = y_2^1,\\
\ldots\\
\dot y_{n_1-1}^1 = y_{n_1}^1,\\
\dot y_{n_1}^1 = -\displaystyle\sum_{\substack{1\le j\le k\\ 0\le i\le n_j-1}} \gamma_{ij}^1 y_{i+1}^j + u_1,
\end{cases}
\quad , \quad \ldots, \quad
\begin{cases}
\dot y_1^k = y_2^k,\\
\ldots\\
\dot y_{n_k-1}^k = y_{n_k}^k,\\
\dot y_{n_k}^k = -\displaystyle\sum_{\substack{1\le j\le k\\ 0\le i\le n_j-1}} \gamma_{ij}^k y_{i+1}^j + u_k,
\end{cases}
\tag{9.29}$$

where

$$x = \sum_{\substack{1 \le i \le k \\ 1 \le j \le n_i}} y_j^i e_j^i. \tag{9.30}$$

We proceed exactly as in the scalar-input case (Section 9.2). The required canonical form (9.29) determines uniquely the last basis vectors in all k groups:

$$e_{n_1}^1 = b_1, \ldots, e_{n_k}^k = b_k. \tag{9.31}$$

Denote the space $B = \operatorname{span}(b_1, \ldots, b_k)$. Then our system

$$\dot{x} = Ax \mod B$$

reads in coordinates as follows:

$$\dot{x} = \sum_{\substack{1 \le i \le k \\ 1 \le j \le n_i}} \dot{y}_j^i e_j^i = \sum_{\substack{1 \le i \le k \\ 1 \le j \le n_i}} y_j^i A e_j^i \mod B.$$

In view of the required equations

$$\dot{y}_j^i = y_{j+1}^i, \qquad 1 \le i \le k, \quad 1 \le j < n_i,$$

we have

$$\sum_{\substack{1 \le i \le k \\ 1 \le j < n_i}} y_{j+1}^i e_j^i = \sum_{\substack{1 \le i \le k \\ 1 \le j \le n_i}} y_j^i A e_j^i \mod B,$$

or, equivalently,

$$\sum_{\substack{1 \le i \le k \\ 2 \le j \le n_i}} y_j^i e_{j-1}^i = \sum_{\substack{1 \le i \le k \\ 1 \le j \le n_i}} y_j^i A e_j^i \mod B.$$

So the following relations should hold for the required basis vectors:

$$Ae_j^i = e_{j-1}^i \mod B, \qquad 1 \le i \le k, \quad 2 \le j \le n_i, \tag{9.32}$$

$$Ae_1^i = 0 \mod B, \qquad 1 \le i \le k. \tag{9.33}$$

We resolve equations (9.32) recursively starting from (9.31), for all $i = 1, \ldots, k$:

$$e_{n_i}^i = b_i,$$

$$e_{n_i-1}^i = Ab_i - \sum_{\alpha=1}^{n} \beta_{i,n_i-1}^{\alpha} b_\alpha,$$

$$e_{n_i-2}^i = A^2 b_i - \sum_{\alpha=1}^{n} \beta_{i,n_i-1}^{\alpha} A b_\alpha - \sum_{\alpha=1}^{n} \beta_{i,n_i-2}^{\alpha} b_\alpha,$$

$$\ldots$$

$$e_1^i = A^{n_i-1} b_i - \sum_{\alpha=1}^{n} \beta_{i,n_i-1}^{\alpha} A^{n_i-2} b_\alpha - \cdots - \sum_{\alpha=1}^{n} \beta_{i,1}^{\alpha} b_\alpha,$$

while (9.33) yields

$$Ae_1^i = \sum_{\alpha=1}^{n} \beta_{i,0}^{\alpha} b_\alpha$$

for some constants $\beta_{i,j}^{\alpha}$, $1 \le i \le k$, $0 \le j \le n_i$, $1 \le \alpha \le n$. We obtain the equation

$$A^{n_i} b_i = \sum_{\alpha=1}^{n} \beta_{i,n_i-1}^{\alpha} A^{n_i-1} b_\alpha + \cdots + \sum_{\alpha=1}^{n} \beta_{i,0}^{\alpha} b_\alpha,$$

which has a unique solution in $\beta_{i,j}^{\alpha}$ in view of (9.27).

So we proved that there exists a unique linear state transformation that maps a linear controllable system (9.20) to the canonical form (9.29).

Choosing $-\sum_{\substack{1\le j\le k \\ 0\le i\le n_j-1}} \gamma_{ij}^{l} y_{i+1}^{j} + u_l$, $l = 1, \ldots, k$, as new controls, we see that each of the k subsystems in (9.29) is feedback equivalent to a system of the form (9.11), or, equivalently, (9.12).

Thus the whole system (9.20) is state-feedback equivalent to the following normal form:

$$\begin{cases} y_1^{(n_1)} = u_1, \\ \cdots \\ y_k^{(n_k)} = u_k, \end{cases} \tag{9.34}$$

called *Brunovsky normal form*.

We proved the following statement.

Theorem 9.1. *Any controllable linear system* (9.20), (9.21) *with k control parameters is state equivalent to a system of the form* (9.29) *and state-feedback equivalent to a system in Brunovsky normal form* (9.34)*, where n_i, $i = 1, \ldots, k$, are Kronecker indices of system* (9.20)*.*

So the action of the group of invertible state-feedback transformations on controllable linear systems with k control parameters has only discrete invariants — Kronecker indices $n_1, \ldots, n_k$.

9.4 State-feedback linearizability

Consider a nonlinear affine in control system:

$$\dot{q} = f(q) + \sum_{j=1}^{k} u_j g_j(q), \qquad u = (u_1, \ldots, u_k) \in \mathbb{R}^k, \ q \in M. \tag{9.35}$$

We are interested, when such a system is locally state-feedback equivalent to a controllable linear system.

Definition 9.2. System (9.35) is called locally state-feedback equivalent to a linear system (9.20) in a neighborhood of a point $q_0 \in M$, if there exist a state transformation — a diffeomorphism

$$\Phi : O_{q_0} \to \widehat{O} \subset \mathbb{R}^n$$

from a neighborhood O_{q_0} of q_0 in M onto an open subset $\widehat{O} \subset \mathbb{R}^n$, and a feedback transformation

$$\varphi : O_{q_0} \times \mathbb{R}^k \to \mathbb{R}^k,$$
$$\varphi(q, u) = \begin{pmatrix} a_1(q) \\ \cdots \\ a_k(q) \end{pmatrix} + D(q)u, \tag{9.36}$$

with an invertible and smooth in q matrix

$$D(q) = (d_{ij}(q)), \qquad i, \ j = 1, \dots, k,$$

such that the state-feedback transformation (Φ, φ) maps system (9.35) restricted to O_{q_0} to a linear system (9.20) restricted to $\widehat{O}$.

We can generalize the construction of the subspaces D^m (9.22) for the case of nonlinear systems (9.35): consider the families of subspaces

$$D_q^m = \operatorname{span}\{(\operatorname{ad} f)^j g_i(q) \mid j = 0, \dots, m-1, \ i = 1, \dots, k\} \subset T_q M.$$

Notice that, in general, $\dim D_q^m \neq \text{const}$, thus D^m is not a distribution.

Observe that for controllable linear systems (9.20), the following properties hold for the family $D_x^m \equiv D^m$, $x \in \mathbb{R}^n$:

1. $\dim D_x^m = \text{const}$,
2. $D_x^n = T_x\mathbb{R}^n$,
3. the distributions D^m, $m = 1, \dots, n$, are integrable (since they are spanned by the constant vector fields $A^j b_i$).

Before formulating conditions for state-feedback linearizability of nonlinear systems, which are given in terms of the families D_q^m, we prove the following property of these families.

Lemma 9.1. *If the families D^m, $m = 1, \dots, n$, are involutive, then they are feedback-invariant.*

Proof. Notice first that feedback transformations (9.36) can be decomposed into transformations of the two kinds:

$$(f, g_1, \dots, g_k) \mapsto (f + a_j g_j, g_1, \dots, g_k), \tag{9.37}$$
$$(f, g_1, \dots, g_k) \mapsto (f, Dg_1, \dots, Dg_k), \tag{9.38}$$

where $D(q) = (d_{ij}(q))$, $i, j = 1, \dots, k$, is invertible and smooth w.r.t. q. We prove the lemma by induction on m.

Let $m = 1$. The family

$$D^1 = \operatorname{span}\{g_i \mid i = 1, \dots, k\}$$

is obviously preserved by the both transformations (9.37) and (9.38).

Induction step: we assume that the statement is proved for $m-1$ and prove it for m. The family

$$D^m = \{[f, X] \mid X \in D^{m-1}\} + D^{m-1}$$

is preserved by transformation (9.38). Consider transformation (9.37). We have

$$[f + a_j g_j, X] = [f, X] - [X, a_j g_j] = [f, X] - (X a_j) g_j - a_j [X, g_j].$$

Further:

$$\begin{aligned}
&X \in D^{m-1} \;\Rightarrow\; [f, X] \in D^m, \\
&(X a_j) g_j \in D^1 \subset D^m, \\
&X \in D^{m-1},\ g_j \in D^1 \subset D^{m-1} \;\Rightarrow\; [X, g_j] \in D^{m-1} \subset D^m,
\end{aligned}$$

thus

$$[f + a_j g_j, X] \in D^m \quad \forall\, X \in D^{m-1}.$$

So D^m is preserved by feedback transformation (9.37). □

Theorem 9.2. *System* (9.35) *is locally state-feedback equivalent to a controllable linear system* (9.20) *if and only if:*

(1) $\dim D_q^m$, $m = 1, \dots, n$, *does not depend on* q, *i.e.,* D^m *are distributions,*

(2) $D_q^n = T_q M$,

(3) *the distributions* D^m, $m = 1, \dots, n$, *are involutive.*

Conditions (1)–(3) are necessary for local state-feedback linearizability, see discussion before Lemma 9.1.

We prove sufficiency in Theorem 9.3 below only in the case of scalar control parameter. For $k = 1$ we have the system

$$\dot q = f(q) + u g(q), \qquad u \in \mathbb{R},\ q \in M, \tag{9.39}$$

and the corresponding families of subspaces

$$D_q^m = \operatorname{span}\{(\operatorname{ad} f)^i g(q) \mid i = 0, 1, \dots, m-1\}, \qquad m = 1, \dots, n,\ q \in M.$$

In this case it happens that involutivity of D^{n-1} implies involutivity of D^m with smaller m.

Theorem 9.3. *System* (9.39) *is locally state-feedback equivalent to a controllable linear system* (9.13) *if and only if:*

(1) $D_q^n = T_q M$,

(2) *the distribution* D^{n-1} *is involutive.*

First we prove the following proposition of general interest: integral manifolds of integrable distributions can be smoothly parametrized.

Lemma 9.2. *Let* $\Delta = \operatorname{span}\{X_1, \dots, X_k\}$ *be an integrable distribution on a smooth* n*-dimensional manifold* M, $\dim \Delta_q = k$. *Then for any point* $q_0 \in M$ *there exist a neighborhood* $q_0 \in O_{q_0} \subset M$ *and a smooth vector-function*

$$\varphi \,:\, O_{q_0} \to \mathbb{R}^{n-k}$$

such that:

(1) $\operatorname{rank} \varphi_{*q} = n - k$, $q \in O_{q_0}$, *and*

(2) $\varphi^{-1}(y)$ *is an integral manifold of* Δ *for any* $y \in \varphi(O_{q_0})$, *or, equivalently,*

(2′) $\ker \varphi_{*q} = \Delta_q$, $q \in O_{q_0}$.

Proof. Complete the vector fields $X_1, \dots, X_k$ to a basis:

$$\operatorname{span}\{Y_1, \dots, Y_{n-k}, X_1, \dots, X_k\} = \operatorname{Vec} O_{q_0},$$

for a sufficiently small neighborhood $q_0 \in O_{q_0} \subset M$. Consider the mapping

$$\begin{aligned} \psi \,:\, (t, s) &\mapsto q_0 \circ e^{t_1 Y_1} \circ \dots \circ e^{t_{n-k} Y_{n-k}} \circ e^{s_1 X_1} \circ \dots \circ e^{s_k X_k}, \\ t &= (t_1, \dots, t_{n-k}) \in \mathbb{R}^{n-k}, \ s = (s_1, \dots, s_k) \in \mathbb{R}^k. \end{aligned}$$

We have

$$\begin{aligned} \left.\frac{\partial \psi}{\partial t_i}\right|_0 &= Y_i, \qquad i = 1, \dots, n-k, \\ \left.\frac{\partial \psi}{\partial s_i}\right|_0 &= X_i, \qquad i = 1, \dots, k, \end{aligned}$$

thus ψ is a local diffeomorphism in a neighborhood of $0 \in \mathbb{R}^n$.

Further, for fixed $t = t^0$, the set

$$\{\psi(t^0, s) \mid s \in \mathbb{R}^k\}$$

is an integral manifold of Δ.

Finally, locally, by the implicit function theorem, there exists a well-defined smooth mapping

$$\varphi \,:\, \psi(t, s) \mapsto t.$$

It is the required vector-function. □

Now we prove Theorem 9.3.

Proof. Necessity is already known since for linear controllable systems both conditions (1), (2) hold, see discussion before Lemma 9.1.

To prove sufficiency, we construct coordinates in which our system (9.39) is simplified, and then apply a feedback transformation which maps this system to the normal form (9.11).

Since the distribution D^{n-1} is integrable, then by Lemma 9.2 there exists a smooth function

$$\varphi_1 : O_{q_0} \to \mathbb{R}$$

such that

$$d_q\varphi_1 \neq 0, \quad \langle d_q\varphi_1, D_q^{n-1}\rangle = 0, \qquad q \in O_{q_0}. \tag{9.40}$$

Define the following functions in the neighborhood O_{q_0}:

$$\begin{aligned}
&\varphi_2 = f\varphi_1 = \langle d\varphi_1, f\rangle, \\
&\varphi_3 = f\varphi_2 = f^2\varphi_1, \\
&\dots \\
&\varphi_n = f\varphi_{n-1} = f^{n-1}\varphi_1
\end{aligned}$$

(iterated directional derivatives along the vector field f).

We claim that the functions $\varphi_1, \dots, \varphi_n$ (which will be the coordinates that simplify (9.39)) have the following property:

$$(\operatorname{ad} f)^j g\varphi_l = \begin{cases} 0, & j + l < n, \\ \pm(\operatorname{ad} f)^{n-1} g\varphi_1 \neq 0, & j + l = n. \end{cases} \tag{9.41}$$

First of all, notice that $b = (\operatorname{ad} f)^{n-1} g\varphi_1\big|_{O_{q_0}} \neq 0$. Indeed, we have

$$\begin{aligned}
&D_q^{n-1} = \operatorname{span}\{g(q), \dots, (\operatorname{ad} f)^{n-2} g(q)\}, \\
&T_q M = \operatorname{span}\{g(q), \dots, (\operatorname{ad} f)^{n-1} g(q)\} = \operatorname{span}\{D_q^{n-1}, (\operatorname{ad} f)^{n-1} g(q)\},
\end{aligned}$$

thus the equality $(\operatorname{ad} f)^{n-1} g\varphi_1(q) = 0$ is incompatible with properties (9.40).

Now we prove (9.41) by induction on l. If $l = 1$, there is nothing to prove.

Assume that equality (9.41) is proved for $l - 1$ and prove it for l. We have

$$\begin{aligned}
(\operatorname{ad} f)^j g\varphi_l &= \left((\operatorname{ad} f)^j g \circ f\right)\varphi_{l-1} \\
&= \left((\operatorname{ad} f)^j g \circ f - f \circ (\operatorname{ad} f)^j g + f \circ (\operatorname{ad} f)^j g\right)\varphi_{l-1} \\
&= \left(-[f, (\operatorname{ad} f)^j g] + f \circ (\operatorname{ad} f)^j g\right)\varphi_{l-1} \\
&= \left(-(\operatorname{ad} f)^{j+1} g + f \circ (\operatorname{ad} f)^j g\right)\varphi_{l-1}.
\end{aligned}$$

If $j+l \leq n$, then $j+l-1 < n$, and $(\operatorname{ad} f)^j g\varphi_{l-1} = 0$ by the induction assumption. Thus

$$(\operatorname{ad} f)^j g\varphi_l = -(\operatorname{ad} f)^{j+1} g\varphi_{l-1} \qquad \text{for } j + l \leq n,$$

and equality (9.41) for l follows from this equality for $l-1$.

So equality (9.41) is proved for all l. Since the vectors $g(q), \dots, (\operatorname{ad} f)^{n-1} g(q)$ span the tangent space $T_q M$ for $q \in O_{q_0}$, the mapping

$$\Phi = \begin{pmatrix} \varphi_1 \\ \cdots \\ \varphi_n \end{pmatrix} : O_{q_0} \to \mathbb{R}^n$$

is a local diffeomorphism: the differentials $d_q \varphi_1, \dots, d_q \varphi_n$ form a basis of $T_q^* M$ dual to $g(q), \dots, (\operatorname{ad} f)^{n-1} g(q) \in T_q M$.

Take Φ as a coordinate mapping, then coordinates of a point $q \in M$ are

$$x_l = \varphi_l(q), \qquad l = 1, \dots, n.$$

Now we write our system $\dot{q} = f(q) + u g(q)$ in these coordinates: we differentiate x_l with respect to this system.

$$\frac{d}{dt} x_l = \frac{d}{dt} \varphi_l(q(t)) = (f + ug)\varphi_l = f\varphi_l + ug\varphi_l.$$

If $l < n$, then $g\varphi_l = 0$ by equality (9.41), thus

$$\frac{d}{dt} x_l = f\varphi_l = \varphi_{l+1} = x_{l+1}, \qquad l = 1, \dots, n-1.$$

And if $l = n$, then

$$\frac{d}{dt} x_n = f\varphi_n + ug\varphi_n = f\varphi_n \pm ub, \qquad b = g\varphi_n \neq 0.$$

So in coordinates $x_1, \dots, x_n$ our system (9.39) takes the form

$$\begin{cases} \dot{x}_1 = x_2, \\ \cdots \\ \dot{x}_{n-1} = x_n, \\ \dot{x}_n = f\varphi_n \pm ub. \end{cases}$$

Now consider the feedback transformation

$$u \mapsto \mp \frac{f\varphi_n - u}{b}.$$

After this transformation the n-th component of our system reads

$$\dot{x}_n = f\varphi_n \pm \left(\mp \frac{f\varphi_n - u}{b} \right) b = f\varphi_n - f\varphi_n + u = u,$$

i.e., the whole system takes the required form (9.11). $\square$

Chapter 10

Optimal control problem

10.1 Problem statement

Consider a control system of the form

$$\dot{q} = f_u(q), \qquad q \in M, \quad u \in U \subset \mathbb{R}^m. \tag{10.1}$$

Here M is, as usual, a smooth manifold, and U an arbitrary subset of $\mathbb{R}^m$. For the right-hand side of the control system, we suppose that:

$$q \mapsto f_u(q) \text{ is a smooth vector field on } M \text{ for any fixed } u \in U, \tag{10.2}$$

$$(q, u) \mapsto f_u(q) \text{ is a continuous mapping for } q \in M,\ u \in \overline{U}, \tag{10.3}$$

and moreover, in any local coordinates on M

$$(q, u) \mapsto \frac{\partial f_u}{\partial q}(q) \text{ is a continuous mapping for } q \in M,\ u \in \overline{U}. \tag{10.4}$$

Admissible controls are measurable locally bounded mappings

$$u : t \mapsto u(t) \in U.$$

Substitute such a control $u = u(t)$ for control parameter into system (10.1), then we obtain a nonautonomous ODE $\dot{q} = f_u(q)$. By the classical Carathéodory's Theorem, for any point $q_0 \in M$, the Cauchy problem

$$\dot{q} = f_u(q), \qquad q(0) = q_0, \tag{10.5}$$

has a unique solution, see Subsec. 2.4.1. We will often fix the initial point q_0 and then denote the corresponding solution to problem (10.5) as $q_u(t)$.

In order to compare admissible controls one with another on a segment $[0, t_1]$, introduce a *cost functional*:

$$J(u) = \int_0^{t_1} \varphi(q_u(t), u(t))\, dt \tag{10.6}$$

with an integrand

$$\varphi : M \times U \to \mathbb{R}$$

satisfying the same regularity assumptions as the right-hand side f, see (10.2)–(10.4).

Take any pair of points $q_0, q_1 \in M$. We consider the following *optimal control problem*:

MINIMIZE THE FUNCTIONAL J AMONG ALL ADMISSIBLE CONTROLS $u = u(t)$, $t \in [0, t_1]$, FOR WHICH THE CORRESPONDING SOLUTION $q_u(t)$ OF CAUCHY PROBLEM (10.5) SATISFIES THE BOUNDARY CONDITION

$$q_u(t_1) = q_1. \tag{10.7}$$

We study two types of problems, with fixed t_1 and free t_1. A solution u of this problem is called an *optimal control*, and the corresponding curve $q_u(t)$ is the *optimal trajectory*.

So the optimal control problem is the minimization problem for $J(u)$ with constraints on u given by control system and the fixed endpoints conditions (10.5), (10.7). These constraints cannot usually be resolved w.r.t. u, thus solving optimal control problems requires special techniques.

10.2 Reduction to study of attainable sets

Fix an initial point $q_0 \in M$. *Attainable set* of control system (10.1) for time $t \geq 0$ from q_0 with measurable locally bounded controls is defined as follows:

$$\mathcal{A}_{q_0}(t) = \{q_u(t) \mid u \in L^\infty([0,t], U)\}.$$

Similarly, one can consider the attainable sets for time not greater than t:

$$\mathcal{A}_{q_0}^t = \bigcup_{0 \leq \tau \leq t} \mathcal{A}_{q_0}(\tau)$$

and for arbitrary nonnegative time:

$$\mathcal{A}_{q_0} = \bigcup_{0 \leq \tau < \infty} \mathcal{A}_{q_0}(\tau).$$

It turns out that optimal control problems on the state space M can be essentially reduced to the study of attainable sets of some auxiliary control systems on the extended state space

$$\widehat{M} = \mathbb{R} \times M = \{\widehat{q} = (y, q) \mid y \in \mathbb{R},\ q \in M\}.$$

Namely, consider the following extended control system on $\widehat{M}$:

$$\frac{d\widehat{q}}{dt} = \widehat{f}_u(\widehat{q}), \qquad \widehat{q} \in \widehat{M},\ u \in U, \tag{10.8}$$

with the right-hand side

$$\widehat{f}_u(\widehat{q}) = \begin{pmatrix} \varphi(q,u) \\ f_u(q) \end{pmatrix}, \qquad q \in M, \quad u \in U,$$

where φ is the integrand of the cost functional J, see (10.6). Then solutions $\widehat{q}_u(t)$ of the extended system (10.8) with the initial conditions

$$\widehat{q}_u(0) = \begin{pmatrix} y(0) \\ q(0) \end{pmatrix} = \begin{pmatrix} 0 \\ q_0 \end{pmatrix}$$

are expressed through solutions $q_u(t)$ of the original system (10.1) as

$$\widehat{q}_u(t) = \begin{pmatrix} J_t(u) \\ q_u(t) \end{pmatrix},$$

where

$$J_t(u) = \int_0^t \varphi(q_u(\tau), u(\tau))\, d\tau.$$

Thus attainable sets of the extended system (10.8) from the point $(0, q_0)$ have the form

$$\widehat{\mathcal{A}}_{(0,q_0)}(t) = \{(J_t(u), q_u(t)) \mid u \in L^\infty([0,t], U)\}.$$

Let $q(t)$, $t \in [0, t_1]$, be an optimal trajectory for the optimal control problem in M. Consider the corresponding trajectory

$$\widehat{q}(t) = \begin{pmatrix} J_t \\ q(t) \end{pmatrix}, \qquad t \in [0, t_1],$$

of the extended control system in $\widehat{M}$. The endpoint $\widehat{q}(t_1)$ must belong to the boundary of the attainable set $\widehat{\mathcal{A}}_{(0,q_0)}(t_1)$; moreover, this set should not intersect the ray

$$\{(y, q_1) \in \widehat{M} \mid y < J_{t_1}\}.$$

Indeed, if there exist points

$$(y, q_1) \in \widehat{\mathcal{A}}_{(0,q_0)}(t_1), \quad y < J_{t_1},$$

then the trajectory of the extended system

$$\widehat{q'}(t) = \begin{pmatrix} J'_t \\ q'(t) \end{pmatrix}$$

that steers $(0, q_0)$ to (y, q_1):

$$\widehat{q'}(0) = \begin{pmatrix} 0 \\ q_0 \end{pmatrix}, \qquad \widehat{q'}(t_1) = \begin{pmatrix} y \\ q_1 \end{pmatrix},$$

gives a trajectory $q'(t)$, $q'(0) = q_0$, $q'(t_1) = q_1$, with a smaller value of the cost functional:

$$J'_{t_1} = y < J_{t_1},$$

a contradiction with optimality of $q(\cdot)$.

So optimal trajectories (more precisely, their lift to the extended state space $\widehat{M}$) must come to the boundary of the attainable set $\widehat{\mathcal{A}}_{(0,q_0)}(t_1)$. In order to find optimal trajectories, we find those coming to the boundary of $\widehat{\mathcal{A}}_{(0,q_0)}(t_1)$, and then select optimal among them. The first step is much more important than the second one, so solving optimal control problems essentially reduces to the study of dynamics of boundary of attainable sets.

10.3 Compactness of attainable sets

Due to the reduction of optimal control problems to the study of attainable sets, existence of optimal solutions to these problems is reduced to compactness of attainable sets.

For control system (10.1), sufficient conditions for compactness of the attainable sets $\mathcal{A}_{q_0}(t)$ for time t and $\mathcal{A}_{q_0}^t$ for time not greater than t are given in the following proposition.

Theorem 10.1 (Filippov). *Let the space of control parameters $U \Subset \mathbb{R}^m$ be compact. Let there exist a compact $K \Subset M$ such that $f_u(q) = 0$ for $q \notin K$, $u \in U$. Moreover, let the velocity sets*

$$f_U(q) = \{f_u(q) \mid u \in U\} \subset T_qM, \qquad q \in M,$$

be convex. Then the attainable sets $\mathcal{A}_{q_0}(t)$ and $\mathcal{A}_{q_0}^t$ are compact for all $q_0 \in M$, $t > 0$.

Remark. The condition of convexity of the velocity sets $f_U(q)$ is natural in view of Theorem 8.2: the flow of the ODE

$$\dot{q} = \alpha(t)f_{u_1}(q) + (1 - \alpha(t))f_{u_2}(q), \qquad 0 \le \alpha(t) \le 1,$$

can be approximated by flows of the systems of the form

$$\dot{q} = f_v(q), \quad \text{where} \quad v(t) \in \{u_1(t), u_2(t)\}.$$

Now we give a sketch of the proof of Theorem 10.1.

Proof. Notice first of all that all nonautonomous vector fields $f_u(q)$ with admissible controls u have a common compact support, thus are complete. Further, under hypotheses of the theorem, velocities $f_u(q)$, $q \in M$, $u \in U$, are uniformly bounded, thus all trajectories $q(t)$ of control system (10.1) starting at q_0 are Lipschitzian with the same Lipschitz constant. Thus the set of admissible trajectories is precompact in the topology of uniform convergence. (We can embed the manifold M into a Euclidean space $\mathbb{R}^N$, then the space of continuous curves $q(t)$ becomes endowed with the uniform topology of continuous mappings from $[0, t_1]$ to $\mathbb{R}^N$.) For any sequence $q_n(t)$ of admissible trajectories:

$$\dot{q}_n(t) = f_{u_n}(q_n(t)), \qquad 0 \le t \le t_1, \quad q_n(0) = q_0,$$

there exists a uniformly converging subsequence, we denote it again by $q_n(t)$:

$$q_n(\cdot) \to q(\cdot) \text{ in } C[0, t_1] \text{ as } n \to \infty.$$

Now we show that $q(t)$ is an admissible trajectory of control system (10.1).

Fix a sufficiently small $\varepsilon > 0$. Then

$$\begin{aligned}\frac{1}{\varepsilon}(q_n(t+\varepsilon) - q_n(t)) &= \frac{1}{\varepsilon}\int_t^{t+\varepsilon} f_{u_n}(q_n(\tau))\, d\tau \\ &\in \operatorname{conv} \bigcup_{\tau \in [t, t+\varepsilon]} f_U(q_n(\tau)) \subset \operatorname{conv} \bigcup_{q \in O_{q(t)}(c\varepsilon)} f_U(q)\end{aligned}$$

where c is the doubled Lipschitz constant of admissible trajectories. Then we pass to the limit $n \to \infty$ and obtain

$$\frac{1}{\varepsilon}(q(t+\varepsilon) - q(t)) \in \operatorname{conv} \bigcup_{q \in O_{q(t)}(c\varepsilon)} f_U(q).$$

Now let $\varepsilon \to 0$. If t is a point of differentiability of $q(t)$, then

$$\dot{q}(t) \in f_U(q)$$

since $f_U(q)$ is convex.

In order to show that $q(t)$ is an admissible trajectory of control system (10.1), we should find a measurable selection $u(t) \in U$ that generates $q(t)$. We do this via the lexicographic order on the set $U = \{(u_1, \dots, u_m)\} \subset \mathbb{R}^m$.

The set

$$V_t = \{v \in U \mid \dot{q}(t) = f_v(q(t))\}$$

is a compact subset of U, thus of $\mathbb{R}^m$. There exists a vector $v_{\min}(t) \in V_t$ minimal in the sense of lexicographic order: to find $v_{\min}(t)$, we minimize the first coordinate v_1 among all $v = (v_1, \dots, v_m) \in V_t$, then minimize the second coordinate v_2 on the compact set found at the first step, etc. The control $u(t) = v_{\min}(t)$ is measurable, thus $q(t)$ is an admissible solution of control system (10.1).

The proof of compactness of the attainable set $\mathcal{A}_{q_0}(t)$ is complete. Compactness of $\mathcal{A}_{q_0}^t$ is proved by a slightly modified argument. □

Remark. In Filippov's theorem, the hypothesis of common compact support of the vector fields in the right-hand side is essential to ensure the uniform boundedness of velocities and completeness of vector fields. On a manifold, sufficient conditions for completeness of a vector field cannot be given in terms of boundedness of the vector field and its derivatives: a constant vector field is not complete on a bounded domain in $\mathbb{R}^n$. Nevertheless, one can prove compactness of attainable sets for many systems without the assumption of common compact support. If for such a system we have a priori bounds on solutions, then we can multiply its right-hand side by a cut-off function, and obtain a system with vector fields having compact support. We can apply Filippov's theorem to the new system. Since trajectories of the initial and new systems coincide in a domain of interest for us, we obtain a conclusion on compactness of attainable sets for the initial system.

For control systems on $M = \mathbb{R}^n$, there exist well-known sufficient conditions for completeness of vector fields: if the right-hand side grows at infinity not faster than a linear field, i.e.,

$$|f_u(x)| \leq C(1 + |x|), \qquad x \in \mathbb{R}^n, \quad u \in U, \tag{10.9}$$

for some constant C, then the nonautonomous vector fields $f_u(x)$ are complete (here $|x| = \sqrt{x_1^2 + \cdots + x_n^2}$ is the norm of a point $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$).

These conditions provide an a priori bound for solutions: any solution $x(t)$ of the control system

$$\dot{x} = f_u(x), \qquad x \in \mathbb{R}^n, \quad u \in U, \tag{10.10}$$

with the right-hand side satisfying (10.9) admits the bound

$$|x(t)| \leq e^{2Ct} \left(|x(0)| + 1\right), \qquad t \geq 0.$$

So Filippov's theorem plus the previous remark imply the following sufficient condition for compactness of attainable sets for systems in $\mathbb{R}^n$.

Corollary 10.1. *Let system* (10.10) *have a compact space of control parameters* $U \Subset \mathbb{R}^m$ *and convex velocity sets* $f_U(x)$, $x \in \mathbb{R}^n$. *Suppose moreover that the right-hand side of the system satisfies a bound of the form* (10.9). *Then the attainable sets* $\mathcal{A}_{x_0}(t)$ *and* $\mathcal{A}^t_{x_0}$ *are compact for all* $x_0 \in \mathbb{R}^n$, $t > 0$.

10.4 Time-optimal problem

Given a pair of points $q_0 \in M$, $q_1 \in \mathcal{A}_{q_0}$, the *time-optimal problem* consists in minimizing the time of motion from q_0 to q_1 via admissible controls of control system (10.1):

$$\min_u \{t_1 \mid q_u(t_1) = q_1\}. \tag{10.11}$$

That is, we consider the optimal control problem described in Sec. 10.1 with the integrand $\varphi(q, u) \equiv 1$ and free terminal time t_1.

Reduction of optimal control problems to the study of attainable sets and Filippov's Theorem yield the following existence result.

Corollary 10.2. *Under hypotheses of Theorem* 10.1, *time-optimal problem* (10.1), (10.11) *has a solution for any points* $q_0 \in M$, $q_1 \in \mathcal{A}_{q_0}$.

10.5 Relaxations

Consider a control system of the form (10.1) with a compact set of control parameters U. There is a standard procedure called *relaxation* of control system (10.1), which extends the velocity set $f_U(q)$ of this system to its convex hull $\operatorname{conv} f_U(q)$.

Recall that the *convex hull* $\operatorname{conv} S$ of a subset S of a linear space is the minimal convex set that contains S. A constructive description of convex hull is given by the following classical proposition: any point in the convex hull of a set S in the n-dimensional linear space is contained in the convex hull of some $n+1$ points in S.

Lemma 10.1 (Carathéodory). *For any subset $S \subset \mathbb{R}^n$, its convex hull has the form*

$$\operatorname{conv} S = \left\{ \sum_{i=0}^{n} \alpha_i x_i \mid x_i \in S,\ \alpha_i \geq 0,\ \sum_{i=0}^{n} \alpha_i = 1 \right\}.$$

For the proof of this lemma, one can consult e.g. [19].

Relaxation of control system (10.1) is constructed as follows. Let $n = \dim M$ be dimension of the state space. The set of control parameters of the relaxed system is

$$V = \Delta^n \times \underbrace{U \times \cdots \times U}_{n+1 \text{ times}},$$

where

$$\Delta^n = \left\{ (\alpha_0, \ldots, \alpha_n) \mid \alpha_i \geq 0,\ \sum_{i=0}^{n} \alpha_i = 1 \right\} \subset \mathbb{R}^{n+1}$$

is the standard n-dimensional simplex. So the control parameter of the new system has the form

$$v = (\alpha, u_0, \ldots, u_n) \in V, \qquad \alpha = (\alpha_0, \ldots, \alpha_n) \in \Delta^n,\ u_i \in U.$$

If U is compact, then V is compact as well.

The *relaxed system* is

$$\dot{q} = g_v(q) = \sum_{i=0}^{n} \alpha_i f_{u_i}(q), \qquad v = (\alpha, u_0, \ldots, u_n) \in V, \quad q \in M. \tag{10.12}$$

By Carathéodory's lemma, the velocity set $g_V(q)$ of system (10.12) is convex, moreover,

$$g_V(q) = \operatorname{conv} f_U(q).$$

If all vector fields in the right-hand side of (10.12) have a common compact support, we obtain by Filippov's theorem that attainable sets for the relaxed system are compact. By Theorem 8.2, any trajectory of relaxed system (10.12) can be uniformly approximated by families of trajectories of initial system (10.1). Thus attainable sets of the relaxed system coincide with closure of attainable sets of the initial system.

Chapter 11

Elements of Exterior Calculus and Symplectic Geometry

In order to state necessary conditions of optimality for optimal control problems on smooth manifolds — Pontryagin Maximum Principle, see Chapter 12 — we make use of some standard technique of Symplectic Geometry. In this chapter we develop such a technique. Before this we recall some basic facts on calculus of exterior differential forms on manifolds. The exposition in this chapter is rather explanatory than systematic, it is not a substitute to a regular textbook. For a detailed treatment of the subject, see e.g. [24], [4], [6].

11.1 Differential 1-forms

11.1.1 Linear forms

Let E be a real vector space of finite dimension n. The set of linear forms on E, i.e., of linear mappings $\xi : E \to \mathbb{R}$, has a natural structure of a vector space called the *dual space* to E and denoted by E^*. If vectors $e_1, \ldots, e_n$ form a basis of E, then the corresponding *dual basis* of E^* is formed by the covectors $e_1^*, \ldots, e_n^*$ such that

$$\langle e_i^*, e_j \rangle = \delta_{ij}, \qquad i,\ j = 1, \ldots n$$

(we use the angle brackets to denote the value of a linear form $\xi \in E^*$ on a vector $v \in E$: $\langle \xi, v \rangle = \xi(v)$). So the dual space has the same dimension as the initial one:

$$\dim E^* = n = \dim E.$$

11.1.2 Cotangent bundle

Let M be a smooth manifold and T_qM its tangent space at a point $q \in M$. The space of linear forms on T_qM, i.e., the dual space $(T_qM)^*$ to T_qM, is called the *cotangent space* to M at q and is denoted as T_q^*M. The disjoint union of all cotangent spaces is called the *cotangent bundle* of M:

$$T^*M \stackrel{\text{def}}{=} \bigcup_{q \in M} T_q^*M.$$

The set T^*M has a natural structure of a smooth manifold of dimension $2n$, where $n = \dim M$. Local coordinates on T^*M are constructed from local coordinates on M.

Let $O \subset M$ be a coordinate neighborhood and let

$$\Phi : O \to \mathbb{R}^n = \{(x_1, \dots, x_n)\}$$

be a local coordinate system. Let $e_1, \dots, e_n \in \operatorname{Vec} \mathbb{R}^n$ be the standard basis vector fields on $\mathbb{R}^n$:

$$e_i = \begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}$$

with the only identity in the i-th row. Then the pull-back vectors

$$\left.\frac{\partial}{\partial x_i}\right|_q \stackrel{\text{def}}{=} \Phi_*^{-1} e_i \in T_qM, \qquad i = 1, \dots, n, \quad q \in O,$$

form a basis of the tangent space T_qM, $q \in O$.

Recall that if $F : M \to N$ is a smooth mapping between smooth manifolds, then the differential

$$F_* : T_qM \to T_{F(q)}N$$

has the adjoint mapping

$$F^* \stackrel{\text{def}}{=} (F_*)^* : T^*_{F(q)}N \to T_q^*M$$

defined as follows:

$$F^*\xi = \xi \circ F_*, \qquad \xi \in T^*_{F(q)}N,$$
$$\langle F^*\xi, v \rangle = \langle \xi, F_* v \rangle, \qquad v \in T_qM.$$

A vector $v \in T_qM$ is pushed forward by the differential F_* to the vector $F_*v \in T_{F(q)}N$, while a covector $\xi \in T^*_{F(q)}N$ is pulled back to the covector $F^*\xi \in T_q^*M$.

The linear forms

$$dx_i\big|_q \stackrel{\text{def}}{=} \Phi^* e_i^* \in T_q^* M, \qquad i = 1, \dots, n, \quad q \in O,$$

make up the basis in the cotangent spaces $T_q^* M$, $q \in O$, dual to the basis $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$ in the tangent spaces $T_q M$:

$$\left\langle dx_i, \frac{\partial}{\partial x_j} \right\rangle = \delta_{ij}, \qquad i,\ j = 1, \dots, n.$$

Any linear form $\xi \in T_q^* M$ can be decomposed via the basis forms:

$$\xi = \sum_{i=1}^{n} \xi_i \, dx_i.$$

So any covector $\xi \in T^* M$ is characterized by n coordinates $(x_1, \dots, x_n)$ of the point $q \in M$ where ξ is attached, and by n coordinates $(\xi_1, \dots, \xi_n)$ of the linear form ξ in the basis $dx_1, \dots, dx_n$. Mappings of the form

$$\xi \mapsto (\xi_1, \dots, \xi_n;\, x_1, \dots, x_n)$$

define local coordinates on the cotangent bundle. Consequently, $T^* M$ is an $2n$-dimensional manifold. Coordinates of the form (ξ, x) are called *canonical coordinates* on $T^* M$.

11.1.3 Differential 1-forms

A *differential* 1-*form* on M is a smooth mapping

$$q \mapsto \omega_q \in T_q^* M, \qquad q \in M,$$

i.e, a family $\omega = \{\omega_q\}$ of linear forms on the tangent spaces $T_q M$ smoothly depending on the point $q \in M$. The set of all differential 1-forms on M has a natural structure of an infinite-dimensional vector space denoted as $\Lambda^1 M$.

Like linear forms on a vector space are dual objects to vectors of the space, differential forms on a manifold are dual objects to smooth curves in the manifold. The pairing operation is the *integral* of a differential 1-form $\omega \in \Lambda^1(M)$ along a smooth oriented curve $\gamma : [t_0, t_1] \to M$, defined as follows:

$$\int_\gamma \omega \stackrel{\text{def}}{=} \int_{t_0}^{t_1} \langle \omega_{\gamma(t)}, \dot\gamma(t) \rangle \, dt.$$

The integral of a 1-form along a curve does not change under orientation-preserving smooth reparametrizations of the curve and changes its sign under change of orientation.

11.2 Differential k-forms

A differential k-form on M is an object to integrate over k-dimensional surfaces in M. Infinitesimally, a k-dimensional surface is presented by its tangent space, i.e., a k-dimensional subspace in T_qM. We thus need a dual object to the set of k-dimensional subspaces in the linear space. Fix a linear space E. A k-dimensional subspace is defined by its basis $v_1, \ldots, v_k \in E$. The dual objects should be mappings

$$(v_1, \ldots, v_k) \mapsto \omega(v_1, \ldots, v_k) \in \mathbb{R}$$

such that $\omega(v_1, \ldots, v_k)$ depend only on the linear hull $\operatorname{span}\{v_1, \ldots, v_k\}$ and the oriented volume of the k-dimensional parallelepiped generated by $v_1, \ldots, v_k$. Moreover, the dependence on the volume should be linear. Recall that the ratio of volumes of the parallelepipeds generated by vectors $w_i = \sum_{j=1}^{k} \alpha_{ij} v_j$, $i = 1, \ldots, k$, and the vectors $v_1, \ldots, v_k$, equals $\det(\alpha_{ij})_{i,j=1}^{k}$, and that determinant of a $k \times k$ matrix is a multilinear skew-symmetric form of the columns of the matrix. This is why the following definition of the "dual objects" is quite natural.

11.2.1 Exterior k-forms

Let E be a finite-dimensional real vector space, $\dim E = n$, and let $k \in \mathbb{N}$. An *exterior k-form* on E is a mapping

$$\omega \; : \; \underbrace{E \times \cdots \times E}_{k \text{ times}} \to \mathbb{R},$$

which is multilinear:

$$\begin{aligned} &\omega(v_1, \ldots, \alpha_1 v_i^1 + \alpha_2 v_i^2, \ldots, v_k) \\ &\qquad = \alpha_1 \omega(v_1, \ldots, v_i^1, \ldots, v_k) + \alpha_2 \omega(v_1, \ldots, v_i^2, \ldots, v_k), \qquad \alpha_1, \; \alpha_2 \in \mathbb{R}, \end{aligned}$$

and skew-symmetric:

$$\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k), \quad i, \; j = 1, \ldots, k.$$

The set of all exterior k-forms on E is denoted by $\Lambda^k E$. By the skew-symmetry, any exterior form of order $k > n$ is zero, thus $\Lambda^k E = \{0\}$ for $k > n$.

Exterior forms can be multiplied by real numbers, and exterior forms of the same order can be added one with another, so each $\Lambda^k E$ is a vector space. We construct a basis of $\Lambda^k E$ after we consider another operation between exterior forms — the exterior product. The exterior product of two forms $\omega_1 \in \Lambda^{k_1} E$, $\omega_2 \in \Lambda^{k_2} E$ is an exterior form $\omega_1 \wedge \omega_2$ of order $k_1 + k_2$.

Given linear 1-forms $\omega_1, \omega_2 \in \Lambda^1 E$, we have a natural (tensor) product for them:

$$\omega_1 \otimes \omega_2 \; : \; (v_1, v_2) \mapsto \omega_1(v_1)\omega_2(v_2), \qquad v_1, v_2 \in E.$$

The result is a bilinear but not a skew-symmetric form. The *exterior product* is the anti-symmetrization of the tensor one:

$$\omega_1 \wedge \omega_2 \ : \ (v_1, v_2) \mapsto \omega_1(v_1)\omega_2(v_2) - \omega_1(v_2)\omega_2(v_1), \qquad v_1, v_2 \in E.$$

Similarly, the tensor and exterior products of forms $\omega_1 \in \Lambda^{k_1} E$ and $\omega_2 \in \Lambda^{k_2} E$ are the following forms of order $k_1 + k_2$:

$$\begin{aligned}
&\omega_1 \otimes \omega_2 \ : \ (v_1, \dots, v_{k_1+k_2}) \mapsto \omega_1(v_1, \dots, v_{k_1})\omega_2(v_{k_1+1}, \dots, v_{k_1+k_2}), \\
&\omega_1 \wedge \omega_2 \ : \ (v_1, \dots, v_{k_1+k_2}) \mapsto \\
&\quad \frac{1}{k_1!\, k_2!} \sum_{\sigma} (-1)^{\nu(\sigma)} \omega_1(v_{\sigma(1)}, \dots, v_{\sigma(k_1)})\omega_2(v_{\sigma(k_1+1)}, \dots, v_{\sigma(k_1+k_2)}),
\end{aligned} \tag{11.1}$$

where the sum is taken over all permutations σ of order $k_1 + k_2$ and $\nu(\sigma)$ is parity of a permutation σ. The factor $\frac{1}{k_1!\,k_2!}$ normalizes the sum in (11.1) since it contains $k_1!\,k_2!$ identically equal terms: e.g., if permutations σ do not mix the first k_1 and the last k_2 arguments, then all terms of the form

$$(-1)^{\nu(\sigma)} \omega_1(v_{\sigma(1)}, \dots, v_{\sigma(k_1)})\omega_2(v_{\sigma(k_1+1)}, \dots, v_{\sigma(k_1+k_2)})$$

are equal to

$$\omega_1(v_1, \dots, v_{k_1})\omega_2(v_{k_1+1}, \dots, v_{k_1+k_2}).$$

This guarantees the associative property of the exterior product:

$$\omega_1 \wedge (\omega_2 \wedge \omega_3) = (\omega_1 \wedge \omega_2) \wedge \omega_3, \qquad \omega_i \in \Lambda^{k_i} E,$$

Further, the exterior product is skew-commutative:

$$\omega_2 \wedge \omega_1 = (-1)^{k_1 k_2} \omega_1 \wedge \omega_2, \qquad \omega_i \in \Lambda^{k_i} E.$$

Let $e_1, \dots, e_n$ be a basis of the space E and $e_1^*, \dots, e_n^*$ the corresponding dual basis of E^*. If $1 \le k \le n$, then the following $\binom{n}{k}$ elements form a basis of the space $\Lambda^k E$:

$$e_{i_1}^* \wedge \dots \wedge e_{i_k}^*, \qquad 1 \le i_1 < i_2 < \dots < i_k \le n.$$

The equalities

$$\begin{aligned}
&(e_{i_1}^* \wedge \dots \wedge e_{i_k}^*)(e_{i_1}, \dots, e_{i_k}) = 1, \\
&(e_{i_1}^* \wedge \dots \wedge e_{i_k}^*)(e_{j_1}, \dots, e_{j_k}) = 0, \quad \text{if } (i_1, \dots, i_k) \ne (j_1, \dots, j_k)
\end{aligned}$$

for $1 \le i_1 < i_2 < \dots < i_k \le n$ imply that any k-form $\omega \in \Lambda^k E$ has a unique decomposition of the form

$$\omega = \sum_{1 \le i_1 < i_2 < \dots < i_k \le n} \omega_{i_1 \dots i_k} e_{i_1}^* \wedge \dots \wedge e_{i_k}^*$$

with

$$\omega_{i_1 \dots i_k} = \omega(e_{i_1}, \dots, e_{i_k}).$$

Exercise 11.1. Show that for any 1-forms $\omega_1, \ldots \omega_p \in \Lambda^1 E$ and any vectors $v_1, \ldots, v_p \in E$ there holds the equality

$$(\omega_1 \wedge \ldots \wedge \omega_p)(v_1, \ldots, v_p) = \det \left(\langle \omega_i, v_j \rangle\right)_{i,j=1}^p . \tag{11.2}$$

Notice that the space of n-forms of an n-dimensional space E is one-dimensional. Any nonzero n-form on E is a volume form. For example, the value of the standard volume form $e_1^* \wedge \ldots \wedge e_n^*$ on an n-tuple of vectors $(v_1, \ldots, v_n)$ is

$$(e_1^* \wedge \ldots \wedge e_n^*)(v_1, \ldots, v_n) = \det \left(\langle e_i^*, v_j \rangle\right)_{i,j=1}^n ,$$

the oriented volume of the parallelepiped generated by the vectors $v_1, \ldots, v_n$.

11.2.2 Differential k-forms

A *differential k-form* on M is a mapping

$$\omega \; : \; q \mapsto \omega_q \in \Lambda^k T_q M, \qquad q \in M,$$

smooth w.r.t. $q \in M$. The set of all differential k-forms on M is denoted by $\Lambda^k M$. It is natural to consider smooth functions on M as 0-forms, so $\Lambda^0 M = C^\infty(M)$.

In local coordinates $(x_1, \ldots, x_n)$ on a domain $O \subset M$, any differential k-form $\omega \in \Lambda^k M$ can be uniquely decomposed as follows:

$$\omega_x = \sum_{i_1 < \cdots < i_k} a_{i_1 \ldots i_k}(x) dx_{i_1} \wedge \ldots \wedge dx_{i_k}, \qquad x \in O, \quad a_{i_1 \ldots i_k} \in C^\infty(O). \tag{11.3}$$

Any smooth mapping

$$F \; : \; M \to N$$

induces a mapping of differential forms

$$\widehat{F} \; : \; \Lambda^k N \to \Lambda^k M$$

in the following way: given a differential k-form $\omega \in \Lambda^k N$, the k-form $\widehat{F}\omega \in \Lambda^k M$ is defined as

$$(\widehat{F}\omega)_q(v_1, \ldots, v_k) = \omega_{F(q)}(F_* v_1, \ldots, F_* v_k), \qquad q \in M, \; v_i \in T_q M.$$

For 0-forms, pull-back is a substitution of variables:

$$\widehat{F} a = F a, \qquad a \in C^\infty(M).$$

The pull-back $\widehat{F}$ is linear w.r.t. forms and preserves the exterior product:

$$\widehat{F}(\omega_1 \wedge \omega_2) = \widehat{F}\omega_1 \wedge \widehat{F}\omega_2.$$

Exercise 11.2. Prove the composition law for pull-back of differential forms:

$$\widehat{F_1 \circ F_2} = \widehat{F}_1 \circ \widehat{F}_2, \tag{11.4}$$

where $F_1 \; : \; M_1 \to M_2$ and $F_2 \; : \; M_2 \to M_3$ are smooth mappings.

Notice that in our notation (when points are written to the left of mappings as $q \circ F$) the hat $\widehat{\ }$ does not reverse the order of mappings F_1, F_2 in composition, unlike the classical notation $F(q)$.

Now we can define the integral of a k-form over an oriented k-dimensional surface. Let $\Pi \subset \mathbb{R}^k$ be a k-dimensional open oriented domain and

$$\Phi : \Pi \to \Phi(\Pi) \subset M$$

a diffeomorphism. Then the *integral* of a k-form $\omega \in \Lambda^k M$ over the k-dimensional oriented surface $\Phi(\Pi)$ is defined as follows:

$$\int_{\Phi(\Pi)} \omega \stackrel{\text{def}}{=} \int_{\Pi} \widehat{\Phi}\omega,$$

it remains only to define the integral over Π in the right-hand side. Since $\widehat{\Phi}\omega \in \Lambda^k(\mathbb{R}^k)$ is a k-form on $\mathbb{R}^k$, it is expressed via the standard volume form $dx_1 \wedge \ldots \wedge dx_k \in \Lambda^k(\mathbb{R}^k)$:

$$(\widehat{\Phi}\omega)_x = a(x)\, dx_1 \wedge \cdots \wedge dx_k, \qquad x \in \Pi.$$

We set

$$\int_{\Pi} \widehat{\Phi}\omega \stackrel{\text{def}}{=} \int_{\Pi} a(x)\, dx_1 \ldots dx_k,$$

a usual multiple integral.

The integral $\int_{\Phi(\Pi)} \omega$ is defined correctly with respect to orientation-preserving reparametrizations of the surface $\Phi(\Pi)$. Although, if a parametrization changes orientation, then the integral changes sign.

The notion of integral is extended to arbitrary submanifolds as follows. Let $N \subset M$ be a k-dimensional submanifold and let $\omega \in \Lambda^k M$. Consider a covering of N by coordinate neighborhoods $O_i \subset M$:

$$N = \bigcup_i (N \cap O_i).$$

Take a partition of unity subordinated to this covering:

$$\alpha_i \in C^\infty(M), \quad \operatorname{supp} \alpha_i \subset O_i, \quad 0 \le \alpha_i \le 1,$$
$$\sum_i \alpha_i \equiv 1.$$

Then

$$\int_N \omega \stackrel{\text{def}}{=} \sum_i \int_{N \cap O_i} \alpha_i \omega.$$

The integral thus defined does not depend upon the choice of partition of unity.

Remark. Another possible approach to definition of integral of a differential form over a submanifold is based upon triangulation of the submanifold.

11.3 Exterior differential

Exterior differential of a function (i.e., a 0-form) is a 1-form: if $a \in C^\infty(M) = \Lambda^0 M$, then its differential

$$d_q a \in T_q^* M$$

is the functional (directional derivative)

$$\langle d_q a, v\rangle = va, \qquad v \in T_q M, \tag{11.5}$$

so

$$da \in \Lambda^1 M.$$

By the Newton-Leibniz formula, if $\gamma \subset M$ is a smooth oriented curve starting at a point $q_0 \in M$ and terminating at $q_1 \in M$, then

$$\int_\gamma da = a(q_1) - a(q_0).$$

The right-hand side can be considered as the integral of the function a over the oriented boundary of the curve: $\partial\gamma = q_1 - q_0$, thus

$$\int_\gamma da = \int_{\partial\gamma} a. \tag{11.6}$$

In the exposition above, Newton-Leibniz formula (11.6) comes as a consequence of definition (11.5) of differential of a function. But one can go the reverse way: if we postulate Newton-Leibniz formula (11.6) for any smooth curve $\gamma \subset M$ and pass to the limit $q_1 \to q_0$, we necessarily obtain definition (11.5) of differential of a function.

Such approach can be realized for higher order differential forms as well. Let $\omega \in \Lambda^k M$. We define the *exterior differential*

$$d\omega \in \Lambda^{k+1} M$$

as the differential $(k+1)$-form for which Stokes formula holds:

$$\int_N d\omega = \int_{\partial N} \omega \tag{11.7}$$

for $(k+1)$-dimensional submanifolds with boundary $N \subset M$ (for simplicity, one can take here N equal to a diffeomorphic image of a $(k+1)$-dimensional polytope). The boundary ∂N is oriented by a frame of tangent vectors $e_1, \dots e_k \in T_q(\partial N)$ in such a way that the frame $e_n, e_1, \dots, e_k \in T_q N$ define a positive orientation of N, where e_n is the outward normal vector to N at q.

The existence of a form $d\omega$ that satisfies Stokes formula (11.7) comes from the fact that the mapping $N \mapsto \int_{\partial N} \omega$ is additive w.r.t. domain: if $N = N_1 \cup N_2$, $N_1 \cap N_2 = \partial N_1 \cap \partial N_2$, then

$$\int_{\partial N} \omega = \int_{\partial N_1} \omega + \int_{\partial N_2} \omega$$

(notice that orientation of the boundaries is coordinated: ∂N_1 and ∂N_2 have mutually opposite orientations at points of their intersection). Thus the integral $\int_{\partial N} \omega$ is a kind of measure w.r.t. N, and one can recover $(d\omega)_q$ passing to limit in (11.7) as the submanifold N contracts to a point q.

We recall some basic properties of exterior differential. First of all, it is obvious from the Stokes formula that $d : \Lambda^k M \to \Lambda^{k+1} M$ is a linear operator. Further, if $F : M \to N$ is a diffeomorphism, then

$$d\widehat{F}\omega = \widehat{F} d\omega, \qquad \omega \in \Lambda^k N. \tag{11.8}$$

Indeed, if $W \subset M$, then

$$\int_{F(W)} \omega = \int_W \widehat{F}\omega, \qquad \omega \in \Lambda^k N,$$

thus

$$\begin{aligned} \int_W d\widehat{F}\omega &= \int_{\partial W} \widehat{F}\omega = \int_{F(\partial W)} \omega = \int_{\partial F(W)} \omega = \int_{F(W)} d\omega \\ &= \int_W \widehat{F} d\omega, \end{aligned}$$

and equality (11.8) follows.

Another basic property of exterior differential is given by the equality

$$d \circ d = 0,$$

which follows since $\partial(\partial N) = \emptyset$ for any submanifold with boundary $N \subset M$.

Exterior differential is an antiderivation:

$$d(\omega_1 \wedge \omega_2) = (d\omega_1) \wedge \omega_2 + (-1)^{k_1} \omega_1 \wedge d\omega_2, \qquad \omega_i \in \Lambda^{k_i} M.$$

In local coordinates exterior differential is computed as follows: if

$$\omega = \sum_{i_1 < \cdots < i_k} a_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}, \qquad a_{i_1 \dots i_k} \in C^\infty,$$

then

$$d\omega = \sum_{i_1 < \cdots < i_k} (da_{i_1 \dots i_k}) \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}.$$

11.4 Lie derivative of differential forms

The "infinitesimal version" of the pull-back $\widehat{P}$ of a differential form by a flow P is given by the following operation.

Lie derivative of a differential form $\omega \in \Lambda^k M$ along a vector field $f \in \operatorname{Vec} M$ is the differential form $L_f \omega \in \Lambda^k M$ defined as follows:

$$L_f \omega \stackrel{\text{def}}{=} \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \widehat{e^{\varepsilon f}} \omega. \tag{11.9}$$

Since

$$\widehat{e^{tf}}(\omega_1 \wedge \omega_2) = \widehat{e^{tf}}\omega_1 \wedge \widehat{e^{tf}}\omega_2,$$

Lie derivative L_f is a derivation of the algebra of differential forms:

$$L_f(\omega_1 \wedge \omega_2) = (L_f\omega_1) \wedge \omega_2 + \omega_1 \wedge L_f\omega_2.$$

Further, we have

$$\widehat{e^{tf}} \circ d = d \circ \widehat{e^{tf}},$$

thus

$$L_f \circ d = d \circ L_f.$$

For 0-forms, Lie derivative is just the directional derivative:

$$L_f a = fa, \qquad a \in C^\infty(M),$$

since

$$\widehat{e^{tf}} a = e^{tf}\, a$$

as substitution of variables.

Now we obtain a useful formula for the action of Lie derivative on differential forms of an arbitrary order.

Consider, along with exterior differential

$$d : \Lambda^k M \to \Lambda^{k+1} M$$

the *interior product* of a differential form ω with a vector field $f \in \operatorname{Vec} M$:

$$\begin{aligned} &i_f : \Lambda^k M \to \Lambda^{k-1} M, \\ &(i_f\omega)(v_1, \dots, v_{k-1}) \stackrel{\text{def}}{=} \omega(f, v_1, \dots, v_{k-1}), \qquad \omega \in \Lambda^k M,\ v_i \in T_q M, \end{aligned}$$

which acts as substitution of f for the first argument of ω. By definition, for 0-order forms

$$i_f a = 0, \qquad a \in \Lambda^0 M.$$

Interior product is an antiderivation, as well as the exterior differential:

$$i_f(\omega_1 \wedge \omega_2) = (i_f\omega_1) \wedge \omega_2 + (-1)^{k_1}\omega_1 \wedge i_f\omega_2, \qquad \omega_i \in \Lambda^{k_i} M.$$

Now we prove that Lie derivative of a differential form of an arbitrary order can be computed by the following formula:

$$L_f = d \circ i_f + i_f \circ d \tag{11.10}$$

called *Cartan's formula*, for short "$L = di + id$". Notice first of all that the right-hand side in (11.10) has the required order:

$$d \circ i_f + i_f \circ d : \Lambda^k M \to \Lambda^k M.$$

Further, $d \circ i_f + i_f \circ d$ is a derivation as it is obtained from two antiderivations. Moreover, this derivation commutes with differential:

$$d \circ (d \circ i_f + i_f \circ d) = d \circ i_f \circ d,$$
$$(d \circ i_f + i_f \circ d) \circ d = d \circ i_f \circ d.$$

Now we check formula (11.10) on 0-forms: if $a \in \Lambda^0 M$, then

$$(d \circ i_f)a = 0,$$
$$(i_f \circ d)a = \langle da, f \rangle = fa = L_f a.$$

So equality (11.10) holds for 0-forms. The properties of the mappings L_f and $d \circ i_f + i_f \circ d$ established and the coordinate representation (11.3) reduce the general case of k-forms, $k > 0$, to the case of 0-forms. Formula (11.10) is proved.

The differential definition (11.9) of Lie derivative can be integrated, i.e., there holds the following equality on $\Lambda^k M$:

$$\widehat{\left(\overrightarrow{\exp} \int_0^t f_\tau \, d\tau \right)} = \overrightarrow{\exp} \int_0^t L_{f_\tau} \, d\tau, \tag{11.11}$$

in the following sense. Denote the flow

$$P_{t_0}^{t_1} = \overrightarrow{\exp} \int_{t_0}^{t_1} f_\tau \, d\tau.$$

The family of operators on differential forms

$$\widehat{P_0^t} : \Lambda^k M \to \Lambda^k M$$

is a unique solution of the Cauchy problem

$$\frac{d}{dt} \widehat{P_0^t} = \widehat{P_0^t} \circ L_{f_t}, \qquad \left. \widehat{P_0^t} \right|_{t=0} = \mathrm{Id}, \tag{11.12}$$

compare with Cauchy problems for the flow P_0^t (2.9) and for the family of operators $\operatorname{Ad} P_0^t$ (2.21), (2.22), and this solution is denoted as

$$\overrightarrow{\exp} \int_0^t L_{f_\tau} \, d\tau \stackrel{\text{def}}{=} \widehat{\left(\overrightarrow{\exp} \int_0^t f_\tau \, d\tau \right)}.$$

In order to verify the ODE in (11.12), we prove first the following equality for operators on forms:

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \widehat{P_t^{t+\varepsilon}} \omega = L_{f_t} \omega, \qquad \omega \in \Lambda^k M. \tag{11.13}$$

This equality is straightforward for 0-order forms:

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \widehat{P_t^{t+\varepsilon}} a = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} P_t^{t+\varepsilon} a = f_t a = L_{f_t} a, \qquad a \in C^\infty(M).$$

Further, the both operators $\frac{d}{d\varepsilon}\big|_{\varepsilon=0} \widehat{P_t^{t+\varepsilon}}$ and L_{f_t} commute with d and satisfy the Leibniz rule w.r.t. product of a function with a differential form. Then equality (11.13) follows for forms of arbitrary order, as in the proof of Cartan's formula.

Now we easily verify the ODE in (11.12):

$$\frac{d}{dt}\widehat{P_0^t} = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \widehat{P_0^{t+\varepsilon}} = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \widehat{P_0^t \circ P_t^{t+\varepsilon}}$$

by the composition rule (11.4)

$$\begin{aligned} &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \widehat{P_0^t} \circ \widehat{P_t^{t+\varepsilon}} = \widehat{P_0^t} \circ \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \widehat{P_t^{t+\varepsilon}} \\ &= \widehat{P_0^t} \circ L_{f_t}. \end{aligned}$$

Exercise 11.3. Prove uniqueness for Cauchy problem (11.12).

For an autonomous vector field $f \in \operatorname{Vec} M$, equality (11.11) takes the form

$$\widehat{e^{tf}} = e^{tL_f}.$$

Notice that the Lie derivatives of differential forms L_f and vector fields $(-\operatorname{ad} f)$ are in a certain sense dual one to another, see equality (11.14) below. That is, the function

$$\langle \omega, X \rangle \;:\; q \mapsto \langle \omega_q, X(q) \rangle, \qquad q \in M,$$

defines a pairing of $\Lambda^1(M)$ and $\operatorname{Vec}(M)$ over $C^\infty(M)$. Then the equality

$$\langle \widehat{P}\omega, X \rangle = P\langle \omega, P_* X \rangle, \qquad P \in \operatorname{Diff} M,\ X \in \operatorname{Vec} M,\ \omega \in \Lambda^1(M),$$

has an infinitesimal version of the form

$$\langle L_Y \omega, X \rangle = Y\langle \omega, X \rangle - \langle \omega, (\operatorname{ad} Y)X \rangle, \qquad X,\ Y \in \operatorname{Vec} M, \quad \omega \in \Lambda^1(M). \tag{11.14}$$

Taking into account Cartan's formula, we immediately obtain the following important equality:

$$d\omega(Y, X) = Y\langle \omega, X \rangle - X\langle \omega, Y \rangle - \langle \omega, [Y, X] \rangle, \qquad X,\ Y \in \operatorname{Vec} M, \quad \omega \in \Lambda^1(M). \tag{11.15}$$

11.5 Elements of Symplectic Geometry

We have already seen that the cotangent bundle $T^*M = \cup_{q \in M} T_q^* M$ of an n-dimensional manifold M is a $2n$-dimensional manifold. Any local coordinates $x = (x_1, \dots, x_n)$ on M determine canonical local coordinates on T^*M of the form $(\xi, x) = (\xi_1, \dots, \xi_n; x_1, \dots, x_n)$ in which any covector $\lambda \in T_{q_0}^* M$ has the decomposition $\lambda = \sum_{i=1}^n \xi_i \, dx_i\big|_{q_0}$.

11.5.1 Liouville form and symplectic form

The *"tautological" 1-form* (or *Liouville 1-form*) on the cotangent bundle

$$s \in \Lambda^1(T^*M)$$

is defined as follows. Let $\lambda \in T^*M$ be a point in the cotangent bundle and $w \in T_\lambda(T^*M)$ a tangent vector to T^*M at λ. Denote by π the canonical projection from T^*M to M:

$$\begin{aligned} \pi \ &: \ T^*M \to M, \\ \pi \ &: \ \lambda \mapsto q, \qquad \lambda \in T_q^*M. \end{aligned}$$

Differential of π is a linear mapping

$$\pi_* \ : \ T_\lambda(T^*M) \to T_qM, \qquad q = \pi(\lambda).$$

The tautological 1-form s at the point λ acts on the tangent vector w in the following way:

$$\langle s_\lambda, w \rangle \ \stackrel{\mathrm{def}}{=} \ \langle \lambda, \pi_* w \rangle.$$

That is, we project the vector $w \in T_\lambda(T^*M)$ to the vector $\pi_* w \in T_qM$, and then act by the covector $\lambda \in T_q^*M$. So

$$s_\lambda \ \stackrel{\mathrm{def}}{=} \ \lambda \circ \pi_*.$$

The title "tautological" is explained by the coordinate representation of the form s. In canonical coordinates (ξ, x) on T^*M, we have:

$$\begin{aligned} \lambda &= \sum_{i=1}^n \xi_i \, dx_i, \qquad (11.16) \\ w &= \sum_{i=1}^n \alpha_i \frac{\partial}{\partial \xi_i} + \beta_i \frac{\partial}{\partial x_i}. \end{aligned}$$

The projection written in canonical coordinates

$$\pi \ : \ (\xi, x) \mapsto x$$

is a linear mapping, its differential acts as follows:

$$\begin{aligned} \pi_* \left(\frac{\partial}{\partial \xi_i} \right) &= 0, \qquad i = 1, \dots, n, \\ \pi_* \left(\frac{\partial}{\partial x_i} \right) &= \frac{\partial}{\partial x_i}, \qquad i = 1, \dots, n. \end{aligned}$$

Thus

$$\pi_* w = \sum_{i=1}^n \beta_i \frac{\partial}{\partial x_i},$$

consequently,

$$\langle s_\lambda, w \rangle = \langle \lambda, \pi_* w \rangle = \sum_{i=1}^{n} \xi_i \beta_i.$$

But $\beta_i = \langle dx_i, w \rangle$, so the form s has in coordinates (ξ, x) exactly the same expression

$$s_\lambda = \sum_{i=1}^{n} \xi_i \, dx_i \tag{11.17}$$

as the covector λ, see (11.16). Although, definition of the form s does not depend on any coordinates.

Remark. In mechanics, the tautological form s is denoted as $p\,dq$.

Consider the exterior differential of the 1-form s:

$$\sigma \stackrel{\text{def}}{=} ds.$$

The differential 2-form $\sigma \in \Lambda^2(T^*M)$ is called the *canonical symplectic structure* on T^*M. In canonical coordinates, we obtain from (11.17):

$$\sigma = \sum_{i=1}^{n} d\xi_i \wedge dx_i. \tag{11.18}$$

This expression shows that the form σ is nondegenerate, i.e., the bilinear skew-symmetric form

$$\sigma_\lambda : T_\lambda(T^*M) \times T_\lambda(T^*M) \to \mathbb{R}$$

has no kernel:

$$\sigma(w, \cdot) = 0 \quad \Rightarrow \quad w = 0, \qquad w \in T_\lambda(T^*M).$$

In the following basis in the tangent space $T_\lambda(T^*M)$

$$\frac{\partial}{\partial x_1}, \frac{\partial}{\partial \xi_1}, \quad \dots, \quad \frac{\partial}{\partial x_n}, \frac{\partial}{\partial \xi_n},$$

the form σ_λ has the block matrix

$$\begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}.$$

The form σ is closed:

$$d\sigma = 0$$

since it is exact: $\sigma = ds$, and $d \circ d = 0$.

Remarks. (1) A closed nondegenerate exterior differential 2-form on a $2n$-dimensional manifold is called a *symplectic structure*. A manifold with a symplectic structure is called a *symplectic manifold*. The cotangent bundle T^*M with the canonical symplectic structure σ is the most important example of a symplectic manifold.

(2) In mechanics, the 2-form σ is known as the form $dp \wedge dq$.

11.5.2 Hamiltonian vector fields

Due to the symplectic structure $\sigma \in \Lambda^2(T^*M)$, we can develop the Hamiltonian formalism on T^*M. A *Hamiltonian* is an arbitrary smooth function on the cotangent bundle:

$$h \in C^\infty(T^*M).$$

To any Hamiltonian h, we associate the *Hamiltonian vector field*

$$\vec{h} \in \operatorname{Vec}(T^*M)$$

by the rule:

$$\sigma_\lambda(\cdot, \vec{h}) = d_\lambda h, \qquad \lambda \in T^*M. \tag{11.19}$$

In terms of the interior product $i_v \omega(\cdot, \cdot) = \omega(v, \cdot)$, the Hamiltonian vector field is a vector field $\vec{h}$ that satisfies

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

Since the symplectic form σ is nondegenerate, the mapping

$$w \mapsto \sigma_\lambda(\cdot, w)$$

is a linear isomorphism

$$T_\lambda(T^*M) \to T_\lambda^*(T^*M),$$

thus the Hamiltonian vector field $\vec{h}$ in (11.19) exists and is uniquely determined by the Hamiltonian function h.

In canonical coordinates (ξ, x) on T^*M we have

$$dh = \sum_{i=1}^n \left(\frac{\partial h}{\partial \xi_i} d\xi_i + \frac{\partial h}{\partial x_i} dx_i \right),$$

then in view of (11.18)

$$\vec{h} = \sum_{i=1}^n \left(\frac{\partial h}{\partial \xi_i} \frac{\partial}{\partial x_i} - \frac{\partial h}{\partial x_i} \frac{\partial}{\partial \xi_i} \right), \tag{11.20}$$

So the *Hamiltonian system* of ODEs corresponding to h

$$\dot{\lambda} = \vec{h}(\lambda), \qquad \lambda \in T^*M,$$

reads in canonical coordinates as follows:

$$\begin{cases} \dot{x}_i = \dfrac{\partial h}{\partial \xi_i}, & i = 1, \dots, n, \\ \dot{\xi}_i = -\dfrac{\partial h}{\partial x_i}, & i = 1, \dots, n. \end{cases}$$

The Hamiltonian function can depend on a parameter: h_t, $t \in \mathbb{R}$. Then the nonautonomous Hamiltonian vector field $\vec{h}_t$, $t \in \mathbb{R}$ is defined in the same way as in the autonomous case.

The flow of a Hamiltonian system preserves the symplectic form σ.

Proposition 11.1. *Let $\vec{h}_t$ ba a nonautonomous Hamiltonian vector field on T^*M. Then*

$$\left(\overrightarrow{\exp} \int_0^t \vec{h}_\tau \, d\tau\right)^{\widehat{}} \sigma = \sigma.$$

Proof. In view of equality (11.11), we have

$$\left(\overrightarrow{\exp} \int_0^t \vec{h}_\tau \, d\tau\right)^{\widehat{}} = \overrightarrow{\exp} \int_0^t L_{\vec{h}_\tau} \, d\tau,$$

thus the statement of this proposition can be rewritten as

$$L_{\vec{h}_t} \sigma = 0.$$

But this Lie derivative is easily computed by Cartan's formula:

$$L_{\vec{h}_t} \sigma = i_{\vec{h}_t} \circ \underbrace{d\sigma}_{=0} + d \circ \underbrace{i_{\vec{h}_t} \sigma}_{=-dh_t} = -d \circ dh_t = 0.$$

□

Moreover, there holds a local converse statement: if a flow preserves σ, then it is locally Hamiltonian. Indeed,

$$\left(\overrightarrow{\exp} \int_0^t f_\tau \, d\tau\right)^{\widehat{}} \sigma = \sigma \quad \Leftrightarrow \quad L_{f_t} \sigma = 0,$$

further

$$L_{f_t} \sigma = i_{f_t} \circ \underbrace{d\sigma}_{=0} + d \circ i_{f_t} \sigma,$$

thus

$$L_{f_t} \sigma = 0 \quad \Leftrightarrow \quad d \circ i_{f_t} \sigma = 0.$$

If the form $i_{f_t} \sigma$ is closed, then it is locally exact (Poincaré's Lemma), i.e., there exists a Hamiltonian h_t such that locally $f_t = \vec{h}_t$.

Essentially, only Hamiltonian flows preserve σ (globally, "multi-valued Hamiltonians" can appear). If a manifold M is simply connected, then there holds a global statement: a flow on T^*M is Hamiltonian if and only if it preserves the symplectic structure.

The *Poisson bracket* of Hamiltonians $a, b \in C^\infty(T^*M)$ is a Hamiltonian

$$\{a, b\} \in C^\infty(T^*M)$$

defined in one of the following equivalent ways:

$$\{a, b\} = \vec{a} b = \langle db, \vec{a} \rangle = \sigma(\vec{a}, \vec{b}) = -\sigma(\vec{b}, \vec{a}) = -\vec{b} a.$$

It is obvious that Poisson bracket is bilinear and skew-symmetric:

$$\{a, b\} = -\{b, a\}.$$

In canonical coordinates (ξ, x) on T^*M,

$$\{a, b\} = \sum_{i=1}^{n} \left(\frac{\partial a}{\partial \xi_i} \frac{\partial b}{\partial x_i} - \frac{\partial a}{\partial x_i} \frac{\partial b}{\partial \xi_i} \right). \tag{11.21}$$

Leibniz rule for Poisson bracket easily follows from definition:

$$\{a, bc\} = \{a, b\}c + b\{a, c\}$$

(here bc is the usual pointwise product of functions b and c).

Moreover, there holds Jacobi identity for Poisson bracket.

Proposition 11.2.

$$\{a, \{b, c\}\} + \{b, \{c, a\}\} + \{c, \{a, b\}\} = 0, \qquad a, b, c \in C^\infty(T^*M). \tag{11.22}$$

Proof. It is easy to see from the coordinate representation (11.21) that each iterated bracket $\{a, \{b, c\}\}$, $\{b, \{c, a\}\}$, $\{c, \{a, b\}\}$ is a sum of products of second- and first-order derivatives of the functions a, b, c. Now we compute all terms in the left-hand side of (11.22) that contain second-order derivatives of a. The bracket $\{a, \{b, c\}\}$ contains only first-order derivatives of a, so remain the last two brackets:

$$\{b, \{c, a\}\} + \{c, \{a, b\}\} = \{b, \{c, a\}\} - \{c, \{b, a\}\} = \vec{b} \circ \vec{c}\, a - \vec{c} \circ \vec{b}\, a = [\vec{b}, \vec{c}]\, a.$$

The Lie bracket $[\vec{b}, \vec{c}]$ is a first-order differential operator, thus $[\vec{b}, \vec{c}]a$ does not contain second-order derivatives of a. So the sum in the left-hand side of (11.22) contains no second-order derivatives of a. Similarly, it contains no second-order derivatives of b and c. Jacobi identity (11.22) follows. □

So the space of all Hamiltonians $C^\infty(T^*M)$ forms a Lie algebra with Poisson bracket as a product. The correspondence

$$a \mapsto \vec{a}, \qquad a \in C^\infty(T^*M), \tag{11.23}$$

is a homomorphism from the Lie algebra of Hamiltonians to the Lie algebra of Hamiltonian vector fields on M. This follows from the next statement.

Corollary 11.1. $\overrightarrow{\{a,b\}} = [\vec{a}, \vec{b}]$ *for any Hamiltonians* $a, b \in C^\infty(T^*M)$.

Proof. Jacobi identity can be rewritten as

$$\{\{a,b\},c\} = \{a,\{b,c\}\} - \{b,\{a,c\}\},$$

i.e.,

$$\overrightarrow{\{a,b\}}\ c = \vec{a} \circ \vec{b}\, c - \vec{b} \circ \vec{a}\, c = [\vec{a}, \vec{b}]\, c, \qquad c \in C^\infty(T^*M).$$

□

Symplectomorphisms of cotangent bundle preserve the Lie algebra of Hamiltonian vector fields; the action of a symplectomorphism $P \in \mathrm{Diff}(T^*M)$ on a Hamiltonian vector field $\vec{h}$ reduces to the action of P on the Hamiltonian function as substitution of variables:

$$(\mathrm{Ad}\, P)\, \vec{h} = \overrightarrow{Ph}\ .$$

This follows from the chain

$$\begin{aligned}\sigma\left(X, (\mathrm{Ad}\, P)\vec{h}\right) &= \sigma\left(X, P_*^{-1}\vec{h}\right) = (\widehat{P}\sigma)\left(X, P_*^{-1}\vec{h}\right) = P\sigma\left(P_* X, \vec{h}\right)\\ &= P\langle dh, P_* X\rangle = X(Ph), \qquad X \in \mathrm{Vec}(T^*M).\end{aligned}$$

In particular, the action of a Hamiltonian flow on a Hamiltonian vector field gives another Hamiltonian vector field

$$\left(\overrightarrow{\exp} \int_0^t \vec{a}_\tau \, d\tau\right)_*^{-1} \vec{b}_\tau = \left(\overrightarrow{\exp} \int_0^t \mathrm{ad}\, \vec{a}_\tau \, d\tau\right) \vec{b}_\tau = \vec{c}_t \tag{11.24}$$

with the Hamiltonian function

$$c_t = \left(\overrightarrow{\exp} \int_0^t \vec{a}_\tau \, d\tau\right) b_t.$$

Corollary 11.1 can be viewed as an infinitesimal version of equality (11.24).

It is easy to see from the coordinate representation (11.20) that the kernel of the mapping $a \mapsto \vec{a}$ consists of constant functions, i.e., this is isomorphism up to constants. On the other hand, this homomorphism is far from being onto all vector fields on T^*M. Indeed, a general vector field on T^*M is locally defined by arbitrary $2n$ smooth real functions of $2n$ variables, while a Hamiltonian vector field is determined by just one real function of $2n$ variables, a Hamiltonian.

Theorem 11.1 (Nöther). *A function* $a \in C^\infty(T^*M)$ *is an integral of a Hamiltonian system of ODEs*

$$\dot{\lambda} = \vec{h}(\lambda), \qquad \lambda \in T^*M, \tag{11.25}$$

i.e.,

$$e^{t\vec{h}} a = a \qquad t \in \mathbb{R},$$

if and only if it Poisson-commutes with the Hamiltonian:

$$\{a, h\} = 0.$$

Proof. $e^{t\vec{h}} a \equiv a \Leftrightarrow 0 = \vec{h} a = \{h, a\}$. □

Corollary 11.2. $e^{t\vec{h}} h = h$, *i.e., any Hamiltonian* $h \in C^\infty(T^*M)$ *is an integral of the corresponding Hamiltonian system* (11.25).

Further, Jacobi identity for Poisson brackets implies that the set of integrals of the Hamiltonian system (11.25) forms a Lie algebra with respect to Poisson brackets.

Corollary 11.3. $\{h, a\} = \{h, b\} = 0 \Rightarrow \{h, \{a, b\}\} = 0$.

Remark. The Hamiltonian formalism developed generalizes for arbitrary symplectic manifolds.

Now we introduce a construction that works only on T^*M. Given a vector field $X \in \operatorname{Vec} M$, we define a Hamiltonian function

$$X^* \in C^\infty(T^*M),$$

which is linear on fibers T_q^*M, as follows:

$$X^*(\lambda) = \langle \lambda, X(q) \rangle, \qquad \lambda \in T^*M, \quad q = \pi(\lambda).$$

In canonical coordinates (ξ, x) on T^*M we have:

$$X = \sum_{i=1}^n a_i(x) \frac{\partial}{\partial x_i},$$

$$X^* = \sum_{i=1}^n \xi_i a_i(x). \tag{11.26}$$

This coordinate representation implies that

$$\{X^*, Y^*\} = [X, Y]^*, \qquad X, Y \in \operatorname{Vec} M,$$

i.e., Poisson brackets of Hamiltonians linear on fibers in T^*M contain usual Lie brackets of vector fields on M.

The Hamiltonian vector field $\overrightarrow{X^*} \in \operatorname{Vec}(T^*M)$ corresponding to the Hamiltonian function X^* is called the *Hamiltonian lift* of the vector field $X \in \operatorname{Vec} M$. It is easy to see from the coordinate representations (11.26), (11.20) that

$$\pi_* \left(\overrightarrow{X^*} \right) = X.$$

Now we pass to nonautonomous vector fields. Let X_t be a nonautonomous vector field and

$$P_{\tau,t} = \overrightarrow{\exp} \int_\tau^t X_\theta \, d\theta$$

the corresponding flow on M. The flow $P = P_{\tau,t}$ acts on M:

$$P : M \to M, \qquad P : q_0 \mapsto q_1,$$

its diferential pushes tangent vectors forward:

$$P_* : T_{q_0}M \to T_{q_1}M,$$

and the dual mapping P^* pulls covectors back:

$$P^* : T^*_{q_1}M \to T^*_{q_0}M.$$

Thus we have a flow on covectors (i.e., on points of the cotangent bundle):

$$P^*_{\tau,t} : T^*M \to T^*M.$$

Let V_t be the nonautonomous vector field on T^*M that generates the flow $P^*_{\tau,t}$:

$$V_t = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} P^*_{t,t+\varepsilon}.$$

Then

$$\frac{d}{dt}P^*_{\tau,t} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} P^*_{\tau,t+\varepsilon} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} P^*_{t,t+\varepsilon} \circ P^*_{\tau,t} = V_t \circ P^*_{\tau,t},$$

so the flow $P^*_{\tau,t}$ is a solution to the Cauchy problem

$$\frac{d}{dt}P^*_{\tau,t} = V_t \circ P^*_{\tau,t}, \qquad P^*_{\tau,\tau} = \mathrm{Id},$$

i.e., it is the left chronological exponential:

$$P^*_{\tau,t} = \overleftarrow{\exp} \int_\tau^t V_\theta \, d\theta.$$

It turns out that the nonautonomous field V_t is simply related with the Hamiltonian vector field corresponding to the Hamiltonian X^*_t:

$$V_t = -\overrightarrow{X^*_t}. \tag{11.27}$$

Indeed, the flow $P^*_{\tau,t}$ preserves the tautological form s, thus

$$L_{V_t} s = 0.$$

By Cartan's formula,

$$i_{V_t}\sigma = -d\langle s, V_t\rangle,$$

i.e., the field V_t is Hamiltonian:

$$V_t = \overrightarrow{\langle s, V_t\rangle}.$$

But $\pi_* V_t = -X_t$, consequently,

$$\langle s, V_t\rangle = -X^*_t,$$

and equality (11.27) follows. Taking into account relation (2.18) between the left and right chronological exponentials, we obtain

$$P^*_{\tau,t} = \overleftarrow{\exp} \int_\tau^t -\overrightarrow{X^*_\theta} \, d\theta = \overrightarrow{\exp} \int_t^\tau \overrightarrow{X^*_\theta} \, d\theta.$$

We proved the following statement.

Proposition 11.3. *Let X_t be a complete nonautonomous vector field on M. Then*

$$\left(\overrightarrow{\exp}\int_\tau^t X_\theta\, d\theta\right)^* = \overrightarrow{\exp}\int_t^\tau \overrightarrow{X_\theta^*}\, d\theta.$$

In particular, for autonomous vector fields $X \in \operatorname{Vec} M$,

$$\left(e^{tX}\right)^* = e^{-t\overrightarrow{X^*}}.$$

11.5.3 Lagrangian subspaces

A linear space Σ endowed with a bilinear skew-symmetric nondegenerate form σ is called a *symplectic space*. For example, $\Sigma = T_\lambda(T^*M)$ with the canonical symplectic form $\sigma = \sigma_\lambda$ is a symplectic space.

Any subspace L of a symplectic space Σ has the skew-orthogonal complement

$$L^\angle = \{x \in \Sigma \mid \sigma(x, L) = 0\}.$$

A subspace $L \subset \Sigma$ is called *isotropic* if

$$L \subset L^\angle.$$

Since the symplectic form σ is nondegenerate, then

$$\dim L^\angle = \operatorname{codim} L.$$

In particular, if a subspace L is isotropic, then $\dim L \le \frac{1}{2}\dim\Sigma$. Isotropic subspaces of maximal dimension:

$$L \subset L^\angle,\ \dim L = \frac{1}{2}\dim\Sigma \quad \Leftrightarrow \quad L = L^\angle,$$

are called *Lagrangian subspaces*.

For example, in canonical coordinates (p, q) on Σ, the vertical subspace $\{q = 0\}$ and the horizontal subspace $\{p = 0\}$ are Lagrangian.

There exists a standard way to construct a Lagrangian subspace that contains any given isotropic subspace. Let $\Gamma \subset \Sigma$ be an isotropic subspace and $\Lambda \subset \Sigma$ a Lagrangian subspace. Then the subspace

$$\Lambda^\Gamma \stackrel{\text{def}}{=} \Lambda \cap \Gamma^\angle + \Gamma = (\Lambda + \Gamma) \cap \Gamma^\angle \tag{11.28}$$

is Lagrangian (check!). It is clear that

$$\Lambda^\Gamma \supset \Gamma.$$

In particular, any line in Σ is contained in some Lagrangian subspace.

Chapter 12

Pontryagin Maximum Principle

In this chapter we prove the fundamental necessary condition of optimality for optimal control problems — Pontryagin Maximum Principle (PMP). In order to obtain a coordinate-free formulation of PMP on manifolds, we apply the technique of Symplectic Geometry developed in the previous chapter. The first classical version of PMP was proved for optimal control problems in $\mathbb{R}^n$ by L. S. Pontryagin and his collaborators [16].

12.1 Geometric statement of PMP and discussion

Consider the optimal control problem stated in Sec. 10.1 for a control system

$$\dot{q} = f_u(q), \qquad q \in M, \quad u \in U \subset \mathbb{R}^m, \tag{12.1}$$

with the initial condition

$$q(0) = q_0. \tag{12.2}$$

Define the following family of Hamiltonians:

$$h_u(\lambda) = \langle \lambda, f_u(q) \rangle, \qquad \lambda \in T_q^* M, \ q \in M, \ u \in U.$$

In terms of the previous section,

$$h_u(\lambda) = f_u^*(\lambda).$$

Fix an arbitrary instant $t_1 > 0$.

In Sec. 10.2 we reduced the optimal control problem to the study of boundary of attainable sets. Now we give a *necessary optimality condition* in this geometric setting.

Theorem 12.1 (PMP). *Let $\tilde{u}(t)$, $t \in [0, t_1]$, be an admissible control and $\tilde{q}(t) = q_{\tilde{u}}(t)$ the corresponding solution of* (12.1), (12.2)*. If*

$$\tilde{q}(t_1) \in \partial \mathcal{A}_{q_0}(t_1),$$

then there exists a Lipschitzian curve in the cotangent bundle

$$\lambda_t \in T^*_{\tilde{q}(t)} M, \qquad 0 \le t \le t_1,$$

such that

$$\lambda_t \ne 0, \tag{12.3}$$

$$\dot{\lambda}_t = \vec{h}_{\tilde{u}(t)}(\lambda_t), \tag{12.4}$$

$$h_{\tilde{u}(t)}(\lambda_t) = \max_{u \in U} h_u(\lambda_t) \tag{12.5}$$

for almost all $t \in [0, t_1]$.

If $u(t)$ is an admissible control and λ_t a Lipschitzian curve in T^*M such that conditions (12.3)–(12.5) hold, then the pair $(u(t), \lambda_t)$ is said to satisfy PMP. In this case the curve λ_t is called an *extremal*, and its projection $\tilde{q}(t) = \pi(\lambda_t)$ is called an *extremal trajectory*.

Remark. If a pair $(\tilde{u}(t), \lambda_t)$ satisfies PMP, then

$$h_{\tilde{u}(t)}(\lambda_t) = \text{const}, \qquad t \in [0, t_1]. \tag{12.6}$$

Indeed, since the admissible control $\tilde{u}(t)$ is bounded, we can take maximum in (12.5) over the compact $\overline{\{\tilde{u}(t) \mid t \in [0, t_1]\}} = \widetilde{U}$. Further, the function

$$\varphi(\lambda) = \max_{u \in \widetilde{U}} h_u(\lambda)$$

is Lipschitzian w.r.t. $\lambda \in T^*M$. We show that this function has zero derivative. For any admissible control $u(t)$,

$$\varphi(\lambda_t) \ge h_{u(\tau)}(\lambda_t), \qquad \varphi(\lambda_\tau) = h_{u(\tau)}(\lambda_\tau),$$

thus

$$\frac{\varphi(\lambda_t) - \varphi(\lambda_\tau)}{t - \tau} \ge \frac{h_{u(\tau)}(\lambda_t) - h_{u(\tau)}(\lambda_\tau)}{t - \tau}, \qquad t > \tau.$$

Consequently,

$$\left.\frac{d}{dt}\right|_{t=\tau} \varphi(\lambda_t) \ge \{h_{u(\tau)}, h_{u(\tau)}\} = 0$$

if τ is a differentiability point of $\varphi(\lambda_t)$. Similarly,

$$\frac{\varphi(\lambda_t) - \varphi(\lambda_\tau)}{t - \tau} \le \frac{h_{u(\tau)}(\lambda_t) - h_{u(\tau)}(\lambda_\tau)}{t - \tau}, \qquad t < \tau,$$

thus

$$\left.\frac{d}{dt}\right|_{t=\tau} \varphi(\lambda_t) \le 0.$$

So
$$\frac{d}{dt}\varphi(\lambda_t) = 0,$$
and identity (12.6) follows.

The Hamiltonian system of PMP
$$\dot{\lambda}_t = \vec{h}_{u(t)}(\lambda_t) \tag{12.7}$$
is an extension of the initial control system (12.1) to the cotangent bundle. Indeed, in canonical coordinates $\lambda = (\xi, x) \in T^*M$, the Hamiltonian system yields
$$\dot{x} = \frac{\partial h_{u(t)}}{\partial \xi} = f_{u(t)}(x).$$
That is, solutions λ_t to (12.7) are Hamiltonian lifts of solutions $q(t)$ to (12.1):
$$\pi(\lambda_t) = q_u(t).$$

Before proving Pontryagin Maximum Principle, we discuss its statement.

First we give a heuristic explanation of the way the covector curve λ_t appears naturally in the study of trajectories coming to boundary of the attainable set. Indeed, let
$$q_1 = \tilde{q}(t_1) \in \partial \mathcal{A}_{q_0}(t_1).$$
Consider a local convex approximation of the attainable set $\mathcal{A}_{q_0}(t_1)$ in the neighborhood of the point q_1, which is a convex cone in $T_{q_1}M$. Then $0 \in T_{q_1}M$ must belong to the boundary of the convex approximation of $\mathcal{A}_{q_0}(t_1)$. Thus the convex approximation has a hyperplane of support at q_1 determined by a covector $\lambda_{t_1} \in T^*M$, $\lambda_{t_1} \neq 0$ (the covector λ_{t_1} is an analog of Lagrange multipliers). In order to construct the whole curve λ_t, $t \in [0, t_1]$, consider the flow generated by the control $\tilde{u}(\cdot)$:
$$P_\tau^{t_1} = \overrightarrow{\exp} \int_\tau^{t_1} f_{\tilde{u}(t)}\, dt, \qquad \tau \in [0, t_1].$$
It is easy to see that
$$P_\tau^{t_1}(\mathcal{A}_{q_0}(\tau)) \subset \mathcal{A}_{q_0}(t_1), \qquad \tau \in [0, t_1].$$
Indeed, if a point $q \in \mathcal{A}_{q_0}(\tau)$ is reachable from q_0 by a control $u(t)$, $t \in [0, \tau]$, then the point $P_\tau^{t_1}(q)$ is reachable from q_0 by the control
$$v(t) = \begin{cases} u(t), & t \in [0, \tau], \\ \tilde{u}(t), & t \in [\tau, t_1]. \end{cases}$$
Further, the flow $P_\tau^{t_1} : M \to M$ satisfies the condition
$$P_\tau^{t_1}(\tilde{q}(\tau)) = \tilde{q}(t_1) = q_1, \qquad \tau \in [0, t_1].$$
Thus if $\tilde{q}(\tau) \in \operatorname{int} \mathcal{A}_{q_0}(\tau)$, then $q_1 \in \operatorname{int} \mathcal{A}_{q_0}(t_1)$. By contradiction, we obtain
$$\tilde{q}(\tau) \in \partial \mathcal{A}_{q_0}(\tau), \qquad \tau \in [0, t_1].$$

Consequently, we can find a hyperplane of support to the convex approximation of $\mathcal{A}_\tau(q_0)$ and the corresponding covector at any instant τ:

$$\lambda_\tau \in T^*_{\tilde{q}(\tau)}, \qquad \tau \in [0, t_1].$$

The covectors λ_t are defined up to nonzero factors. They can be renormalized so that satisfy the Hamiltonian system $\dot{\lambda}_t = \vec{h}_{\tilde{u}(t)}(\lambda_t)$.

So the covector curve λ_t in Pontryagin Maximum Principle appears naturally from hyperplanes of support to convex approximations of attainable sets.

Now we show the power of PMP by the following statement.

Proposition 12.1. *Assume that the maximized Hamiltonian of PMP*

$$H(\lambda) = \max_{u \in U} h_u(\lambda), \qquad \lambda \in T^*M,$$

is defined and C^2-smooth on $T^*M \setminus \{\lambda = 0\}$.

If a pair $(\tilde{u}(t), \lambda_t)$, $t \in [0, t_1]$, satisfies PMP, then

$$\dot{\lambda}_t = \vec{H}(\lambda_t), \qquad t \in [0, t_1]. \tag{12.8}$$

Conversely, if a Lipschitzian curve $\lambda_t \neq 0$ is a solution to the Hamiltonian system (12.8)*, then one can choose an admissible control $\tilde{u}(t)$, $t \in [0, t_1]$, such that the pair $(\tilde{u}(t), \lambda_t)$ satisfies PMP.*

That is, in the favorable case when the maximized Hamiltonian H is C^2-smooth, PMP reduces the problem to the study of solutions to just one Hamiltonian system (12.8). From the point of view of dimension, this reduction is the best one we can expect. Indeed, for a full-dimensional attainable set $(\dim \mathcal{A}_{q_0}(t_1) = n)$ we have $\dim \partial \mathcal{A}_{q_0}(t_1) = n - 1$, i.e., we need an $(n-1)$-parameter family of curves to describe the boundary $\partial \mathcal{A}_{q_0}(t_1)$. On the other hand, the family of solutions to Hamiltonian system (12.8) with the initial condition $\pi(\lambda_0) = q_0$ is n-dimensional. Taking into account that the Hamiltonian H is homogeneous:

$$H(c\lambda) = cH(\lambda), \qquad c > 0,$$

thus

$$e^{t\vec{H}}(c\lambda_0) = ce^{t\vec{H}}(\lambda_0), \qquad \pi \circ e^{t\vec{H}}(c\lambda_0) = \pi \circ e^{t\vec{H}}(\lambda_0),$$

we obtain the required $(n-1)$-dimensional family of curves.

Now we prove Proposition 12.1.

Proof. We show that if an admissible control $\tilde{u}(t)$ satisfies the maximality condition (12.5), then

$$\vec{h}_{\tilde{u}(t)}(\lambda_t) = \vec{H}(\lambda_t), \qquad t \in [0, t_1]. \tag{12.9}$$

By definition of the maximized Hamiltonian H,

$$H(\lambda) - h_{\tilde{u}(t)}(\lambda) \geq 0 \qquad \lambda \in T^*M, \quad t \in [0, t_1].$$

On the other hand, by the maximality condition of PMP (12.5), along the extremal λ_t this inequality turns into equality:

$$H(\lambda_t) - h_{\tilde{u}(t)}(\lambda_t) = 0, \qquad t \in [0, t_1].$$

That is why

$$d_{\lambda_t} H = d_{\lambda_t} h_{\tilde{u}(t)}, \qquad t \in [0, t_1].$$

But a Hamiltonian vector field is obtained from differential of the Hamiltonian by a standard linear transformation, thus equality (12.9) follows.

Conversely, let $\lambda_t \neq 0$ be a trajectory of the Hamiltonian system $\dot{\lambda}_t = \vec{H}(\lambda_t)$. In the same way as in the proof of Filippov's theorem, one can choose an admissible control $\tilde{u}(t)$ that realizes maximum along λ_t:

$$H(\lambda_t) = h_{\tilde{u}(t)}(\lambda_t) = \max_{u \in U} h_u(\lambda_t).$$

As we have shown above, then there holds equality (12.9). So the pair $(\tilde{u}(t), \lambda_t)$ satisfies PMP. $\square$

12.2 Proof of PMP

We start from two auxiliary propositions.

Denote the positive orthant in $\mathbb{R}^m$ as

$$\mathbb{R}^m_+ = \{(x_1, \ldots, x_m) \in \mathbb{R}^m \mid x_i \geq 0, \ i = 1, \ldots, m\}.$$

Lemma 12.1. *Let a vector-function $F : \mathbb{R}^m \to \mathbb{R}^n$ be Lipschitzian, $F(0) = 0$, and differentiable at 0:*

$$\exists \ F'_0 = \left. \frac{dF}{dx} \right|_0 .$$

Assume that

$$F'_0(\mathbb{R}^m_+) = \mathbb{R}^n.$$

Then for any neighborhood of the origin $O_0 \subset \mathbb{R}^m$

$$0 \in \operatorname{int} F(O_0 \cap \mathbb{R}^m_+).$$

Remark. The statement of the previous lemma holds if the orthant $\mathbb{R}^m_+$ is replaced by an arbitrary convex cone $C \subset \mathbb{R}^m$. In this case the proof given below works without any changes.

Proof. Choose points $y_0, \dots, y_n \in \mathbb{R}^n$ that generate an n-dimensional simplex centered at the origin:

$$\frac{1}{n+1} \sum_{i=0}^{n} y_i = 0.$$

Since the mapping $F_0' : \mathbb{R}^m_+ \to \mathbb{R}^n$ is surjective and the positive orthant $\mathbb{R}^m_+$ is convex, it is easy to show that restriction to the interior $F_0'|_{\operatorname{int} \mathbb{R}^m_+}$ is also surjective:

$$\exists\ v_i \in \operatorname{int} \mathbb{R}^m_+ \quad \text{such that} \quad F_0' v_i = y_i, \quad i = 0, \dots, n.$$

The points $y_0, \dots, y_n$ are affinely independent in $\mathbb{R}^n$, thus their preimages $v_0, \dots, v_n$ are also affinely independent in $\mathbb{R}^m$. The mean

$$v = \frac{1}{n+1} \sum_{i=0}^{n} v_i$$

belongs to $\operatorname{int} \mathbb{R}^m_+$ and satisfies the equality

$$F_0' v = 0.$$

Further, the subspace

$$W = \operatorname{span}\{v_i - v \mid i = 0, \dots, n\} \subset \mathbb{R}^m$$

is n-dimensional. Since $v \in \operatorname{int} \mathbb{R}^m_+$, we can find an n-dimensional ball $B_\delta \subset W$ of a sufficiently small radius δ centered at the origin such that

$$v + B_\delta \subset \operatorname{int} \mathbb{R}^m_+.$$

Since $F_0'(v_i - v) = F_0' v_i$, then $F_0' W = \mathbb{R}^n$, i.e., the linear mapping $F_0' : W \to \mathbb{R}^n$ is invertible.

Consider the following family of mappings:

$$\begin{aligned} &G_\alpha : B_\delta \to \mathbb{R}^n, \qquad \alpha \in [0, \alpha_0), \\ &G_\alpha(w) = \frac{1}{\alpha} F(\alpha(v + w)), \qquad \alpha > 0, \\ &G_0(w) = F_0' w. \end{aligned}$$

By the hypotheses of the proposition,

$$F(x) = F_0' x + o(x), \qquad x \in \mathbb{R}^m, \ x \to 0,$$

thus

$$G_\alpha(w) = F_0' w + o(1), \qquad \alpha \to 0, \quad w \in B_\delta. \tag{12.10}$$

Since the mapping F is Lipschitzian, all mappings G_α are Lipschitzian with a common constant. Thus the family G_α is equicontinuous. Equality (12.10) means that

$$G_\alpha \to G_0, \qquad \alpha \to 0,$$

pointwise, thus uniformly.

But $0 \in \operatorname{int} G_0 B_\delta$, and it is easy to show that $0 \in \operatorname{int} G_\alpha B_\delta$, $\alpha > 0$. Indeed, the mapping

$$w \mapsto \frac{G_\alpha(w)}{\|G_\alpha(w)\|}, \qquad w \in \partial B_\delta \tag{12.11}$$

from the $(n-1)$-dimensional sphere ∂B_δ to the unit sphere S^{n-1} has degree ± 1 for $\alpha = 0$, thus for small $\alpha > 0$. Then $0 \in \operatorname{int} G_\alpha B_\delta$, $\alpha > 0$. Consequently, $0 \in \operatorname{int} F(\alpha B_\delta)$ for small $\alpha > 0$. □

Now we start to compute a convex approximation of the attainable set $\mathcal{A}_{q_0}(t_1)$ at the point $q_1 = \tilde{q}(t_1)$. Take any admissible control $u(t)$ and express the endpoint of a trajectory via Variations Formula (2.28):

$$\begin{aligned} q_u(t_1) &= q_0 \circ \overrightarrow{\exp} \int_0^{t_1} f_{u(\tau)}\, d\tau = q_0 \circ \overrightarrow{\exp} \int_0^{t_1} f_{\tilde{u}(\tau)} + (f_{u(\tau)} - f_{\tilde{u}(\tau)})\, d\tau \\ &= q_0 \circ \overrightarrow{\exp} \int_0^{t_1} f_{\tilde{u}(\tau)}\, d\tau \circ \overrightarrow{\exp} \int_0^{t_1} \left(P_\tau^{t_1}\right)_* (f_{u(\tau)} - f_{\tilde{u}(\tau)})\, d\tau \\ &= q_1 \circ \overrightarrow{\exp} \int_0^{t_1} \left(P_\tau^{t_1}\right)_* (f_{u(\tau)} - f_{\tilde{u}(\tau)})\, d\tau. \end{aligned}$$

Introduce the following vector field depending on two parameters:

$$g_{\tau,u} = \left(P_\tau^{t_1}\right)_* (f_u - f_{\tilde{u}(\tau)}), \qquad \tau \in [0, t_1], \quad u \in U. \tag{12.12}$$

We showed that

$$q_u(t_1) = q_1 \circ \overrightarrow{\exp} \int_0^{t_1} g_{\tau, u(\tau)}\, d\tau. \tag{12.13}$$

Notice that

$$g_{\tau,\tilde{u}(\tau)} \equiv 0, \qquad \tau \in [0, t_1].$$

Lemma 12.2. *Let $\mathcal{T} \subset [0, t_1]$ be the set of Lebesgue points of the control $\tilde{u}(\cdot)$. If*

$$T_{q_1} M = \operatorname{cone}\{g_{\tau,u}(q_1) \mid \tau \in \mathcal{T},\ u \in U\},$$

then

$$q_1 \in \operatorname{int} \mathcal{A}_{q_0}(t_1).$$

Remark. The set $\operatorname{cone}\{g_{\tau,u}(q_1) \mid \tau \in \mathcal{T},\ u \in U\} \subset T_{q_1} M$ is a local convex approximation of the attainable set $\mathcal{A}_{q_0}(t_1)$ at the point q_1.

Recall that a point $\tau \in [0, t_1]$ is called a *Lebesgue point* of a function $u \in L^1[0, t_1]$ if

$$\lim_{t \to \tau} \frac{1}{|t - \tau|} \int_\tau^t |u(\theta) - u(\tau)| \, d\theta = 0.$$

At Lebesgue points of u, the integral $\displaystyle\int_0^t u(\theta)\, d\theta$ is differentiable and

$$\frac{d}{dt}\left(\int_0^t u(\theta)\, d\theta\right) = u(t).$$

The set of Lebesgue points has the full measure in the domain $[0, t_1]$. For details on this subject, see e.g. [21].

Now we prove Lemma 12.2.

Proof. We can choose vectors

$$g_{\tau_i, u_i}(q_1) \in T_{q_1}M, \qquad \tau_i \in \mathcal{T}, \quad u_i \in U, \quad i = 1, \ldots, k,$$

that generate the whole tangent space as a positive convex cone:

$$\operatorname{cone}\{g_{\tau_i, u_i}(q_1) \mid i = 1, \ldots, k\} = T_{q_1}M,$$

moreover, we can choose points τ_i distinct: $\tau_i \neq \tau_j$, $i \neq j$. Indeed, if $\tau_i = \tau_j$ for some $i \neq j$, we can find a sufficiently close Lebesgue point $\tau_j' \neq \tau_j$ such that the difference $g_{\tau_j', u_j}(q_1) - g_{\tau_j, u_j}(q_1)$ is as small as we wish. This is possible since for any $\tau \in \mathcal{T}$ and any $\varepsilon > 0$

$$\frac{1}{|t - \tau|} \operatorname{meas}\{t' \in [\tau, t] \mid |u(t') - u(\tau)| \leq \varepsilon\} \to 1 \text{ as } t \to \tau.$$

We suppose that $\tau_1 < \tau_2 < \cdots < \tau_k$.

We define a family of variations of controls that follow the control $\tilde{u}(\cdot)$ everywhere except neighborhoods of τ_i, and follow u_i near τ_i (such variations are called *needle-like*).

More precisely, for any $s = (s_1, \ldots, s_k) \in \mathbb{R}^k_+$ consider a control of the form

$$u_s(t) = \begin{cases} u_i, & t \in [\tau_i, \tau_i + s_i], \\ \tilde{u}(t), & t \notin \cup_{i=1}^k [\tau_i, \tau_i + s_i]. \end{cases} \tag{12.14}$$

For small s, the segments $[\tau_i, \tau_i + s_i]$ do not overlap since $\tau_i \neq \tau_j$, $i \neq j$. In view of formula (12.13), the endpoint of the trajectory corresponding to the control constructed is expressed as follows:

$$\begin{aligned} q_{u_s}(t_1) &= q_0 \circ \overrightarrow{\exp} \int_0^{t_1} f_{u_s(t)}\, dt \\ &= q_1 \circ \overrightarrow{\exp} \int_{\tau_1}^{\tau_1 + s_1} g_{t, u_1}\, dt \circ \overrightarrow{\exp} \int_{\tau_2}^{\tau_2 + s_2} g_{t, u_2}\, dt \circ \cdots \\ &\qquad \circ \overrightarrow{\exp} \int_{\tau_k}^{\tau_k + s_k} g_{t, u_k}\, dt. \end{aligned}$$

The mapping

$$F \ : \ s = (s_1, \dots, s_k) \mapsto q_{u_s}(t_1)$$

is Lipschitzian, differentiable at $s = 0$, and

$$\left. \frac{\partial F}{\partial s_i} \right|_{s=0} = g_{\tau_i, u_i}(q_1).$$

By Lemma 12.1,

$$F(0) = q_1 \in \operatorname{int} F(O_0 \cap \mathbb{R}^k_+)$$

for any neighborhood $O_0 \subset \mathbb{R}^k$. But the curve $q_{u_s}(t)$, $t \in [0, t_1]$, is an admissible trajectory for small $s \in \mathbb{R}^k_+$, thus $F(O_0 \cap \mathbb{R}^k_+) \subset \mathcal{A}_{q_0}(t_1)$ and $q_1 \in \operatorname{int} \mathcal{A}_{q_0}(t_1)$. □

Now we can prove the geometric statement of Pontryagin Maximum Principle, Theorem 12.1.

Proof. Let the endpoint of the reference trajectory

$$q_1 = \tilde{q}(t_1) \in \partial \mathcal{A}_{q_0}(t_1).$$

By Lemma 12.2, the origin $0 \in T_{q_1} M$ belongs to the boundary of the convex set $\operatorname{cone}\{g_{t,u}(q_1) \mid t \in \mathcal{T}, \ u \in U\}$, so this set has a hyperplane of support at the origin:

$$\exists \ \lambda_{t_1} \in T^*_{q_1} M, \quad \lambda_{t_1} \neq 0,$$

such that

$$\langle \lambda_{t_1}, g_{t,u}(q_1) \rangle \leq 0 \qquad \forall \text{ a.e. } t \in [0, t_1], \quad u \in U.$$

Taking into account definition (12.12) of the field $g_{t,u}$, we rewrite this inequality as follows:

$$\langle \lambda_{t_1}, \left(P^{t_1}_{t*} f_u \right) (q_1) \rangle \leq \langle \lambda_{t_1}, \left(P^{t_1}_{t*} f_{\tilde{u}(t)} \right) (q_1) \rangle,$$

i.e.,

$$\langle \left(P^{t_1}_t \right)^* \lambda_{t_1}, f_u(\tilde{q}(t)) \rangle \leq \langle \left(P^{t_1}_t \right)^* \lambda_{t_1}, f_{\tilde{u}(t)}(\tilde{q}(t)) \rangle.$$

The action of the flow $P^{t_1}_t$ on covectors defines the curve in the cotangent bundle:

$$\lambda_t \ \stackrel{\text{def}}{=} \ \left(P^{t_1}_t \right)^* \lambda_{t_1} \in T^*_{\tilde{q}(t)} M, \qquad t \in [0, t_1].$$

In terms of this covector curve, the inequality above reads

$$\langle \lambda_t, f_u(\tilde{q}(t)) \rangle \leq \langle \lambda_t, f_{\tilde{u}(t)}(\tilde{q}(t)) \rangle$$

Thus the maximality condition of PMP (12.5) holds along the reference trajectory:

$$h_u(\lambda_t) \leq h_{\tilde{u}(t)}(\lambda_t) \qquad \forall \ u \in U \quad \forall \text{ a.e. } t \in [0, t_1].$$

By Proposition 11.3, the curve λ_t is a trajectory of the nonautonomous Hamiltonian flow with the Hamiltonian function $f^*_{\tilde{u}(t)} = h_{\tilde{u}(t)}$:

$$\lambda_t = \lambda_{t_1} \circ \left(\overrightarrow{\exp} \int_t^{t_1} f_{\tilde{u}(\theta)}\, d\theta \right)^* = \lambda_{t_1} \circ\, \overrightarrow{\exp} \int_{t_1}^{t} \vec{h}_{\tilde{u}(\theta)}\, d\theta,$$

thus it satisfies the Hamiltonian equation of PMP (12.4)

$$\dot{\lambda}_t = \vec{h}_{\tilde{u}(t)}(\lambda_t).$$

□

12.3 Geometric statement of PMP for free time

In the previous section we proved Pontryagin Maximum Principle for the case of fixed terminal time t_1. Now we consider the case of free t_1.

Theorem 12.2. *Let $\tilde{u}(\cdot)$ be an admissible control for control system* (12.1) *such that*

$$\tilde{q}(t_1) \in \partial \left(\cup_{|t-t_1|<\varepsilon} \mathcal{A}_{q_0}(t) \right)$$

for some $t_1 > 0$ and $\varepsilon \in (0, t_1)$. Then there exists a Lipschitzian curve

$$\lambda_t \in T^*_{\tilde{q}(t)} M, \quad \lambda_t \neq 0, \qquad 0 \leq t \leq t_1,$$

such that

$$\begin{aligned} &\dot{\lambda}_t = \vec{h}_{\tilde{u}(t)}(\lambda_t), \\ &h_{\tilde{u}(t)}(\lambda_t) = \max_{u \in U} h_u(\lambda_t), \\ &h_{\tilde{u}(t)}(\lambda_t) = 0 \end{aligned} \tag{12.15}$$

for almost all $t \in [0, t_1]$.

Remark. In problems with free time, there appears one more variable, the terminal time t_1. In order to eliminate it, we have one additional condition — equality (12.15). This condition is indeed scalar since the previous two equalities imply that $h_{\tilde{u}(t)}(\lambda_t) = \text{const}$, see remark after formulation of Theorem 12.1.

Proof. We reduce the case of free time to the case of fixed time by extension of the control system via substitution of time. Admissible trajectories of the extended system are reparametrized admissible trajectories of the initial system (the positive direction of time on trajectories is preserved).

Let a new time be a smooth function

$$\varphi : \mathbb{R} \to \mathbb{R}, \quad \dot{\varphi} > 0.$$

We find an ODE for a reparametrized trajectory:

$$\frac{d}{dt} q_u(\varphi(t)) = \dot{\varphi}(t) f_{u(\varphi(t))}(q_u(\varphi(t))),$$

so the required equation is

$$\dot{q} = \dot{\varphi}(t) f_{u(\varphi(t))}(q).$$

Now consider along with the initial control system

$$\dot{q} = f_u(q), \qquad u \in U,$$

an extended system of the form

$$\dot{q} = v f_u(q), \qquad u \in U, \ |v - 1| < \delta, \tag{12.16}$$

where $\delta = \varepsilon / t_1 \in (0, 1)$. Admissible controls of the new system are

$$w(t) = (v(t), u(t)),$$

and the reference control corresponding to the control $\tilde{u}(\cdot)$ of the initial system is

$$\tilde{w}(t) = (1, \tilde{u}(t)).$$

It is easy to see that since $\tilde{q}(t_1) \in \partial \left(\cup_{|t - t_1| < \varepsilon} \mathcal{A}_{q_0}(t) \right)$, then the trajectory of the new system through the point q_0 corresponding to the control $\tilde{w}(\cdot)$ comes at the moment t_1 to the boundary of the attainable set of the new system for time t_1. Thus $\tilde{w}(t)$ satisfies PMP with fixed time. We apply Theorem 12.1 to the new system (12.16). The Hamiltonian for the new system is $v h_u(\lambda)$. Then the maximality condition (12.5) reads

$$1 \cdot h_{\tilde{u}(t)}(\lambda_t) = \max_{u \in U, \ |v-1| < \delta} v h_u(\lambda_t).$$

We take $u = \tilde{u}(t)$ under the maximum and obtain

$$h_{\tilde{u}(t)}(\lambda_t) = 0,$$

then we restrict the maximum to the set $v = 1$ and come to

$$h_{\tilde{u}(t)}(\lambda_t) = \max_{u \in U} h_u(\lambda_t).$$

The Hamiltonian systems along $\tilde{w}(\cdot)$ and $\tilde{u}(\cdot)$ coincide one with another, thus the proposition follows. □

12.4 PMP for optimal control problems

Now we apply PMP in geometric form to optimal control problems, starting from problems with fixed time.

For a control system

$$\dot{q} = f_u(q), \qquad q \in M, \quad u \in U, \tag{12.17}$$

with the boundary conditions

$$q(0) = q_0, \quad q(t_1) = q_1, \quad q_0, q_1 \in M \text{ fixed}, \tag{12.18}$$

$$t_1 > 0 \text{ fixed}, \tag{12.19}$$

and the cost functional

$$J(u) = \int_0^{t_1} \varphi(q_u(t), u(t))\, dt \tag{12.20}$$

we consider the optimal control problem

$$J(u) \to \min. \tag{12.21}$$

We transform the problem as in Sec. 10.2. We extend the state space:

$$\widehat{q} = \begin{pmatrix} y \\ q \end{pmatrix} \in \mathbb{R} \times M,$$

define the extended vector field $\widehat{f}_u \in \operatorname{Vec}(\mathbb{R} \times M)$:

$$\widehat{f}_u(q) = \begin{pmatrix} \varphi(q, u) \\ f_u(q) \end{pmatrix},$$

and come to the new control system:

$$\frac{d\widehat{q}}{dt} = \widehat{f}_u(q) \quad \Leftrightarrow \quad \begin{cases} \dot{y} = \varphi(q, u), \\ \dot{q} = f_u(q) \end{cases} \tag{12.22}$$

with the boundary conditions

$$\widehat{q}(0) = \widehat{q}_0 = \begin{pmatrix} 0 \\ q_0 \end{pmatrix}, \qquad \widehat{q}(t_1) = \begin{pmatrix} J(u) \\ q_1 \end{pmatrix}.$$

If a control $\tilde{u}(\cdot)$ is optimal for problem (12.17)–(12.21), then the trajectory $\widehat{q}_{\tilde{u}}(t)$ of the extended system (12.22) starting from $\widehat{q}_0$ satisfies the condition

$$\widehat{q}_{\tilde{u}}(t_1) \in \partial \widehat{\mathcal{A}}_{\widehat{q}_0}(t_1),$$

where $\widehat{\mathcal{A}}_{\widehat{q}_0}(t_1)$ is the attainable set of system (12.22) from the point $\widehat{q}_0$ for time t_1. So we can apply Theorem 12.1.

But the geometric form of PMP applied to the extended system (12.22) does not distinguish minimum and maximum of the cost $J(u)$. In order to have conditions valid only for minimum, we introduce a new control parameter v and consider a new system of the form

$$\begin{cases} \dot{y} = \varphi(q, u) + v, \\ \dot{q} = f_u(q), \end{cases} \qquad v \geq 0, \quad u \in U. \tag{12.23}$$

Now the trajectory of system (12.23) corresponding to the controls $\tilde{v}(t) \equiv 0$, $\tilde{u}(t)$, comes to the boundary of the attainable set of this system at time t_1. We apply Theorem 12.1 to system (12.23). We have

$$T_{(y,q)}(\mathbb{R} \times M) = \mathbb{R} \oplus T_q M,$$
$$T^*_{(y,q)}(\mathbb{R} \times M) = \mathbb{R} \oplus T^*_q M = \{(\nu, \lambda)\}.$$

The Hamiltonian function for system (12.23) has the form

$$\widehat{h}_{(v,u)}(\nu, \lambda) = \langle \lambda, f_u \rangle + \nu(\varphi + v),$$

and the Hamiltonian system of PMP is

$$\begin{cases} \dot{\nu} = \frac{\partial \widehat{h}}{\partial y} = 0, \\ \dot{y} = \varphi(q, u) + v, \\ \dot{\lambda} = \vec{h}_{\tilde{u}(t)}(\nu, \lambda). \end{cases} \tag{12.24}$$

Here $\vec{h}_u(\nu, \lambda)$ is the Hamiltonian vector field with the Hamiltonian function

$$h_u(\nu, \lambda) = \langle \lambda, f_u \rangle + \nu\varphi.$$

The first of equations (12.24) means that

$$\nu = \text{const}$$

along the reference trajectory.

The maximality condition has the form

$$\langle \lambda_t, f_{\tilde{u}(t)} \rangle + \nu\varphi(\tilde{q}(t), \tilde{u}(t)) = \max_{u \in U,\ v \geq 0} \left(\langle \lambda_t, f_u \rangle + \nu\varphi(\tilde{q}(t), u) + \nu v \right).$$

Since the previous maximum is attained, we have

$$\nu \leq 0,$$

thus we can set $v = 0$ in the right-hand side of the maximality condition:

$$\langle \lambda_t, f_{\tilde{u}(t)} \rangle + \nu\varphi(\tilde{q}(t), \tilde{u}(t)) = \max_{u \in U} \left(\langle \lambda_t, f_u \rangle + \nu\varphi(\tilde{q}(t), u) \right).$$

So we proved the following statement.

Theorem 12.3. *Let* $\tilde{u}(t)$, $t \in [0, t_1]$, *be an optimal control for problem* (12.17)–(12.21)*:*

$$J(\tilde{u}) = \min\{J(u) \mid q_u(t_1) = q_1\}.$$

Define a Hamiltonian function

$$h_u^\nu(\lambda) = \langle \lambda, f_u \rangle + \nu\varphi(q, u), \qquad \lambda \in T_q^*M, \quad u \in U, \quad \nu \in \mathbb{R}.$$

Then there exists a nontrivial pair:

$$(\nu, \lambda_t) \neq 0, \qquad \nu \in \mathbb{R}, \quad \lambda_t \in T_{\tilde{q}(t)}^*M,$$

such that the following conditions hold:

$$\begin{aligned}
&\dot{\lambda}_t = \vec{h}_{\tilde{u}(t)}^\nu(\lambda_t), \\
&h_{\tilde{u}(t)}^\nu(\lambda_t) = \max_{u \in U} h_u^\nu(\lambda_t) \quad \forall \text{ a.e. } t \in [0, t_1], \\
&\nu \leq 0.
\end{aligned}$$

Remarks. (1) If we have a maximization problem instead of minimization problem (12.21), then the preceding inequality for ν should be reversed:

$$\nu \geq 0.$$

(2) For the problem with free time t_1: (12.17), (12.18), (12.20), (12.21), necessary optimality conditions of PMP are the same as in Theorem 12.3 plus one additional scalar equality $h_{\tilde{u}(t)}^\nu(\lambda_t) \equiv 0$ (exercise).

There are two distinct possibilities for the constant parameter ν in Theorem 12.3:

(a) if $\nu \neq 0$, then the curve λ_t is called a *normal extremal*. Since the pair (ν, λ_t) can be multiplied by any positive number, we can normalize $\nu < 0$ and assume that $\nu = -1$ in the normal case;

(b) if $\nu = 0$, then λ_t is an *abnormal extremal*.

So we can always assume that $\nu = -1$ or 0.

Now consider the time-optimal problem:

$$\begin{aligned}
&\dot{q} = f_u(q), \qquad q \in M, \quad u \in U, \\
&q(0) = q_0, \quad q(t_1) = q_1, \quad q_0, q_1 \text{ fixed}, \\
&t_1 = \int_0^{t_1} 1 \, dt \to \min.
\end{aligned}$$

For the time-optimal problem, Pontryagin Maximum Principle takes the following form.

Corollary 12.1. *Let an admissible control* $\tilde{u}(t)$, $t \in [0, t_1]$, *be time-optimal. Define a Hamiltonian function*

$$h_u(\lambda) = \langle \lambda, f_u \rangle, \qquad \lambda \in T_q^*M, \quad u \in U.$$

Then there exists a Lipschitzian curve

$$\lambda_t \in T^*M, \quad \lambda_t \neq 0, \qquad t \in [0, t_1],$$

such that the following conditions hold for almost all $t \in [0, t_1]$*:*

$$\begin{aligned} \dot{\lambda}_t &= \vec{h}_{\tilde{u}(t)}(\lambda_t), \\ h_{\tilde{u}(t)}(\lambda_t) &= \max_{u \in U} h_u(\lambda_t), \\ h_{\tilde{u}(t)}(\lambda_t) &\geq 0. \end{aligned} \tag{12.25}$$

Proof. Apply Theorem 12.3 and the second remark after it, taking $\varphi \equiv 1$. Then the Hamiltonian system and the maximality condition follow. Inequality (12.25) is equivalent to conditions $h_{\tilde{u}(t)}(\lambda_t) + \nu = 0$ and $\nu \leq 0$.

The inequality $\lambda_t \neq 0$ is obtained as follows: if $\lambda_t = 0$, then $h_{\tilde{u}(t)}(\lambda_t) = 0$, thus $\nu = 0$. But the pair (ν, λ_t) must be nontrivial, consequently, $\lambda_t \neq 0$. □

12.5 PMP with general boundary conditions

In this section we prove versions of Pontryagin Maximum Principle for optimal control problems in which boundary points of trajectories belong to prescribed manifolds.

First consider the following problem:

$$\dot{q} = f_u(q), \qquad q \in M, \quad u \in U \subset \mathbb{R}^m, \tag{12.26}$$

$$q(0) \in N_0, \qquad q(t_1) \in N_1, \tag{12.27}$$

$$t_1 > 0 \text{ fixed}, \tag{12.28}$$

$$J(u) = \int_0^{t_1} \varphi(q(t), u(t))\, dt \to \min. \tag{12.29}$$

Here N_0 and N_1 are given immersed submanifolds of the state space M. So the boundary points $q(0)$ and $q(t_1)$ are not fixed as before, but should belong to N_0 and N_1 respectively.

If a trajectory $\tilde{q}(t)$ is optimal for this problem, then it is optimal as well for the problem with the fixed boundary points $\tilde{q}(0)$, $\tilde{q}(t_1)$ considered in Section 12.4. Consequently, the statement of Theorem 12.3 should be satisfied for $\tilde{q}(t)$. But now we need additional conditions that select boundary points $\tilde{q}(0) \in N_0$ and $\tilde{q}(t_1) \in N_1$. It is reasonable to expect that they should be determined by $(\dim N_0 + \dim N_1)$ scalar equalities. Such conditions can easily be formulated in the Hamiltonian framework, they are called *transversality conditions*, see (12.34) below.

Theorem 12.4. *Let* $\tilde{u}(t)$, $t \in [0, t_1]$, *be an optimal control in problem* (12.26)–(12.29). *Define a family of Hamiltonians:*

$$h_u^\nu(\lambda) = \langle \lambda, f_u(q) \rangle + \nu \varphi(q, u), \qquad \lambda \in T_q^*M, \ q \in M, \quad \nu \in \mathbb{R}, \ u \in U.$$

Then there exists a Lipschitzian curve $\lambda_t \in T^*_{\tilde q(t)}M$, $t \in [0, t_1]$, *and a number* $\nu \in \mathbb{R}$ *such that:*

$$\dot\lambda_t = \overrightarrow{h^\nu_{\tilde u(t)}}(\lambda_t), \tag{12.30}$$

$$h^\nu_{\tilde u(t)}(\lambda_t) = \max_{u\in U} h^\nu_u(\lambda_t), \tag{12.31}$$

$$(\lambda_t, \nu) \not\equiv (0,0), \qquad t \in [0, t_1], \tag{12.32}$$

$$\nu \le 0, \tag{12.33}$$

$$\lambda_0 \perp T_{\tilde q(0)}N_0, \qquad \lambda_{t_1} \perp T_{\tilde q(t_1)}N_1. \tag{12.34}$$

Remarks. (1) Any linear functional on a linear space acts naturally on a subspace by restriction, so transversality conditions (12.34) read respectively as follows:

$$\langle \lambda_0, v\rangle = 0, \qquad v \in T_{\tilde q(0)}N_0,$$
$$\langle \lambda_{t_1}, w\rangle = 0, \qquad w \in T_{\tilde q(t_1)}N_1.$$

(2) The problem with free time: (12.26), (12.27), (12.29), is reduced to the case of fixed t_1 in the same way as in Section 12.4, so for this problem holds the previous theorem with the additional condition $h^\nu_{\tilde u(t)}(\lambda_t) \equiv 0$.

Now we prove Theorem 12.4.

Proof. The scheme of proof of PMP developed in Theorems 12.1, 12.3 can be applied to much more general problems after appropriate modifications. Now we only indicate how the proofs of these theorems should be changed in order to cover the new boundary conditions $q(0) \in N_0$, $q(t_1) \in N_1$.

(1) First consider the special case where the initial point is fixed: let

$$N_0 = \{q_0\}$$

for some point $q_0 \in M$.

As in the proof of Theorem 12.3, we introduce an extended system on $\mathbb{R}\times M$:

$$\widehat q = \begin{pmatrix} y \\ q \end{pmatrix} \in \mathbb{R}\times M,$$

$$\widehat f_u(q) = \begin{pmatrix} \varphi(q,u)+v \\ f_u(q) \end{pmatrix} \in T_{(y,\widehat q)}(\mathbb{R}\times M) = \mathbb{R}\times T_qM,$$

$$\frac{d\widehat q}{dt} = \widehat f_u(q) \quad \Leftrightarrow \quad \begin{cases} \dot y = \varphi(q,u)+v, \\ \dot q = f_u(q), \end{cases} \tag{12.35}$$

$$\widehat q(0) = \widehat q_0 = \begin{pmatrix} 0 \\ q_0 \end{pmatrix}.$$

Further, in the case of fixed terminal point $q(t_1)$, the necessary condition for optimality of the trajectory $q_{\tilde u}(t)$ was the following:

$$\widehat q_1 \in \partial \widehat{\mathcal{A}}_{\widehat q_0}(t_1). \tag{12.36}$$

Here $\widehat{\mathcal{A}}$ is the attainable set of the extended system (12.35) and $\widehat{q}_1 = \widehat{q}_{\tilde{u}}(t_1)$.

Now, when the target manifold N_1 is not a point, we should modify the argument. In a sense, we reduce the target manifold to a point defining it locally by an equation $\Phi = 0$. Choose a submersion

$$\Phi : O_{q_{\tilde{u}(t_1)}} \to \mathbb{R}^p, \qquad p = \dim M - \dim N_1,$$

of a small neighborhood $O_{q_{\tilde{u}(t_1)}} \subset M$, so that locally

$$\Phi^{-1}(0) = N_1.$$

Further, extend the submersion: define the mapping

$$\widehat{\Phi} : \mathbb{R} \times O_{q_{\tilde{u}(t_1)}} \to \mathbb{R}^{1+p}, \qquad \widehat{\Phi}\begin{pmatrix} y \\ q \end{pmatrix} = \begin{pmatrix} y \\ \Phi(q) \end{pmatrix}.$$

Since the control $\tilde{u}(t)$ is optimal in our problem (12.26)–(12.29), then

$$\widehat{\Phi}(\widehat{q}_1) \in \partial \widehat{\Phi}(\widehat{\mathcal{A}}_{\widehat{q}_0}(t_1)). \tag{12.37}$$

So we replace the necessary optimality condition (12.36) by (12.37) and return to the scheme of proof of Theorems 12.1, 12.3.

Take any $k \in \mathbb{N}$ and any needle-like variation (12.14) of the optimal control:

$$u_s(t), \quad s \in \mathbb{R}^k_+, \qquad u_0(t) = \tilde{u}(t), \quad t \in [0, t_1].$$

Define the mappings

$$G : \mathbb{R}^k \to \mathbb{R} \times M, \qquad G(s) = \widehat{q}_{u_s}(t_1) = \widehat{q}_0 \circ \overrightarrow{\exp} \int_0^{t_1} \widehat{f}_{u_s(t)}\, dt, \tag{12.38}$$

$$F : \mathbb{R}^k \to \mathbb{R}^{1+p}, \qquad F(s) = \widehat{\Phi}(G(s)) = \widehat{q}_0 \circ \overrightarrow{\exp} \int_0^{t_1} \widehat{f}_{u_s(t)}\, dt \circ \widehat{\Phi}. \tag{12.39}$$

Then it follows from inclusion (12.37) that

$$\widehat{\Phi}(\widehat{q}_1) = F(0) \in \partial F(\mathbb{R}^k_+). \tag{12.40}$$

By Lemma 12.1,

$$F_0'(\mathbb{R}^k_+) = \operatorname{cone}\left\{ \left.\frac{\partial F}{\partial s_i}\right|_0 \mid i = 1, \dots, k \right\} \neq \mathbb{R}^{1+p},$$

thus there exists a plane of support, i.e.,

$$\exists\, \widehat{\xi} \in \left(\mathbb{R}^{1+p}\right)^*, \quad \widehat{\xi} \neq 0,$$

such that

$$\left\langle \widehat{\xi}, \left.\frac{\partial F}{\partial s_i}\right|_0 \right\rangle \leq 0, \qquad i = 1, \dots, k. \tag{12.41}$$

We compute the derivative by the chain rule:

$$\left.\frac{\partial F}{\partial s_i}\right|_0 = \widehat{\Phi}_* \left.\frac{\partial G}{\partial s_i}\right|_0, \tag{12.42}$$

and rewrite inequalities (12.41) as follows:

$$\left\langle \widehat{\Phi}^* \widehat{\xi}, \left.\frac{\partial G}{\partial s_i}\right|_0 \right\rangle = \left\langle \widehat{\xi}, \widehat{\Phi}_* \left.\frac{\partial G}{\partial s_i}\right|_0 \right\rangle \le 0, \qquad i = 1, \dots, k. \tag{12.43}$$

Then we denote the covector

$$\widehat{\lambda}_{t_1} = \widehat{\Phi}^* \widehat{\xi} = \begin{pmatrix} \nu \\ \lambda_{t_1} \end{pmatrix} \in T_{\widehat{q}_1}(\mathbb{R} \times M) \tag{12.44}$$

and obtain conclusions (12.30)–(12.33) in the same way as in Theorem 12.3. The only distinction now is that the covector $\widehat{\lambda}_{t_1}$ is not arbitrary: equality (12.44) implies the second of the transversality conditions (12.34). Indeed, we have

$$\lambda_{t_1} = \Phi^* \xi, \qquad \xi \in (\mathbb{R}^p)^*,$$

thus

$$\langle \lambda_{t_1}, T_{q_{\tilde{u}}(t_1)} N_1 \rangle = \langle \Phi^* \xi, T_{q_{\tilde{u}}(t_1)} N_1 \rangle = \langle \xi, \underbrace{\Phi_* T_{q_{\tilde{u}}(t_1)} N_1}_{=0} \rangle = 0.$$

The first transversality condition (12.34) is now trivially satisfied, so the proof of this theorem in the case $N_0 = \{q_0\}$ is complete.

(2) Let now the initial manifold N_0 be an arbitrary immersed submanifold of M. We can modify the scheme presented above to cover this case as well. Since now the initial point $q(0)$ is not fixed, we add variations of $q(0)$.

Replace mappings (12.38), (12.39) by the following ones:

$$G : N_0 \times \mathbb{R}^k \to \mathbb{R} \times M, \qquad G(q, s) = \widehat{q} \circ \overrightarrow{\exp} \int_0^{t_1} \widehat{f}_{u_s(t)}\, dt,$$

$$F : N_0 \times \mathbb{R}^k \to \mathbb{R}^{1+p}, \qquad F(q, s) = \widehat{\Phi}(G(q, s)) = \widehat{q} \circ \overrightarrow{\exp} \int_0^{t_1} \widehat{f}_{u_s(t)}\, dt \circ \widehat{\Phi},$$

where $\widehat{q} = (0, q) \in \mathbb{R} \times M$. Then the necessary optimality condition (12.40) is replaced by the inclusion

$$F(\tilde{q}(0), 0) \in \partial F(N_0 \times \mathbb{R}^k_+). \tag{12.45}$$

Apply Lemma 12.1 to restriction of the mapping F to the space

$$\mathbb{R}^m \cong O_{\tilde{q}(0)} \times \mathbb{R}^k, \qquad m = l + k, \quad l = \dim N_0,$$

where $O_{\tilde{q}(0)} \subset N_0$ is a small neighborhood of $\tilde{q}(0)$. By the remark after Lemma 12.1, inclusion (12.45) implies that

$$F'_{(\tilde{q}(0),0)}(\mathbb{R}^l \oplus \mathbb{R}^k_+) \ne \mathbb{R}^{1+p},$$

i.e., there exists a covector

$$\widehat{\xi} \in \left(\mathbb{R}^{1+p}\right)^*, \qquad \widehat{\xi} \neq 0, \quad \widehat{\xi} = \begin{pmatrix} \nu \\ \xi \end{pmatrix},$$

such that

$$\begin{aligned} \left\langle \widehat{\xi}, \frac{\partial F}{\partial q} v \right\rangle &\leq 0, \qquad v \in T_{\tilde{q}(0)} N_0, \\ \left\langle \widehat{\xi}, \frac{\partial F}{\partial s_i} \right\rangle &\leq 0, \qquad i = 1, \dots, k. \end{aligned} \tag{12.46}$$

In the first inequality v belongs to a linear space, thus it turns into equality:

$$\left\langle \widehat{\xi}, \frac{\partial F}{\partial q} v \right\rangle = 0, \qquad v \in T_{\tilde{q}(0)} N_0. \tag{12.47}$$

Compute by Leibniz rule the partial derivative:

$$\begin{aligned} \left.\frac{\partial F}{\partial q}\right|_{(\tilde{q}(0),0)} &: T_{\tilde{q}(0)} N_0 \to \mathbb{R}^{1+p}, \\ \left.\frac{\partial F}{\partial q}\right|_{(\tilde{q}(0),0)} v &= \begin{pmatrix} 0 \\ v \end{pmatrix} \circ \overrightarrow{\exp} \int_0^{t_1} \widehat{f}_{\tilde{u}(t)} \, dt \circ \widehat{\Phi} = \begin{pmatrix} 0 \\ v \circ P^{t_1} \circ \Phi \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \Phi_* P_*^{t_1} v \end{pmatrix}, \qquad v \in T_{\tilde{q}(0)} N_0. \end{aligned}$$

Here we applied formula (2.19) to the flow

$$P^{t_1} = \overrightarrow{\exp} \int_0^{t_1} f_{\tilde{u}(t)} \, dt.$$

Then conditions (12.47), (12.46) read as follows:

$$\begin{aligned} \left\langle \xi, \Phi_* P_*^{t_1} v \right\rangle &= 0, \qquad v \in T_{\tilde{q}(0)} N_0, \\ \left\langle \widehat{\Phi}^* \widehat{\xi}, \left.\frac{\partial G}{\partial s_i}\right|_{(\tilde{q}(0),0)} \right\rangle &\leq 0, \qquad i = 1, \dots, k. \end{aligned} \tag{12.48}$$

As before, introduce the covector $\widehat{\lambda}_{t_1} = (\nu, \lambda_{t_1})$ by equality (12.44), then conclusions (12.30)–(12.33) of this theorem and the second transversality condition (12.34) follow.

The first transversality condition is also satisfied: equality (12.48) can be rewritten as

$$\left\langle \lambda_{t_1}, P_*^{t_1} v \right\rangle = 0, \qquad v \in T_{\tilde{q}(0)} N_0.$$

But $\lambda_0 = P_{t_1}^* \lambda_{t_1}$, thus

$$\langle \lambda_0, v \rangle = \left\langle P_{t_1}^* \lambda_{t_1}, v \right\rangle = 0, \qquad v \in T_{\tilde{q}(0)} N_0.$$

The theorem is completely proved. □

Now consider even more general problem with mixed boundary conditions, see inclusion (12.50) below. Pontryagin Maximum Principle easily generalizes to this case, both in formulation and in proof.

We study optimal control problem of the form:

$$\dot{q} = f_u(q), \qquad q \in M, \quad u \in U \subset \mathbb{R}^m, \tag{12.49}$$

$$(q(0), q(t_1)) \in N \subset M \times M, \tag{12.50}$$

$$t_1 > 0 \text{ fixed}, \tag{12.51}$$

$$J(u) = \int_0^{t_1} \varphi(q(t), u(t))\, dt \to \min, \tag{12.52}$$

where N is a smooth immersed submanifold of $M \times M$.

Theorem 12.5. *Let $\tilde{u}$ be an optimal control in problem* (12.49)–(12.52)*. Then there hold all statements of Theorem* 12.4 *except its transversality condition* (12.34)*, which is replaced now by the relation*

$$(-\lambda_0, \lambda_{t_1}) \perp T_{(\tilde{q}(0), \tilde{q}(t_1))} N. \tag{12.53}$$

Remarks. (1) We identify

$$T^*_{(q_0,q_1)}(M \times M) \cong T^*_{q_0} M \oplus T^*_{q_1} M,$$

so the transversality condition (12.53) makes sense.

(2) An important particular case of mixed boundary conditions (12.50) is the case of periodic trajectories:

$$q(t_1) = q(0). \tag{12.54}$$

Indeed, then

$$N = \Delta \overset{\text{def}}{=} \{(q, q) \mid q \in M\} \subset M \times M, \tag{12.55}$$

the diagonal of the product $M \times M$. In this case the transversality condition (12.53) reads

$$\langle(-\lambda_0, \lambda_{t_1}), (v, v)\rangle = -\langle\lambda_0, v\rangle + \langle\lambda_{t_1}, v\rangle = 0, \qquad v \in T_{q(0)}M = T_{q(t_1)}M,$$

i.e.,

$$\lambda_0 = \lambda_{t_1}.$$

That is, an optimal trajectory in the problem with periodic boundary conditions (12.54) possesses a periodic Hamiltonian lift (extremal).

Now we prove Theorem 12.5.

Proof. We reduce our problem to the case of separated boundary conditions by introducing an auxiliary problem on $M \times M$:

$$\begin{cases} \dot{x} = 0, \\ \dot{q} = f_u(q), \end{cases} \qquad (x, q) \in M \times M, \quad u \in U,$$

$$(x(0), q(0)) \in \Delta, \qquad (x(t_1), q(t_1)) \in N,$$

(the diagonal Δ is defined in (12.55) above)

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

It is obvious that this problem is equivalent to our problem (12.49)–(12.52). We apply a version of PMP (Theorem 12.4) to the auxiliary problem. The Hamiltonian is the same as for the initial problem:

$$h_u^\nu(\eta, \lambda) = h_u^\nu(\lambda) = \langle \lambda, f_u(q) \rangle + \nu\varphi(q, u), \qquad (\eta, \lambda) \in T^*M \oplus T^*M.$$

The corresponding Hamiltonian system is

$$\begin{cases} \dot{\eta}_t = 0, \\ \dot{\lambda}_t = \overrightarrow{h^\nu_{\tilde{u}(t)}}(\lambda_t). \end{cases} \tag{12.56}$$

All required statements of PMP obviously follow, we should only check transversality conditions.

At the initial instant $t = 0$ the first of conditions (12.34) reads:

$$\langle (\eta_0, \lambda_0), (v, v) \rangle = \langle \eta_0, v \rangle + \langle \lambda_0, v \rangle = 0, \qquad v \in T_{\tilde{q}(0)}M,$$

i.e.,

$$\eta_0 + \lambda_0 = 0,$$

or, taking into account the first of equations (12.56),

$$\eta_{t_1} = -\lambda_0 .$$

And at the terminal instant $t = t_1$:

$$(\eta_{t_1}, \lambda_{t_1}) \perp T_{(\tilde{x}(t_1), \tilde{q}(t_1))}N,$$

that is,

$$(-\lambda_0, \lambda_{t_1}) \perp T_{(\tilde{q}(0), \tilde{q}(t_1))}N,$$

which is the required transversality condition (12.53). □

Remarks. (1) Needless to say, if the terminal time t_1 is free, then one should add to statements of Theorem 12.5 the additional equality $h^\nu_{\tilde{u}(t)}(\lambda_t) \equiv 0$.

(2) Pontryagin Maximum Principle withstands further generalizations to wider classes of cost functionals and boundary conditions. After certain modifications of argument, the general scheme provides necessary optimality conditions for more general problems.

Chapter 13

Examples of optimal control problems

In this chapter we apply Pontryagin Maximum Principle to solve concrete optimal control problems.

13.1 The fastest stop of a train at a station

Consider a train moving on a railway. The problem is to drive the train to a station and stop it there in a minimal time.

Describe position of the train by a coordinate x_1 on the real line; the origin $0 \in \mathbb{R}$ corresponds to the station. Assume that the train moves without friction, and we can control acceleration of the train by applying a force bounded by absolute value. Using rescaling if necessary, we can assume that absolute value of acceleration is bounded by 1.

We obtain the control system

$$\ddot{x}_1 = u, \qquad x_1 \in \mathbb{R}, \quad |u| \leq 1,$$

or, in the standard form,

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u, \end{cases} \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2, \quad |u| \leq 1.$$

The time-optimal control problem is

$$\begin{aligned} & x(0) = x^0, \quad x(t_1) = 0, \\ & t_1 \to \min . \end{aligned}$$

First we verify existence of optimal controls by Filippov's theorem. The set of control parameters $U = [-1, 1]$ is compact, the vector fields in the right-hand side

$$f(x,u) = \begin{pmatrix} x_2 \\ u \end{pmatrix}, \qquad |u| \leq 1,$$

are linear, and the set of admissible velocities at a point

$$f(x,U) = \{f(x,u) \mid |u| \leq 1\}$$

is convex. By Corollary 10.2, the time-optimal control problem has a solution if the origin $0 \in \mathbb{R}^2$ is attainable from the initial point x^0. We will show that any point $x \in \mathbb{R}^2$ can be connected with the origin by an extremal curve.

Now we apply Pontryagin Maximum Principle. Introduce canonical coordinates on the cotangent bundle:

$$M = \mathbb{R}^2,$$
$$T^*M = T^*\mathbb{R}^2 = \mathbb{R}^{2*} \times \mathbb{R}^2 = \left\{ \lambda = (\xi, x) \mid x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \xi = (\xi_1, \xi_2) \right\}.$$

The control-dependent Hamiltonian function of PMP is

$$h_u(\xi, x) = (\xi_1, \xi_2) \begin{pmatrix} x_2 \\ u \end{pmatrix} = \xi_1 x_2 + \xi_2 u,$$

and the corresponding Hamiltonian system has the form

$$\begin{cases} \dot{x} = \dfrac{\partial h_u}{\partial \xi}, \\ \dot{\xi} = -\dfrac{\partial h_u}{\partial x}. \end{cases}$$

In coordinates this system splits into two independent subsystems:

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u, \end{cases} \qquad \begin{cases} \dot{\xi}_1 = 0, \\ \dot{\xi}_2 = -\xi_1. \end{cases} \tag{13.1}$$

By PMP, if a control $\tilde{u}(\cdot)$ is time-optimal, then the Hamiltonian system has a nontrivial solution $(\xi(t), x(t))$, $\xi(t) \not\equiv 0$, such that

$$h_{\tilde{u}(t)}(\xi(t), x(t)) = \max_{|u| \leq 1} h_u(\xi(t), x(t)) \geq 0.$$

From this maximality condition, if $\xi_2(t) \neq 0$, then $\tilde{u}(t) = \operatorname{sgn} \xi_2(t)$. Notice that the maximized Hamiltonian

$$\max_{|u| \leq 1} h_u(\xi, x) = \xi_1 x_2 + |\xi_2|$$

is not smooth. So we cannot apply Proposition 12.1, but we can obtain description of optimal controls directly from Pontryagin Maximum Principle, without preliminary maximization of Hamiltonian.

Since

$$\ddot{\xi}_2 = 0,$$

then ξ_2 is linear:

$$\xi_2(t) = \alpha + \beta t, \qquad \alpha,\ \beta = \text{const},$$

hence the optimal control has the form

$$\tilde{u}(t) = \operatorname{sgn}(\alpha + \beta t).$$

So $\tilde{u}(t)$ is piecewise constant, takes only the extremal values ± 1, and has not more than one switching (discontinuity point).

New we find all trajectories $x(t)$ that correspond to such controls and come to the origin. For controls $u = \pm 1$, the first of subsystems (13.1) reads

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = \pm 1. \end{cases}$$

Trajectories of this system satisfy the equation

$$\frac{d\,x_1}{d\,x_2} = \pm x_2,$$

thus are parabolas of the form

$$x_1 = \pm\frac{x_2^2}{2} + C, \qquad C = \text{const}\,.$$

First we find trajectories from this family that come to the origin without switchings: these are two semiparabolas

$$x_1 = \frac{x_2^2}{2}, \qquad x_2 < 0, \quad \dot{x}_2 > 0, \tag{13.2}$$

and

$$x_1 = -\frac{x_2^2}{2}, \qquad x_2 > 0, \quad \dot{x}_2 < 0, \tag{13.3}$$

for $u = +1$ and -1 respectively.

Now we find all extremal trajectories with one switching. Let $(x_{1s}, x_{2s}) \in \mathbb{R}^2$ be a switching point for anyone of curves (13.2), (13.3). Then extremal trajectories with one switching coming to the origin have the form

$$x_1 = \begin{cases} -x_2^2/2 + x_{2s}^2/2 + x_{1s}, & x_2 > x_{2s}, \quad \dot{x}_2 < 0, \\ x_2^2/2 & 0 > x_2 > x_{2s}, \quad \dot{x}_2 > 0, \end{cases} \tag{13.4}$$

and

$$x_1 = \begin{cases} x_2^2/2 - x_{2s}^2/2 + x_{1s}, & x_2 < x_{2s}, \quad \dot{x}_2 > 0, \\ -x_2^2/2 & 0 < x_2 < x_{2s}, \quad \dot{x}_2 < 0. \end{cases} \tag{13.5}$$

It is easy to see that through any point (x_1, x_2) of the plane passes exactly one curve of the forms (13.2)–(13.5). So for any point of the plane there exists exactly one extremal trajectory steering this point to the origin. Since optimal trajectories exist, then the solutions found are optimal.

13.2 Control of a linear oscillator

Consider a linear oscillator whose motion can be controlled by force bounded in absolute value. The corresponding control system (after appropriate rescaling) is

$$\ddot{x}_1 + x_1 = u, \qquad |u| \leq 1, \quad x_1 \in \mathbb{R},$$

or, in the canonical form:

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -x_1 + u, \end{cases} \qquad |u| \leq 1, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2.$$

We consider the time-optimal problem for this system.

By Filippov's theorem, optimal control exists. Similarly to the previous problem, we apply Pontryagin Maximum Principle: the Hamiltonian function is

$$h_u(\xi, x) = \xi_1 x_2 - \xi_2 x_1 + \xi_2 u, \qquad (\xi, x) \in T^*\mathbb{R}^2 = \mathbb{R}^{2*} \times \mathbb{R}^2,$$

and the Hamiltonian system reads

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -x_1 + u, \end{cases} \qquad \begin{cases} \dot{\xi}_1 = \xi_2, \\ \dot{\xi}_2 = -\xi_1. \end{cases}$$

The maximality condition of PMP yields

$$\xi_2(t)\tilde{u}(t) = \max_{|u| \leq 1} \xi_2(t) u,$$

thus optimal controls satisfy the condition

$$\tilde{u}(t) = \operatorname{sgn} \xi_2(t) \qquad \text{if } \xi_2(t) \neq 0.$$

For the variable ξ_2 we have the ODE

$$\ddot{\xi}_2 = -\xi_2,$$

hence

$$\xi_2 = \alpha \sin(t + \beta), \qquad \alpha, \ \beta = \text{const}.$$

Notice that $\alpha \neq 0$: indeed, if $\xi_2 \equiv 0$, then $\xi_1 = -\dot{\xi}_2(t) \equiv 0$, thus $\xi(t) = (\xi_1(t), \xi_2(t)) \equiv 0$, which is impossible by PMP. Consequently,

$$\tilde{u}(t) = \operatorname{sgn}(\alpha \sin(t + \beta)).$$

This equality yields a complete description of possible structure of optimal control. The interval between successive switching points of $\tilde{u}(t)$ has the length π. Let $\tau \in [0, \pi)$ be the first switching point of $\tilde{u}(t)$. Then

$$\tilde{u}(t) = \begin{cases} \operatorname{sgn} \tilde{u}(0), & t \in [0, \tau) \cup [\tau + \pi, \tau + 2\pi) \cup [\tau + 3\pi, \tau + 4\pi) \cup \dots \\ -\operatorname{sgn} \tilde{u}(0), & t \in [\tau, \tau + \pi) \cup [\tau + 2\pi, \tau + 3\pi) \cup \dots \end{cases}$$

That is, $\tilde{u}(t)$ is parametrized by two numbers: the first switching time $\tau \in [0, \pi)$ and the initial sign $\operatorname{sgn} \tilde{u}(0) \in \{\pm 1\}$.

Optimal control $\tilde{u}(t)$ takes only the extremal values ± 1. Thus optimal trajectories $(x_1(t), x_2(t))$ consist of pieces that satisfy the system

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -x_1 \pm 1, \end{cases} \tag{13.6}$$

i.e., arcs of the circles

$$(x_1 \pm 1)^2 + x_2^2 = C, \qquad C = \text{const},$$

passed clockwise.

Now we describe all optimal trajectories coming to the origin. Let γ be any such trajectory. If γ has no switchings, then it is an arc belonging to one of the semicircles

$$(x_1 - 1)^2 + x_2^2 = 1, \qquad x_2 \le 0, \tag{13.7}$$

$$(x_1 + 1)^2 + x_2^2 = 1, \qquad x_2 \ge 0 \tag{13.8}$$

and containing the origin. If γ has switchings, then the last switching can occur at any point of these semicircles except the origin. Assume that γ has the last switching on semicircle (13.7). Then the part of γ before the last switching and after the next to last switching is a semicircle of the circle $(x_1 + 1)^2 + x_2^2 = C$ passing through the last switching point. The next to last switching of γ occurs on the curve obtained by rotation of semicircle (13.7) around the point $(-1, 0)$ in the plane (x_1, x_2) by the angle π, i.e., on the semicircle

$$(x_1 + 3)^2 + x_2^2 = 1, \qquad x_2 \ge 0. \tag{13.9}$$

To obtain the geometric locus of the previous switching of γ, we have to rotate semicircle (13.9) around the point $(1, 0)$ by the angle π; we come to the semicircle

$$(x_1 - 5)^2 + x_2^2 = 1, \qquad x_2 \le 0.$$

The previous switching of γ takes place on the semicircle

$$(x_1 + 7)^2 + x_2^2 = 1, \qquad x_2 \ge 0,$$

and so on.

The case when the last switching of γ occurs on semicircle (13.8) is obtained from the case just considered by the central symmetry of the plane (x_1, x_2)

w.r.t. the origin: $(x_1, x_2) \mapsto (-x_1, -x_2)$. Then the successive switchings of γ (in the reverse order starting from the end) occur on the semicircles

$$\begin{aligned} &(x_1+1)^2 + x_2^2 = 1, && x_2 \geq 0, \\ &(x_1-3)^2 + x_2^2 = 1, && x_2 \leq 0, \\ &(x_1+5)^2 + x_2^2 = 1, && x_2 \geq 0, \\ &(x_1-7)^2 + x_2^2 = 1, && x_2 \leq 0, \end{aligned}$$

etc. We obtained the switching curve in the plane (x_1, x_2):

$$\begin{aligned} &(x_1 - (2k-1))^2 + x_2^2 = 1, && x_2 \leq 0, \quad k \in \mathbb{N}, \\ &(x_1 + (2k-1))^2 + x_2^2 = 1, && x_2 \geq 0, \quad k \in \mathbb{N}. \end{aligned} \tag{13.10}$$

This switching curve divides the plane (x_1, x_2) into two parts. Any extremal trajectory $(x_1(t), x_2(t))$ in the upper part of the plane is a solution of ODE (13.6) with -1 in the second equation, and in the lower part it is a solution of (13.6) with $+1$. For any point of the plane (x_1, x_2) there exists exactly one curve of this family of extremal trajectories that comes to the origin (it has the form of a "spiral" with a finite number of switchings). Since optimal trajectories exist, the constructed extremal trajectories are optimal.

The time-optimal control problem is solved: in the part of the plane (x_1, x_2) over the switching curve (13.10) the optimal control is $\tilde{u} = -1$, and below this curve $\tilde{u} = +1$. Through any point of the plane passes one optimal trajectory which corresponds to this optimal control rule. After finite number of switchings, any optimal trajectory comes to the origin.

Now we consider optimal control problems with the same dynamics as in the previous two sections, but with another cost functional.

13.3 The cheapest stop of a train

As in Section 13.1, we control motion of a train. Now the goal is to stop the train at a fixed instant of time with a minimum expenditure of energy, which is assumed proportional to the integral of squared acceleration.

So the optimal control problem is as follows:

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u, \end{cases} \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2, \quad u \in \mathbb{R},$$

$$x(0) = x^0, \quad x(t_1) = 0, \quad t_1 \text{ fixed},$$

$$\frac{1}{2}\int_0^{t_1} u^2 \, dt \to \min .$$

Filippov's theorem cannot be applied directly since the right-hand side of the control system is not compact. Although, one can choose a new time $t \mapsto$

$\frac{1}{2}\int_0^t u^2(\tau)\,d\tau + C$ and obtain a bounded right-hand side, then compactify it and apply Filippov's theorem. In such a way existence of optimal control can be proved. See also the general theory of linear quadratic problems below in Chapter 16.

To find optimal control, we apply PMP. The Hamiltonian function is

$$h_u^\nu(\xi, x) = \xi_1 x_2 + \xi_2 u + \frac{\nu}{2}u^2, \qquad (\xi, x) \in \mathbb{R}^{2*} \times \mathbb{R}^2.$$

Along optimal trajectories

$$\nu \le 0, \quad \nu = \text{const}\,.$$

From the Hamiltonian system of PMP, we have

$$\begin{cases} \dot{\xi}_1 = 0, \\ \dot{\xi}_2 = -\xi_1. \end{cases} \tag{13.11}$$

Consider first the case of abnormal extremals:

$$\nu = 0.$$

The triple (ξ_1, ξ_2, ν) must be nonzero, thus

$$\xi_2(t) \not\equiv 0.$$

But the maximality condition of PMP yields

$$\tilde{u}(t)\xi_2(t) = \max_{u \in \mathbb{R}} u\,\xi_2(t). \tag{13.12}$$

Since $\xi_2(t) \neq 0$, the maximum above does not exist. Consequently, there are no abnormal extremals.

Consider the normal case: $\nu \neq 0$, we can take $\nu = -1$. The normal Hamiltonian function is

$$h_u(\xi, x) = h_u^{-1}(\xi, x) = \xi_1 x_2 + \xi_2 u - \frac{1}{2}u^2.$$

Maximality condition of PMP is equivalent to $\frac{\partial h_u}{\partial u} = 0$, thus

$$\tilde{u}(t) = \xi_2(t)$$

along optimal trajectories. Taking into account system (13.11), we conclude that optimal control is linear:

$$\tilde{u}(t) = \alpha t + \beta, \qquad \alpha,\ \beta = \text{const}\,.$$

The maximized Hamiltonian function

$$H(\xi, x) = \max_u h_u(\xi, x) = \xi_1 x_2 + \frac{1}{2}\xi_2^2$$

is smooth. That is why optimal trajectories satisfy the Hamiltonian system

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = \xi_2, \\ \dot{\xi}_1 = 0, \\ \dot{\xi}_2 = -\xi_1. \end{cases}$$

For the variable x_1 we obtain the boundary value problem

$$x_1^{(4)} = 0,$$
$$x_1(0) = x_1^0, \quad \dot{x}_1(0) = x_2^0, \quad x_1(t_1) = 0, \quad \dot{x}_1(t_1) = 0. \tag{13.13}$$

For any (x_1^0, x_2^0), there exists exactly one solution $x_1(t)$ of this problem — a cubic spline. The function $x_2(t)$ is found from the equation $x_2 = \dot{x}_1$.

So through any initial point $x^0 \in \mathbb{R}^2$ passes a unique extremal trajectory arriving at the origin. It is a curve $(x_1(t), x_2(t))$, $t \in [0, t_1]$, where $x_1(t)$ is a cubic polynomial that satisfies the boundary conditions (13.13), and $x_2(t) = \dot{x}_1(t)$. In view of existence, this is an optimal trajectory.

13.4 Control of a linear oscillator with cost

We control a linear oscillator, say a pendulum with a small amplitude, by an unbounded force u, but take into account expenditure of energy measured by the integral $\frac{1}{2}\int_0^{t_1} u^2(t)\,dt$. The optimal control problem reads

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -x_1 + u, \end{cases} \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2, \quad u \in \mathbb{R},$$

$$x(0) = x^0, \quad x(t_1) = 0, \quad t_1 \text{ fixed},$$

$$\frac{1}{2}\int_0^{t_1} u^2\,dt \to \min.$$

Existence of optimal control can be proved by the same argument as in the previous section.

The Hamiltonian function of PMP is

$$h_u^\nu(\xi, x) = \xi_1 x_2 - \xi_2 x_1 + \xi_2 u + \frac{\nu}{2}u^2.$$

The corresponding Hamiltonian system yields

$$\begin{cases} \dot{\xi}_1 = \xi_2, \\ \dot{\xi}_2 = -\xi_1. \end{cases}$$

In the same way as in the previous problem, we show that there are no abnormal extremals, thus we can assume $\nu = -1$. Then the maximality condition yields

$$\tilde{u}(t) = \xi_2(t).$$

In particular, optimal control is a harmonic:

$$\tilde{u}(t) = \alpha \sin(t + \beta), \qquad \alpha,\ \beta = \text{const}\,.$$

The system of ODEs for extremal trajectories

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -x_1 + \alpha \sin(t + \beta) \end{cases}$$

is solved explicitly:

$$\begin{aligned} x_1(t) &= -\frac{\alpha}{2} t \cos(t + \beta) + a \sin(t + b), \\ x_2(t) &= \frac{\alpha}{2} t \sin(t + \beta) - \frac{\alpha}{2} \cos(t + \beta) + a \cos(t + b), \quad a,\ b \in \mathbb{R}. \end{aligned} \tag{13.14}$$

Exercise 13.1. Show that exactly one extremal trajectory of the form (13.14) satisfies the boundary conditions.

In view of existence, these extremal trajectories are optimal.

13.5 Dubins car

Consider a car moving in the plane. The car can move forward with a fixed linear velocity and simultaneously rotate with a bounded angular velocity. Given initial and terminal position and orientation of the car in the plane, the problem is to drive the car from the initial configuration to the terminal one for a minimal time.

Admissible paths of the car are curves with bounded curvature. Suppose that curves are parametrized by length, then our problem can be stated geometrically. Given two points in the plane and two unit velocity vectors attached respectively at these points, one has to find a curve in the plane that starts at the first point with the first velocity vector and comes to the second point with the second velocity vector, has curvature bounded by a given constant, and has the minimal length among all such curves.

Remark. If curvature is unbounded, then the problem, in general, has no solutions. Indeed, the infimum of lengths of all curves that satisfy the boundary conditions without bound on curvature is the distance between the initial and terminal points: the segment of the straight line through these points can be approximated by smooth curves with the required boundary conditions. But this infimum is not attained when the boundary velocity vectors do not lie on the line through the boundary points and are not collinear one to another.

After rescaling, we obtain a time-optimal problem for a nonlinear system:

$$\begin{cases} \dot{x}_1 = \cos\theta, \\ \dot{x}_2 = \sin\theta, \\ \dot{\theta} = u, \end{cases} \tag{13.15}$$

$$x = (x_1, x_2) \in \mathbb{R}^2, \quad \theta \in S^1, \quad |u| \leq 1,$$
$$x(0),\ \theta(0),\ x(t_1),\ \theta(t_1) \text{ fixed},$$
$$t_1 \to \min.$$

Existence of solutions is guaranteed by Filippov's Theorem. We apply Pontryagin Maximum Principle.

We have $(x_1, x_2, \theta) \in M = \mathbb{R}^2_x \times S^1_\theta$, let (ξ_1, ξ_2, μ) be the corresponding coordinates of the adjoint vector. Then

$$\lambda = (x, \theta, \xi, \mu) \in T^*M,$$

and the control-dependent Hamiltonian is

$$h_u(\lambda) = \xi_1 \cos\theta + \xi_2 \sin\theta + \mu u.$$

The Hamiltonian system of PMP yields

$$\dot{\xi} = 0, \tag{13.16}$$
$$\dot{\mu} = \xi_1 \sin\theta - \xi_2 \cos\theta, \tag{13.17}$$

and the maximality condition reads

$$\mu(t)u(t) = \max_{|u|\leq 1} \mu(t)u. \tag{13.18}$$

Equation (13.16) means that ξ is constant along optimal trajectories, thus the right-hand side of (13.17) can be rewritten as

$$\xi_1 \sin\theta - \xi_2 \cos\theta = \alpha \sin(\theta + \beta), \qquad \alpha,\ \beta = \text{const}, \quad \alpha = \sqrt{\xi_1^2 + \xi_2^2} \geq 0. \tag{13.19}$$

So the Hamiltonian system of PMP (13.15)–(13.17) yields the following system:

$$\begin{cases} \dot{\mu} = \alpha \sin(\theta + \beta), \\ \dot{\theta} = u. \end{cases}$$

Maximality condition (13.18) implies that

$$u(t) = \operatorname{sgn} \mu(t) \quad \text{if } \mu(t) \neq 0. \tag{13.20}$$

If $\alpha = 0$, then $(\xi_1, \xi_2) \equiv 0$ and $\mu = \text{const} \neq 0$, thus $u = \text{const} = \pm 1$. So the curve $x(t)$ is an arc of a circle of radius 1.

Let $\alpha \neq 0$, then in view of (13.19), we have $\alpha > 0$. Conditions (13.16), (13.17), (13.18) are preserved if the adjoint vector (ξ, μ) is multiplied by any positive constant. Thus we can choose (ξ, μ) such that $\alpha = \sqrt{\xi_1^2 + \xi_2^2} = 1$. That is why we suppose in the sequel that

$$\alpha = 1.$$

Condition (13.20) means that behavior of sign of the function $\mu(t)$ is crucial for the structure of optimal control. We consider several possibilities for $\mu(t)$.

(0) If the function $\mu(t)$ does not vanish on the segment $[0, t_1]$, then the optimal control is constant:

$$u(t) = \text{const} = \pm 1, \qquad t \in [0, t_1], \tag{13.21}$$

and the optimal trajectory $x(t)$, $t \in [0, t_1]$, is an arc of a circle. Notice that an optimal trajectory cannot contain a full circle: a circle can be eliminated so that the resulting trajectory satisfy the same boundary conditions and is shorter. Thus controls (13.21) can be optimal only if $t_1 < 2\pi$.

In the sequel we can assume that the set

$$N = \{\tau \in [0, t_1] \mid \mu(\tau) = 0\}$$

is nonempty. Since N is open, it is a union of open intervals in $[0, t_1]$, plus, may be, semiopen intervals of the form $[0, \tau_1)$, $(\tau_2, t_1]$.

(1) Suppose that the set N contains an interval of the form

$$(\tau_1, \tau_2) \subset [0, t_1], \qquad \tau_1 < \tau_2. \tag{13.22}$$

We can assume that the interval (τ_1, τ_2) is maximal w.r.t. inclusion:

$$\mu(\tau_1) = \mu(\tau_2) = 0, \qquad \mu|_{(\tau_1, \tau_2)} \neq 0.$$

From PMP we have the inequality

$$h_{u(t)}(\lambda(t)) = \cos(\theta(t) + \beta) + \mu(t) u(t) \geq 0.$$

Thus

$$\cos(\theta(\tau_1) + \beta) \geq 0.$$

This inequality means that the angle

$$\widehat{\theta} = \theta(\tau_1) + \beta$$

satisfies the inclusion

$$\widehat{\theta} \in \left[0, \frac{\pi}{2}\right] \cup \left[\frac{3\pi}{2}, 2\pi\right).$$

Consider first the case

$$\widehat{\theta} \in \left(0, \frac{\pi}{2}\right].$$

Then $\dot{\mu}(\tau_1) = \sin \widehat{\theta} > 0$, thus at τ_1 control switches from -1 to $+1$, so

$$\dot{\theta}(t) = u(t) \equiv 1, \qquad t \in (\tau_1, \tau_2).$$

We evaluate the distance $\tau_2 - \tau_1$. Since

$$\mu(\tau_2) = \int_{\tau_1}^{\tau_2} \sin(\widehat{\theta} + \tau - \tau_1)\, d\tau = 0,$$

then $\tau_2 - \tau_1 = 2(\pi - \widehat{\theta})$, thus

$$\tau_2 - \tau_1 \in [\pi, 2\pi). \tag{13.23}$$

In the case

$$\widehat{\theta} \in \left[\frac{3\pi}{2}, 2\pi\right)$$

inclusion (13.23) is proved similarly, and in the case $\widehat{\theta} = 0$ we obtain no optimal controls (the curve $x(t)$ contains a full circle, which can be eliminated).

Inclusion (13.23) means that successive roots τ_1, τ_2 of the function $\mu(t)$ cannot be arbitrarily close one to another. Moreover, the previous argument shows that at such instants τ_i optimal control switches from one extremal value to another, and along any optimal trajectory the distance between any successive switchings τ_i, τ_{i+1} is the same.

So in case (1) an optimal control can only have the form

$$u(t) = \begin{cases} \varepsilon, & t \in (\tau_{2k-1}, \tau_{2k}), \\ -\varepsilon, & t \in (\tau_{2k}, \tau_{2k+1}), \end{cases} \tag{13.24}$$

$$\varepsilon = \pm 1,$$

$$\tau_{i+1} - \tau_i = \text{const} \in [\pi, 2\pi), \qquad i = 1, \dots, N-1, \tag{13.25}$$

$$\tau_1 \in (0, 2\pi),$$

here we do not indicate values of u in the intervals before the first switching, $t \in (0, \tau_1)$, and after the last switching, $t \in (\tau_N, t_1)$. For such trajectories, control takes only extremal values ± 1 and the number of switchings is finite on any compact time segment. Such a control is called *bang-bang*.

Controls $u(t)$ given by (13.24), (13.25) satisfy PMP for arbitrarily large t, but they are not optimal if the number of switchings is $N > 3$. Indeed, suppose that such a control has at least 4 switchings. Then the piece of trajectory $x(t)$, $t \in [\tau_1, \tau_4]$, is a concatenation of three arcs of circles corresponding to the segments of time $[\tau_1, \tau_2]$, $[\tau_2, \tau_3]$, $[\tau_3, \tau_4]$ with

$$\tau_4 - \tau_3 = \tau_3 - \tau_2 = \tau_2 - \tau_1 \in [\pi, 2\pi).$$

Draw the segment of line

$$\tilde{x}(t), \qquad t \in [(\tau_1 + \tau_2)/2, (\tau_3 + \tau_4)/2], \qquad \left|\frac{d\,\tilde{x}}{d\,t}\right| \equiv 1,$$

the common tangent to the first and third circles through the points $x((\tau_1+\tau_2)/2)$ and $x((\tau_3+\tau_4)/2)$. Then the curve

$$y(t) = \begin{cases} x(t), & t \notin [(\tau_1+\tau_2)/2, (\tau_3+\tau_4)/2], \\ \tilde{x}(t), & t \in [(\tau_1+\tau_2)/2, (\tau_3+\tau_4)/2], \end{cases}$$

is an admissible trajectory and shorter than $x(t)$. We proved that optimal bang-bang control can have not more than 3 switchings.

(2) It remains to consider the case where the set N does not contain intervals of the form (13.22). Then N consists of at most two semiopen intervals:

$$N = [0, \tau_1) \cup (\tau_2, t_1], \qquad \tau_1 \le \tau_2,$$

where one or both intervals may be absent. If $\tau_1 = \tau_2$, then the function $\mu(t)$ has a unique root on the segment $[0, t_1]$, and the corresponding optimal control is determined by condition (13.20). Otherwise

$$\tau_1 < \tau_2,$$

and

$$\mu|_{[0,\tau_1)} \ne 0, \quad \mu|_{[\tau_1,\tau_2]} \equiv 0, \quad \mu|_{(\tau_2,t_1]} \ne 0. \tag{13.26}$$

In this case the maximality condition of PMP (13.20) does not determine optimal control $u(t)$ uniquely since the maximum is attained for more than one value of control parameter u. Such a control is called *singular*. Nevertheless, singular controls in this problem can be determined from PMP. Indeed, the following identities hold on the interval (τ_1, τ_2):

$$\dot{\mu} = \sin(\theta+\beta) = 0 \quad \Rightarrow \quad \theta+\beta = \pi k \quad \Rightarrow \quad \theta = \text{const} \quad \Rightarrow \quad u = 0.$$

Consequently, if an optimal trajectory $x(t)$ has a singular piece, which is a line, then τ_1 and τ_2 are the only switching times of the optimal control. Then

$$u|_{(0,\tau_1)} = \text{const} = \pm 1, \quad u|_{(\tau_2,t_1)} = \text{const} = \pm 1,$$

and the whole trajectory $x(t)$, $t \in [0, t_1]$, is a concatenation of an arc of a circle of radius 1

$$x(t), \quad u(t) = \pm 1, \qquad t \in [0, \tau_1],$$

a line

$$x(t), \quad u(t) = 0, \qquad t \in [\tau_1, \tau_2],$$

and one more arc of a circle of radius 1

$$x(t), \quad u(t) = \pm 1, \qquad t \in [\tau_2, t_1].$$

So optimal trajectories in the problem have one of the following two types:

(1) concatenation of a bang-bang piece (arc of a circle, $u = \pm 1$), a singular piece (segment of a line, $u = 0$), and a bang-bang piece, or

(2) concatenation of bang-bang pieces with not more than 3 switchings, the arcs of circles between switchings having the same central angle $\in [\pi, 2\pi)$.

If boundary points $x(0)$, $x(t_1)$ are sufficiently far one from another, then they can be connected only by trajectories containing singular piece. For such boundary points, we obtain a simple algorithm for construction of an optimal trajectory. Through each of the points $x(0)$ and $x(t_1)$, construct a pair of circles of radius 1 tangent respectively to the velocity vectors $\dot{x}(0) = (\cos\theta(0), \sin\theta(0))$ and $\dot{x}(t_1) = (\cos\theta(t_1), \sin\theta(t_1))$. Then draw common tangents to the circles at $x(0)$ and $x(t_1)$ respectively, so that direction of motion along these tangents was compatible with direction of rotation along the circles determined by the boundary velocity vectors $\dot{x}(0)$ and $\dot{x}(t_1)$. Finally, choose the shortest curve among the candidates obtained. This curve is the optimal trajectory.

Chapter 14

Hamiltonian systems with convex Hamiltonians

A well-known theorem states that if a level surface of a Hamiltonian is convex, then it contains a periodic trajectory of the Hamiltonian system [18], [25]. In this chapter we prove a more general statement as an application of optimal control theory for linear systems.

Theorem 14.1. *Let S be a strongly convex compact subset of $\mathbb{R}^n$, n even, and let the boundary of S be a level surface of a Hamiltonian $H \in C^\infty(\mathbb{R}^n)$. Then for any vector $v \in \mathbb{R}^n$ there exists a chord in S parallel to v such that there exists a trajectory of the Hamiltonian system $\dot{x} = \vec{H}(x)$ passing through the endpoints of the chord.*

We assume here that $\mathbb{R}^n$ is endowed with the standard symplectic structure

$$\sigma(x, x) = \langle x, Jx \rangle, \qquad J = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix},$$

i.e., the Hamiltonian vector field corresponding to a Hamiltonian H has the form $\vec{H} = J \operatorname{grad} H$.

The theorem on periodic trajectories of Hamiltonian systems is a particular case of the previous theorem with $v = 0$. Now we prove Th. 14.1.

Proof. Without loss of generality, we can assume that $0 \in \operatorname{int} S$.

Consider the polar of the set S:

$$S^\circ = \{u \in \mathbb{R}^n \mid \sup_{x \in S} \langle u, x \rangle \leq 1\}.$$

It follows from the separation theorem that

$$(S^\circ)^\circ = S, \qquad 0 \in \operatorname{int} S^\circ,$$

and that S° is a strongly convex compact subset of $\mathbb{R}^n$.

Introduce the following linear optimal control problem:

$$\begin{aligned}
&\dot{x} = u, \qquad u \in S^\circ, \quad x \in \mathbb{R}^n, \\
&x(0) = a, \qquad x(1) = b, \\
&\int_0^1 \langle x, Ju \rangle \, dt \to \min .
\end{aligned} \tag{14.1}$$

Here a and b are any points in S° sufficiently close to the origin and such that the vector $J(b-a)$ is parallel to v. By Filippov's theorem, this problem has optimal solutions. We use these solutions in order to construct the required trajectory of the Hamiltonian system on ∂S.

The control-dependent Hamiltonian of PMP has the form:

$$h_u^\nu(p, x) = pu + \nu \langle x, Ju \rangle.$$

We show first that abnormal trajectories cannot be optimal. Let $\nu = 0$. Then the adjoint equation is $\dot{p} = 0$, thus

$$p = p_0 = \text{const} .$$

The maximality condition of PMP reads

$$p_0 u(t) = \max_{v \in S^\circ} p_0 v.$$

Since the polar S° is strictly convex, then

$$u(t) = \text{const}, \qquad u(t) \in \partial S^\circ .$$

Consequently, abnormal trajectories are lines with velocities separated from zero. For points a, b sufficiently close to the origin, abnormal trajectories cannot meet the boundary conditions.

Thus optimal trajectories are normal, so we can set $\nu = -1$. The normal Hamiltonian is

$$h_u(p, x) = pu - \langle x, Ju \rangle,$$

and the corresponding Hamiltonian system reads

$$\begin{cases} \dot{p} = Ju, \\ \dot{x} = u. \end{cases}$$

The normal Hamiltonian can be written as

$$\begin{aligned}
&h_u(p, x) = \langle y, u \rangle, \\
&y = p + Jx,
\end{aligned}$$

where the vector y satisfies the equation

$$\dot{y} = 2Ju.$$

Along a normal trajectory

$$h_{u(t)}(p(t), x(t)) = \langle y(t), u(t)\rangle = \max_{v \in S^\circ} \langle y(t), v\rangle = C = \text{const}\,. \tag{14.2}$$

Consider first the case $C \neq 0$, thus $C > 0$. Then

$$z(t) = \frac{1}{C} y(t) \in (S^\circ)^\circ = S,$$

i.e., $z(t) \in S$. Moreover, $z(t) \in \partial S$ and the vector $u(t)$ is a normal to ∂S at the point $z(t)$. Consequently, the curve $z(t)$ is, up to reparametrization, a trajectory of the Hamiltonian field $\vec{H} = J \operatorname{grad} H$. Compute the boundary conditions:

$$\begin{aligned}
&p(1) - p(0) = J(x(1) - x(0)),\\
&y(1) - y(0) = 2J(x(1) - x(0)) = 2J(b - a),\\
&z(1) - z(0) = \frac{2}{C} J(b - a).
\end{aligned}$$

Thus $z(t)$ is the required trajectory: the chord $z(1) - z(0)$ is parallel to the vector v.

In order to complete the proof, we show now that the case $C = 0$ in (14.2) is impossible. Indeed, if $C = 0$, then $y(t) \equiv 0$, thus $u(t) \equiv 0$. If $a \neq b$, then the boundary conditions for x are not satisfied. And if $a = b$, then the pair $(u(t), x(t)) \equiv (0, 0)$ does not realize minimum of functional (14.1), which can take negative values: for any admissible 1-periodic trajectory $x(t)$, the trajectory $\hat{x}(t) = x(1 - t)$ is periodic with the cost

$$\int_0^1 \langle \hat{x}, J\dot{\hat{x}}\rangle \, dx = -\int_0^1 \langle x, Ju\rangle \, dx.$$

$\square$

Chapter 15

Linear time-optimal problem

15.1 Problem statement

In this chapter we study the following optimal control problem:

$$\begin{aligned}
&\dot{x} = Ax + Bu, \qquad x \in \mathbb{R}^n, \quad u \in U \subset \mathbb{R}^m, \\
&x(0) = x_0, \quad x(t_1) = x_1, \quad x_0,\ x_1 \in \mathbb{R}^n \text{ fixed}, \\
&t_1 \to \min,
\end{aligned} \tag{15.1}$$

where U is a compact convex polytope in $\mathbb{R}^m$, and A and B are constant matrices of order $n \times n$ and $n \times m$ respectively. Such problem is called *linear time-optimal problem*.

The polytope U is the convex hull of a finite number of points $a_1, \dots, a_k$ in $\mathbb{R}^m$:

$$U = \operatorname{conv}\{a_1, \dots, a_k\}.$$

We assume that the points a_i do not belong to the convex hull of all the rest points a_j, $j \neq i$, so that each a_i is a vertex of the polytope U.

In the sequel we assume the following *General Position Condition*:

For any edge $[a_i, a_j]$ of U, the vector $e_{ij} = a_j - a_i$ satisfies the equality

$$\operatorname{span}(Be_{ij}, ABe_{ij}, \dots, A^{n-1}Be_{ij}) = \mathbb{R}^n. \tag{15.2}$$

This condition means that no vector Be_{ij} belongs to a proper invariant subspace of the matrix A. By Theorem 3.1, this is equivalent to controllability of the linear system $\dot{x} = Ax + Bu$ with the set of control parameters $u \in \mathbb{R}e_{ij}$. Condition (15.2) can be achieved by a small perturbation of matrices A, B.

We already considered examples of linear time-optimal problems in Sections 13.1, 13.2. Here we study the structure of optimal control, prove its uniqueness, evaluate the number of switchings.

Existence of optimal control for any points x_0, x_1 such that $x_1 \in \mathcal{A}(x_0)$ is guaranteed by Filippov's theorem. Notice that for the analogous problem with an unbounded set of control parameters, optimal control may not exist: it is easy to show this using linearity of the system.

Before proceeding with the study of linear time-optimal problems, we recall some basic facts on polytopes.

15.2 Geometry of polytopes

The convex hull of a finite number of points $a_1, \ldots, a_k \in \mathbb{R}^m$ is the set

$$U = \operatorname{conv}\{a_1, \ldots, a_k\} \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^{k} \alpha_i a_i \mid \alpha_i \geq 0, \ \sum_{i=1}^{k} \alpha_i = 1 \right\}.$$

An affine hyperplane in $\mathbb{R}^m$ is a set of the form

$$\Pi = \{u \in \mathbb{R}^m \mid \langle \xi, u \rangle = c\}, \qquad \xi \in \mathbb{R}^{m*} \setminus \{0\}, \quad c \in \mathbb{R}.$$

A supporting hyperplane to a polytope U is a hyperplane Π such that

$$\langle \xi, u \rangle \leq c \quad \forall u \in U$$

for the covector ξ and number c that define Π, and this inequality turns into equality at some point $u \in \partial U$, i.e., $\Pi \cap U \neq \emptyset$.

A polytope $U = \operatorname{conv}\{a_1, \ldots, a_k\}$ intersects with any its supporting hyperplane $\Pi = \{u \mid \langle \xi, u \rangle = c\}$ by another polytope:

$$\begin{aligned}
&U \cap \Pi = \operatorname{conv}\{a_{i1}, \ldots, a_{il}\}, \\
&\langle \xi, a_{i1} \rangle = \cdots = \langle \xi, a_{il} \rangle = c, \\
&\langle \xi, a_j \rangle < c, \quad j \notin \{i_1, \ldots, i_l\}.
\end{aligned}$$

Such polytopes $U \cap \Pi$ are called faces of the polytope U. Zero-dimensional and one-dimensional faces are called respectively vertices and edges. A polytope has a finite number of faces, each of which is the convex hull of a finite number of vertices. A face of a face is a face of the initial polytope. Boundary of a polytope is a union of all its faces. This is a straightforward corollary of the separation theorem for convex sets (or the Hahn-Banach Theorem).

15.3 Bang-bang theorem

Optimal control in the linear time-optimal problem is bang-bang, i.e., it is piecewise constant and takes values in vertices of the polytope U.

Theorem 15.1. *Let $u(t)$, $0 \leq t \leq t_1$, be an optimal control in the linear time-optimal control problem* (15.1). *Then there exists a finite subset*

$$\mathcal{T} \subset [0, t_1], \qquad \#\mathcal{T} < \infty,$$

such that

$$u(t) \in \{a_1, \dots, a_k\}, \qquad t \in [0, t_1] \setminus \mathcal{T}, \tag{15.3}$$

and restriction $u(t)|_{t \in [0,t_1] \setminus \mathcal{T}}$ *is locally constant.*

Proof. Apply Pontryagin Maximum Principle to the linear time-optimal problem (15.1). State and adjoint vectors are

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \qquad \xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^{n*},$$

and a point in the cotangent bundle is

$$\lambda = (\xi, x) \in \mathbb{R}^{n*} \times \mathbb{R}^n = T^*\mathbb{R}^n.$$

The control-dependent Hamiltonian is

$$h_u(\xi, x) = \xi A x + \xi B u$$

(we multiply raws by columns). The Hamiltonian system and maximality condition of PMP take the form:

$$\begin{aligned} &\begin{cases} \dot{x} = Ax + Bu, \\ \dot{\xi} = -\xi A, \end{cases} \\ &\xi(t) \neq 0, \\ &\xi(t) B u(t) = \max_{u \in U} \xi(t) B u. \end{aligned} \tag{15.4}$$

The Hamiltonian system implies that adjoint vector

$$\xi(t) = \xi(0) e^{-tA}, \qquad \xi(0) \neq 0, \tag{15.5}$$

is analytic along the optimal trajectory.

Consider the set of indices corresponding to vertices where maximum (15.4) is attained:

$$J(t) = \left\{ 1 \leq j \leq k \mid \xi(t) B a_j = \max_{u \in U} \xi(t) B u = \max\{\xi(t) B a_i \mid i = 1, \dots, k\} \right\}.$$

At each instant t the linear function $\xi(t)B$ attains maximum at vertices of the polytope U. We show that this maximum is attained at one vertex always except a finite number of moments.

Define the set

$$\mathcal{T} = \{t \in [0, t_1] \mid \#J(t) > 1\}.$$

By contradiction, suppose that $\mathcal{T}$ is infinite: there exists a sequence of distinct moments

$$\{\tau_1, \dots, \tau_n, \dots\} \subset \mathcal{T}.$$

Since there is a finite number of choices for the subset $J(\tau_n) \subset \{1, \dots, k\}$, we can assume, without loss of generality, that

$$J(\tau_1) = J(\tau_2) = \cdots = J(\tau_n) = \cdots .$$

Denote $J = J(\tau_i)$.

Further, since the convex hull

$$\operatorname{conv}\{a_j \mid j \in J\}$$

is a face of U, then there exist indices $j_1, j_2 \in J$ such that the segment $[a_{j_1}, a_{j_2}]$ is an edge of U. We have

$$\xi(\tau_i)Ba_{j_1} = \xi(\tau_i)Ba_{j_2}, \qquad i = 1, 2, \dots .$$

For the vector $e = a_{j_2} - a_{j_1}$ we obtain

$$\xi(\tau_i)Be = 0, \qquad i = 1, 2, \dots .$$

But $\xi(\tau_i) = \xi(0)e^{-\tau_i A}$ by (15.5), so the analytic function

$$t \mapsto \xi(0)e^{-tA}Be$$

has an infinite number of zeros on the segment $[0, t_1]$, thus it is identically zero:

$$\xi(0)e^{-tA}Be \equiv 0.$$

We differentiate this identity successively at $t = 0$ and obtain

$$\xi(0)Be = 0, \quad \xi(0)ABe = 0, \quad \dots, \quad \xi(0)A^{n-1}Be = 0.$$

By General Position Condition (15.2), we have $\xi(0) = 0$, a contradiction to (15.5). So the set $\mathcal{T}$ is finite.

Out of the set $\mathcal{T}$, the function $\xi(t)B$ attains maximum on U at one vertex $a_{j(t)}$, $\{j(t)\} = J(t)$, thus the optimal control $u(t)$ takes value in the vertex $a_{j(t)}$. Condition (15.3) follows. Further,

$$\xi(t)Ba_{j(t)} > \xi(t)Ba_i, \qquad i \neq j(t).$$

But all functions $t \mapsto \xi(t)Ba_i$ are continuous, so the preceding inequality preserves for instants close to t. The function $t \mapsto j(t)$ is locally constant on $[0, t_1] \setminus \mathcal{T}$, thus the optimal control $u(t)$ is also locally constant on $[0, t_1] \setminus \mathcal{T}$. □

In the sequel we will need the following statement proved in the preceding argument.

Corollary 15.1. *Let $\xi(t)$, $t \in [0, t_1]$, be a nonzero solution of the adjoint equation $\dot{\xi} = -\xi A$. Then everywhere in the segment $[0, t_1]$, except a finite number of points, there exists a unique control $u(t) \in U$ such that $\xi(t)Bu(t) = \max_{u \in U} \xi(t)Bu$.*

15.4 Uniqueness of optimal controls and extremals

Theorem 15.2. *Let the terminal point x_1 be reachable from the initial point x_0:*

$$x_1 \in \mathcal{A}(x_0).$$

Then linear time-optimal control problem (15.1) *has a unique solution.*

Proof. As we already noticed, existence of an optimal control follows from Filippov's Theorem.

Suppose that there exist two optimal controls: $u_1(t)$, $u_2(t)$, $t \in [0, t_1]$. By Cauchy's formula:

$$x(t_1) = e^{t_1 A}\left(x_0 + \int_0^{t_1} e^{-tA} Bu(t)\,dt\right),$$

we obtain

$$e^{t_1 A}\left(x_0 + \int_0^{t_1} e^{-tA} Bu_1(t)\,dt\right) = e^{t_1 A}\left(x_0 + \int_0^{t_1} e^{-tA} Bu_2(t)\,dt\right),$$

thus

$$\int_0^{t_1} e^{-tA} Bu_1(t)\,dt = \int_0^{t_1} e^{-tA} Bu_2(t)\,dt. \tag{15.6}$$

Let $\xi_1(t) = \xi_1(0)e^{-tA}$ be the adjoint vector corresponding by PMP to the control $u_1(t)$. Then equality (15.6) can be written in the form

$$\int_0^{t_1} \xi_1(t) Bu_1(t)\,dt = \int_0^{t_1} \xi_1(t) Bu_2(t)\,dt. \tag{15.7}$$

By the maximality condition of PMP

$$\xi_1(t)Bu_1(t) = \max_{u \in U} \xi_1(t)Bu,$$

thus

$$\xi_1(t)Bu_1(t) \geq \xi_1(t)Bu_2(t).$$

But this inequality together with equality (15.7) implies that almost everywhere on $[0, t_1]$

$$\xi_1(t)Bu_1(t) = \xi_1(t)Bu_2(t).$$

By Corollary 15.1,

$$u_1(t) \equiv u_2(t)$$

almost everywhere on $[0, t_1]$. □

So for linear time-optimal problem, optimal control is unique. The standard procedure to find the optimal control for a given pair of boundary points x_0, x_1 is to find all extremals $(\xi(t), x(t))$ steering x_0 to x_1 and then to seek for the best among them. In the examples considered in Sections 13.1, 13.2, there was one extremal for each pair x_0, x_1 with $x_1 = 0$. We prove now that this is a general property of linear time-optimal problems.

Theorem 15.3. *Let $x_1 = 0 \in \mathcal{A}(x_0)$ and $0 \in U \setminus \{a_1, \dots, a_k\}$. Then there exists a unique control $u(t)$ that steers x_0 to 0 and satisfies Pontryagin Maximum Principle.*

Proof. Assume that there exist two controls

$$u_1(t), \quad t \in [0, t_1], \qquad \text{and} \qquad u_2(t), \quad t \in [0, t_2],$$

that steer x_0 to 0 and satisfy PMP.

If $t_1 = t_2$, then the argument of the proof of preceding theorem shows that $u_1(t) \equiv u_2(t)$ a.e., so we can assume that

$$t_1 > t_2.$$

Cauchy's formula gives

$$e^{t_1 A}\left(x_0 + \int_0^{t_1} e^{-tA} B u_1(t)\, dt\right) = 0,$$
$$e^{t_2 A}\left(x_0 + \int_0^{t_2} e^{-tA} B u_2(t)\, dt\right) = 0,$$

thus

$$\int_0^{t_1} e^{-tA} B u_1(t)\, dt = \int_0^{t_2} e^{-tA} B u_2(t)\, dt. \tag{15.8}$$

According to PMP, there exists an adjoint vector $\xi_1(t)$, $t \in [0, t_1]$, such that

$$\xi_1(t) = \xi_1(0) e^{-tA}, \qquad \xi_1(0) \neq 0, \tag{15.9}$$
$$\xi_1(t) B u_1(t) = \max_{u \in U} \xi_1(t) B u. \tag{15.10}$$

Since $0 \in U$, then

$$\xi_1(t) B u_1(t) \geq 0, \qquad t \in [0, t_1]. \tag{15.11}$$

Equality (15.8) can be rewritten as

$$\int_0^{t_1} \xi_1(t) B u_1(t)\, dt = \int_0^{t_2} \xi_1(t) B u_2(t)\, dt. \tag{15.12}$$

Taking into account inequality (15.11), we obtain

$$\int_0^{t_2} \xi_1(t) B u_1(t)\, dt \leq \int_0^{t_2} \xi_1(t) B u_2(t)\, dt. \tag{15.13}$$

But maximality condition (15.10) implies that

$$\xi_1(t)Bu_1(t) \geq \xi_1(t)Bu_2(t), \qquad t \in [0, t_2]. \tag{15.14}$$

Now inequalities (15.13) and (15.14) are compatible only if

$$\xi_1(t)Bu_1(t) = \xi_1(t)Bu_2(t), \qquad t \in [0, t_2],$$

thus inequality (15.13) should turn into equality. In view of (15.12), we have

$$\int_{t_1}^{t_2} \xi_1(t)Bu_1(t)\, dt = 0.$$

Since the integrand is nonnegative, see (15.11), then it vanishes identically:

$$\xi_1(t)Bu_1(t) \equiv 0, \qquad t \in [t_1, t_2].$$

By the argument of Theorem 15.1, the control $u_1(t)$ is bang-bang, so there exists an interval $I \subset [t_1, t_2]$ such that

$$u_1(t)|_I \equiv a_j \neq 0.$$

Thus

$$\xi_1(t)Ba_j \equiv 0, \qquad t \in I.$$

But $\xi_1(t)0 \equiv 0$, this is a contradiction with uniqueness of the control for which maximum in PMP is obtained, see Corollary 15.1. $\square$

15.5 Switchings of optimal control

Now we evaluate the number of switchings of optimal control in linear time-optimal problems. In the examples of Sections 13.1, 13.2 we had respectively one switching and an arbitrarily large number of switchings, although finite on any segment. It turns out that in general there are two cases: non-oscillating and oscillating, depending on whether the matrix A of the control system has real spectrum or not. Recall that in the example with one switching, Section 13.1, we had

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathrm{Sp}(A) = \{0\} \subset \mathbb{R},$$

and in the example with arbitrarily large number of switchings, Section 13.2,

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \mathrm{Sp}(A) = \{\pm i\} \not\subset \mathbb{R}.$$

We consider systems with scalar control:

$$\dot{x} = Ax + ub, \qquad u \in U = [\alpha, \beta] \subset \mathbb{R}, \quad x \in \mathbb{R}^n,$$

under the General Position Condition

$$\operatorname{span}(b, Ab, \dots, A^{n-1}b) = \mathbb{R}^n.$$

Then attainable set of the system is full-dimensional for arbitrarily small times. We can evaluate the minimal number of switchings necessary to fill a full-dimensional domain. Optimal control is piecewise constant with values in $\{\alpha, \beta\}$. Assume that we start from the initial point x_0 with the control α. Without switchings we fill a piece of a 1-dimensional curve $e^{(Ax+\alpha b)t}x_0$, with 1 switching we fill a piece of a 2-dimensional surface $e^{(Ax+\beta b)t_2} \circ e^{(Ax+\alpha b)t_1}x_0$, with 2 switchings we can attain points in a 3-dimensional surface, etc. So the minimal number of switchings required to reach an n-dimensional domain is $n-1$.

We prove now that in the non-oscillating case we never need more than $n-1$ switchings of optimal control.

Theorem 15.4. *Assume that the matrix A has only real eigenvalues:*

$$\operatorname{Sp}(A) \subset \mathbb{R}.$$

Then any optimal control in linear time-optimal problem (15.1) *has no more than $n-1$ switchings.*

Proof. Let $u(t)$ be an optimal control and $\xi(t) = \xi(0)e^{-tA}$ the corresponding solution of the adjoint equation $\dot{\xi} = -\xi A$. The maximality condition of PMP reads

$$\xi(t)bu(t) = \max_{u \in [\alpha,\beta]} \xi(t)bu,$$

thus

$$u(t) = \begin{cases} \beta & \text{if } \xi(t)b > 0, \\ \alpha & \text{if } \xi(t)b < 0. \end{cases}$$

So the number of switchings of the control $u(t)$, $t \in [0, t_1]$, is equal to the number of changes of sign of the function

$$y(t) = \xi(t)b, \qquad t \in [0, t_1].$$

We show that $y(t)$ has not more than $n-1$ real roots.

Derivatives of the adjoint vector have the form

$$\xi^{(k)}(t) = \xi(0)e^{-tA}(-A)^k.$$

By Cayley Theorem, the matrix A satisfies its characteristic equation:

$$A^n + c_1 A^{n-1} + \dots + c_n \operatorname{Id} = 0,$$

where

$$\det(t \operatorname{Id} - A) = t^n + c_1 t^{n-1} + \dots + c_n,$$

thus

$$(-A)^n - c_1(-A)^{n-1} + \cdots + (-1)^n c_n \,\mathrm{Id} = 0.$$

Then the function $y(t)$ satisfies an n-th order ODE:

$$y^{(n)}(t) - c_1 y^{(n-1)}(t) + \cdots + (-1)^n c_n y(t) = 0. \tag{15.15}$$

It is well known (see e.g. [5]) that any solution of this equation is a quasipolynomial:

$$\begin{aligned} &y(t) = \sum_{i=1}^{k} e^{-\lambda_i t} P_i(t), \\ &P_i(t) \text{ a polynomial}, \\ &\lambda_i \neq \lambda_j \text{ for } i \neq j, \end{aligned}$$

where λ_i are eigenvalues of the matrix A and degree of each polynomial P_i is less than multiplicity of the corresponding eigenvalue λ_i, thus

$$\sum_{i=1}^{k} \deg P_i \leq n - k.$$

Now the statement of this theorem follows from the next general lemma. □

Lemma 15.1. *A quasipolynomial*

$$\begin{aligned} &y(t) = \sum_{i=1}^{k} e^{\lambda_i t} P_i(t), \qquad \sum_{i=1}^{k} \deg P_i \leq n - k, \\ &\lambda_i \neq \lambda_j \textit{ for } i \neq j, \end{aligned} \tag{15.16}$$

has no more than $n-1$ real roots.

Proof. Apply induction on k.

If $k = 1$, then a quasipolynomial

$$y(t) = e^{\lambda t} P(t), \qquad \deg P \leq n - 1,$$

has no more than $n - 1$ roots.

We prove the induction step for $k > 1$. Denote

$$n_i = \deg P_i, \qquad i = 1, \ldots, k.$$

Suppose that the quasipolynomial $y(t)$ has n real roots. Rewrite the equation

$$y(t) = \sum_{i=1}^{k-1} e^{\lambda_i t} P_i(t) + e^{\lambda_k t} P_k(t) = 0$$

as follows:

$$\sum_{i=1}^{k-1} e^{(\lambda_i-\lambda_k)t} P_i(t) + P_k(t) = 0. \tag{15.17}$$

The quasipolynomial in the left-hand side has n roots. We differentiate this quasipolynomial successively (n_k+1) times so that the polynomial $P_k(t)$ disappear. After (n_k+1) differentiations we obtain a quasipolynomial

$$\sum_{i=1}^{k-1} e^{(\lambda_i-\lambda_k)t} Q_i(t), \qquad \deg Q_i \leq \deg P_i,$$

which has $(n-n_k-1)$ real roots by Rolle's Theorem. But by induction assumption the maximal possible number of real roots of this quasipolynomial is

$$\sum_{i=1}^{k-1} n_i + k - 2 < n - n_k - 1.$$

The contradiction finishes the proof of the lemma. □

So we completed the proof of Theorem 15.4: in the non-oscillating case an optimal control has no more than $n-1$ switchings on the whole domain (recall that $n-1$ switchings are always necessary even on short time segments since the attainable sets $\mathcal{A}_{q_0}(t)$ are full-dimensional for all $t>0$).

For an arbitrary matrix A, one can obtain the upper bound of $(n-1)$ switchings for sufficiently short intervals of time.

Theorem 15.5. *Consider the characteristic polynomial of the matrix A:*

$$\det(t\,\mathrm{Id} - A) = t^n + c_1 t^{n-1} + \cdots + c_n,$$

and let

$$c = \max_{1\leq i\leq n} |c_i|.$$

Then for any time-optimal control $u(t)$ and any $\bar{t}\in\mathbb{R}$, the real segment

$$\left[\bar{t}, \bar{t} + \ln\left(1+\frac{1}{c}\right)\right]$$

contains not more than $(n-1)$ switchings of an optimal control $u(t)$.

In the proof of this theorem we will require the following general proposition, which we learned from S. Yakovenko.

Lemma 15.2. *Consider an ODE*

$$y^{(n)} + c_1(t) y^{(n-1)} + \cdots + c_n(t) y = 0$$

with measurable and bounded coefficients:

$$c_i = \max_{t\in[\bar t,\bar t+\delta]} |c_i(t)|.$$

If

$$\sum_{k=1}^{n} c_k \frac{\delta^k}{k!} < 1, \tag{15.18}$$

then any nonzero solution $y(t)$ of the ODE has not more than $n-1$ roots on the segment $t \in [\bar t, \bar t+\delta]$.

Proof. By contradiction, suppose that the function $y(t)$ has at least n roots on the segment $t \in [\bar t, \bar t+\delta]$. By Rolle's Theorem, derivative $\dot y(t)$ has not less than $n-1$ roots, etc. Then $y^{(n-1)}$ has a root $t_{n-1} \in [\bar t, \bar t+\delta]$. Thus

$$y^{(n-1)}(t) = \int_{t_{n-1}}^{t} y^{(n)}(\tau)\, d\tau.$$

Let $t_{n-2} \in [\bar t, \bar t+\delta]$ be a root of $y^{(n-2)}(t)$, then

$$y^{(n-2)}(t) = \int_{t_{n-2}}^{t} d\tau_1 \int_{t_{n-1}}^{\tau_1} y^{(n)}(\tau_2)\, d\tau_2.$$

We continue this procedure by integrating $y^{(n-i+1)}(t)$ from a root $t_{n-i} \in [\bar t, \bar t+\delta]$ of $y^{(n-i)}(t)$ and obtain

$$y^{(n-i)}(t) = \int_{t_{n-i}}^{t} d\tau_1 \int_{t_{n-i+1}}^{\tau_1} d\tau_2 \cdots \int_{t_{n-1}}^{\tau_{i-1}} y^{(n)}(\tau_i)\, d\tau_i, \qquad i = 1, \dots, n.$$

There holds a bound:

$$\begin{aligned} \left|y^{(n-i)}(t)\right| &\le \int_{t_{n-i}}^{t} d\tau_1 \int_{t_{n-i+1}}^{\tau_1} d\tau_2 \cdots \int_{t_{n-1}}^{\tau_{i-1}} \left|y^{(n)}(\tau_i)\right| d\tau_i \\ &\le \int_{\bar t}^{\bar t+\delta} d\tau_1 \int_{\bar t}^{\tau_1} d\tau_2 \cdots \int_{\bar t}^{\tau_{i-1}} \left|y^{(n)}(\tau_i)\right| d\tau_i \le \frac{\delta^k}{k!} \sup_{t\in[\bar t,\bar t+\delta]} \left|y^{(n)}(t)\right|. \end{aligned}$$

Then

$$\left|\sum_{i=1}^{n} c_i(t) y^{(n-i)}(t)\right| \le \sum_{i=1}^{n} |c_i(t)| \left|y^{(n-i)}(t)\right| \le \sum_{i=1}^{n} c_i \frac{\delta^k}{k!} \sup_{t\in[\bar t,\bar t+\delta]} \left|y^{(n)}(t)\right|,$$

i.e.,

$$\left|y^{(n)}(t)\right| \le \sum_{i=1}^{n} c_i \frac{\delta^k}{k!} \sup_{t\in[\bar t,\bar t+\delta]} \left|y^{(n)}(t)\right|,$$

a contradiction with (15.18). The lemma is proved. □

Now we prove Theorem 15.5.

Proof. As we showed in the proof of Theorem 15.4, the number of switchings of $u(t)$ is not more than the number of roots of the function $y(t) = \xi(t)b$, which satisfies ODE (15.15).

We have

$$\sum_{k=1}^{n} |c_k| \frac{\delta^k}{k!} < c(e^\delta - 1) \quad \forall \delta > 0.$$

By Lemma 15.2, if

$$c(e^\delta - 1) \leq 1, \tag{15.19}$$

then the function $y(t)$ has not more than $n - 1$ real roots on any interval of length δ. But inequality (15.19) is equivalent to the following one:

$$\delta \leq \ln \left(1 + \frac{1}{c}\right),$$

so $y(t)$ has not more than $n-1$ roots on any interval of the length $\ln\left(1 + \frac{1}{c}\right)$. □

Chapter 16

Linear-quadratic problem

16.1 Problem statement and assumptions

In this chapter we study a class of optimal control problems very popular in applications, *linear-quadratic problems*. That is, we consider linear systems with quadratic cost functional:

$$
\begin{aligned}
&\dot{x} = Ax + Bu, \qquad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m, \\
&x(0) = x_0, \quad x(t_1) = x_1, \qquad x_0, x_1, t_1 \text{ fixed}, \\
&J(u) = \frac{1}{2}\int_0^{t_1} \langle Ru(t), u(t)\rangle + \langle Px(t), u(t)\rangle + \langle Qx(t), x(t)\rangle \, dt \to \min .
\end{aligned}
\tag{16.1}
$$

Here A, B, R, P, Q are constant matrices of appropriate dimensions, R, Q are symmetric:

$$
R^* = R, \qquad Q^* = Q,
$$

and angle brackets $\langle \cdot, \cdot \rangle$ denote the standard inner product in $\mathbb{R}^m$ and $\mathbb{R}^n$.

One can show that the condition $R \geq 0$ is necessary for existence of optimal control. We do not touch here the case of degenerate R and assume that $R > 0$. The substitution of variables $u \mapsto v = R^{1/2}u$ transforms the functional $J(u)$ to a similar functional with the identity matrix instead of R. That is why we assume in the sequel that $R = \mathrm{Id}$. Another change of variables kills the matrix P (exercise: find this change of variables). So we can write the cost functional as follows:

$$
J(u) = \frac{1}{2}\int_0^{t_1} |u(t)|^2 + \langle Qx(t), x(t)\rangle \, dt.
$$

For dynamics of the problem, we assume that the linear system is controllable:

$$
\operatorname{rank}(B, AB, \dots, A^{n-1}B) = n. \tag{16.2}
$$

16.2 Existence of optimal control

Since the set of control parameters $U = \mathbb{R}^m$ is noncompact, Filippov's Theorem does not apply, and existence of optimal controls in linear-quadratic problems is a nontrivial problem.

In this chapter we assume that admissible controls are square-integrable:

$$u \in L_2^m[0, t_1]$$

and use the L_2^m norm for controls:

$$\|u\| = \left(\int_0^{t_1} |u(t)|^2 \, dt \right)^{1/2} = \left(\int_0^{t_1} u_1^2(t) + \cdots + u_m^2(t) \, dt \right)^{1/2}.$$

Consider the set of all admissible controls that steer the initial point to the terminal one:

$$U(x_0, x_1) = \{ u \in L_2^m[0, t_1] \mid x(t_1, u, x_0) = x_1 \}.$$

We denote by $x(t, u, x_0)$ the trajectory of system (16.1) corresponding to an admissible control $u \in L_2^m$ starting at a point $x_0 \in \mathbb{R}^n$. By Cauchy's formula, the endpoint mapping

$$u \mapsto x(t_1, u, x_0) = e^{t_1 A} x_0 + \int_0^{t_1} e^{(t_1 - \tau)A} B u(\tau) \, d\tau$$

is an affine mapping from $L_2^m[0, t_1]$ to $\mathbb{R}^n$. Controllability of the linear system (16.1) means that for any $x_0 \in \mathbb{R}^n$, $t_1 > 0$, the image of the endpoint mapping is the whole $\mathbb{R}^n$. Thus

$$U(x_0, x_1) \subset L_2^m[0, t_1]$$

is an affine subspace,

$$U(0, 0) \subset L_2^m[0, t_1]$$

is a linear subspace, and

$$U(x_0, x_1) = u + U(0, 0) \quad \text{for any } u \in U(x_0, x_1).$$

Thus it is natural that existence of optimal controls is closely related to behavior of the cost functional $J(u)$ on the linear subspace $U(0, 0)$.

Proposition 16.1. (1) *If there exist points $x_0, x_1 \in \mathbb{R}^n$ such that*

$$\inf_{u \in U(x_0, x_1)} J(u) > -\infty, \tag{16.3}$$

then

$$J(u) \geq 0 \quad \forall u \in U(0, 0).$$

(2) *Conversely, if*

$$J(u) > 0 \quad \forall u \in U(0,0) \setminus 0,$$

then the minimum is attained:

$$\exists \min_{u \in U(x_0,x_1)} J(u) \quad \forall x_0, \; x_1 \in \mathbb{R}^n.$$

Remark. That is, the inequality

$$J|_{U(0,0)} \geq 0$$

is necessary for existence of optimal controls, at least for one pair (x_0, x_1), and the strict inequality

$$J|_{U(0,0)\setminus 0} > 0$$

is sufficient for existence of optimal controls for all pairs (x_0, x_1).

In the proof of Proposition 16.1, we will need the following auxiliary proposition.

Lemma 16.1. *If $J(v) > 0$ for all $v \in U(0,0) \setminus 0$, then*

$$J(v) \geq \alpha \|v\|^2 \qquad \text{for some } \alpha > 0 \text{ and all } v \in U(0,0),$$

or, which is equivalent,

$$\inf\{J(v) \mid \|v\| = 1, \; v \in U(0,0)\} > 0.$$

Proof. Let v_n be a minimizing sequence of the functional $J(v)$ on the sphere $\{\|v\| = 1\} \cap U(0,0)$. Closed balls in Hilbert spaces are weakly compact, thus we can find a subsequence weakly converging in the unit ball and preserve the notation v_n for its terms, so that

$$\begin{gathered} v_n \to \widehat{v} \text{ weakly as } n \to \infty, \qquad \|\widehat{v}\| \leq 1, \quad \widehat{v} \in U(0,0), \\ J(v_n) \to \inf\{J(v) \mid \|v\| = 1, \; v \in U(0,0)\}, \qquad n \to \infty. \end{gathered} \tag{16.4}$$

We have

$$J(v_n) = \frac{1}{2} + \frac{1}{2} \int_0^{t_1} \langle Q x_n(\tau), x_n(\tau) \rangle \, d\tau.$$

Since the controls converge weakly, then the corresponding trajectories converge strongly:

$$x_n(\cdot) \to x_{\widehat{v}}(\cdot), \qquad n \to \infty,$$

thus

$$J(v_n) \to \frac{1}{2} + \frac{1}{2} \int_0^{t_1} \langle Q x_{\widehat{v}}(\tau), x_{\widehat{v}}(\tau) \rangle \, d\tau, \qquad n \to \infty.$$

In view of (16.4), the infimum in question is equal to

$$\frac{1}{2} + \frac{1}{2} \int_0^{t_1} \langle Q x_{\widehat{v}}(\tau), x_{\widehat{v}}(\tau) \rangle \, d\tau = \frac{1}{2} \left(1 - \|\widehat{v}\|^2\right) + J(\widehat{v}) > 0.$$

□

Now we prove Proposition 16.1

Proof. (1) By contradiction, suppose that there exists $v \in U(0,0)$ such that $J(v) < 0$. Take any $u \in U(x_0, x_1)$, then $u + sv \in U(x_0, x_1)$ for any $s \in \mathbb{R}$.

Let $y(t)$, $t \in [0, t_1]$, be the solution to the Cauchy problem

$$\dot{y} = Ay + Bv, \qquad y(0) = 0,$$

and

$$J(u, v) = \frac{1}{2} \int_0^{t_1} \langle u(\tau), v(\tau) \rangle + \langle Qx(\tau), y(\tau) \rangle \, d\tau.$$

Then the quadratic functional J on the family of controls $u + sv$, $s \in \mathbb{R}$, is computed as follows:

$$J(u + sv) = J(u) + 2sJ(u, v) + s^2 J(v).$$

Since $J(v) < 0$, then $J(u + sv) \to -\infty$ as $s \to \infty$. The contradiction with hypothesis (16.3) finishes the proof of item (1) of this proposition.

(2) We have

$$J(u) = \frac{1}{2}\|u\|^2 + \frac{1}{2} \int_0^{t_1} \langle Qx(\tau), x(\tau) \rangle \, d\tau.$$

The norm $\|u\|$ is lower semicontinuous in the weak topology on L_2^m, and the functional $\int_0^{t_1} \langle Qx(\tau), x(\tau) \rangle \, d\tau$ is weakly continuous on L_2^m. Thus $J(u)$ is weakly lower semicontinuous on L_2^m. Since balls are weakly compact in L_2^m and the affine subspace $U(x_0, x_1)$ is weakly compact, it is enough to prove that $J(u) \to \infty$ when $u \to \infty$, $u \in U(x_0, x_1)$.

Take any control $u \in U(x_0, x_1)$. Then for any $v \in U(0,0) \setminus 0$, the control $u + v$ belongs to $U(x_0, x_1)$ and

$$J(u + v) = J(u) + 2\|v\| J\left(u, \frac{v}{\|v\|}\right) + J(v).$$

Denote $J(u) = C_0$. Further, $\left| J\left(u, \frac{v}{\|v\|}\right) \right| \leq C_1 = \text{const}$ for all $v \in U(0,0) \setminus 0$. Finally, by Lemma 16.1, $J(v) \geq \alpha \|v\|^2$, $\alpha > 0$, for all $v \in U(0,0) \setminus 0$. Consequently,

$$J(u + v) \geq C_0 - 2\|v\| C_1 + \alpha \|v\|^2 \to \infty, \qquad v \to \infty, \; v \in U(0,0).$$

Item (2) of this proposition follows. □

So we reduced the question of existence of optimal controls in linear-quadratic problems to the study of the restriction $J|_{U(0,0)}$. We will consider this restriction in detail later.

16.3 Extremals

Now we write PMP for linear-quadratic problems. The control-dependent Hamiltonian is

$$h_u(\xi, x) = \xi A x + \xi B u - \frac{\nu}{2}(\|u\|^2 + \langle Qx, x\rangle), \qquad x \in \mathbb{R}^n, \ \xi \in \mathbb{R}^{n*}.$$

Consider first the abnormal case:

$$\nu = 0.$$

By PMP, adjoint vector along an extremal satisfies the ODE $\dot{\xi} = -\xi A$, thus $\xi(t) = \xi(0)e^{-tA}$. The maximality condition

$$\xi(t)Bu(t) = \max_{u \in \mathbb{R}^n} \xi(t)Bu \tag{16.5}$$

implies that

$$0 \equiv \xi(t)B = \xi(0)e^{-tA}B.$$

We differentiate this identity $n-1$ times, take into account the controllability condition (16.2) and obtain $\xi(0) = 0$. This contradicts PMP, thus there are no abnormal extremals.

In the sequel we consider the normal case: $\nu \neq 0$, thus we can assume

$$\nu = 1.$$

Then the control-dependent Hamiltonian takes the form

$$h_u(\xi, x) = \xi A x + \xi B u - \frac{1}{2}(\|u\|^2 + \langle Qx, x\rangle), \qquad x \in \mathbb{R}^n, \ \xi \in \mathbb{R}^{n*}.$$

The term $\xi B u - \frac{1}{2}\|u\|^2$ depending on u has a unique maximum in $u \in \mathbb{R}^m$ at the point where

$$\frac{\partial h_u}{\partial u} = \xi B - u^* = 0,$$

thus

$$u = B^* \xi^*.$$

So the maximized Hamiltonian is

$$\begin{aligned} H(\xi, x) &= \max_{u \in \mathbb{R}^m} h_u(\xi, x) = \xi A x - \frac{1}{2}\langle Qx, x\rangle + \frac{1}{2}|B^* \xi^*|^2 \\ &= \xi A x - \frac{1}{2}\langle Qx, x\rangle + \frac{1}{2}|B\xi|^2. \end{aligned}$$

The Hamiltonian function $H(\xi, x)$ is smooth, thus extremals are solutions of the corresponding Hamiltonian system

$$\begin{cases} \dot{x} = Ax + BB^*\xi^*, \\ \dot{\xi} = x^* Q - \xi A. \end{cases}$$

16.4 Conjugate points

Now we study conditions of existence and uniqueness of optimal controls depending upon the terminal time. So we write the cost functional to be minimized as follows:

$$J_t(u) = \frac{1}{2}\int_0^t |u(\tau)|^2 + \langle Qx(\tau), x(\tau)\rangle \, d\tau.$$

Denote

$$\begin{aligned} U_t(0,0) &= \{u \in L_2^m[0,t] \mid x(t,u,x_0) = x_1\}, \\ \mu(t) &\stackrel{\text{def}}{=} \inf\{J_t(u) \mid u \in U_t(0,0), \ \|u\| = 1\}. \end{aligned} \tag{16.6}$$

We showed in Proposition 16.1 that if $\mu(t) > 0$ then the problem has solution for any boundary conditions, and if $\mu(t) < 0$ then there are no solutions for any boundary conditions. The case $\mu(t) = 0$ is doubtful. Now we study properties of the function $\mu(t)$ in detail.

Proposition 16.2. (1) *The function $t \mapsto \mu(t)$ is monotone nonincreasing and continuous.*

(2)

$$1 \geq 2\mu(t) \geq 1 - \frac{t^2}{2} e^{2t\|A\|}\|B\|^2\|Q\|. \tag{16.7}$$

(3) *If $1 > 2\mu(t)$, then the infimum in* (16.6) *is attained, i.e., it is minimum.*

Proof. (3) Denote

$$I_t(u) = \frac{1}{2}\int_0^t \langle Qx(\tau), x(\tau)\rangle \, d\tau,$$

the functional $I_t(u)$ is weakly continuous on L_2^m. Notice that

$$J_t(u) = \frac{1}{2} + I_t(u) \quad \text{on the sphere } \|u\| = 1.$$

Take a minimizing sequence of the functional $I_t(u)$ on the sphere $\{\|u\| = 1\} \cap U_t(0,0)$. Since the ball $\{\|u\| \leq 1\}$ is weakly compact, we can find a weakly converging subsequence:

$$\begin{aligned} &u_n \to \widehat{u} \text{ weakly as } n \to \infty, \qquad \|\widehat{u}\| \leq 1, \quad \widehat{u} \in U_t(0,0), \\ &I_t(u_n) \to I_t(\widehat{u}) = \inf\{I_t(u) \mid \|u\| = 1, \ u \in U_t(0,0)\}, \qquad n \to \infty. \end{aligned}$$

If $\widehat{u} = 0$, then $I_t(\widehat{u}) = 0$, thus $\mu(t) = \frac{1}{2}$, which contradicts hypothesis of item (3).

So $\widehat{u} \neq 0$, $I_t(\widehat{u}) < 0$, and $I_t\left(\frac{\widehat{u}}{\|\widehat{u}\|}\right) \leq I_t(\widehat{u})$. Thus $\|\widehat{u}\| = 1$, and $J_t(u)$ attains minimum on the sphere $\{\|u\| = 1\} \cap U_t(0,0)$ at the point $\widehat{u}$.

(2) Let $\|u\| = 1$ and $x_0 = 0$. By Cauchy's formula,

$$x(t) = \int_0^t e^{(t-\tau)A} Bu(\tau) \, d\tau,$$

thus

$$|x(t)| \leq \int_0^t e^{(t-\tau)\|A\|}\|B\| \cdot |u(\tau)|\, d\tau$$

by Cauchy-Schwartz inequality

$$\leq \|u\| \left(\int_0^t e^{(t-\tau)2\|A\|}\|B\|^2\, d\tau \right)^{1/2}$$
$$= \left(\int_0^t e^{(t-\tau)2\|A\|}\|B\|^2\, d\tau \right)^{1/2}.$$

We substitute this estimate of $x(t)$ into J_t and obtain the second inequality in (16.7).

The first inequality in (16.7) is obtained by considering a weakly converging sequence $u_n \to 0$, $n \to \infty$, in the sphere $\|u_n\| = 1$, $u_n \in U_t(0,0)$.

(1) Monotonicity of $\mu(t)$. Take any $\hat{t} > t$. Then the space $U_t(0,0)$ is isometrically embedded into $U_{\hat{t}}(0,0)$ by extending controls $u \in U_t(0,0)$ by zero:

$$u \in U_t(0,0) \quad \Rightarrow \quad \widehat{u} \in U_{\hat{t}}(0,0),$$
$$\widehat{u}(\tau) = \begin{cases} u(\tau), & \tau \leq t, \\ 0, & \tau > t. \end{cases}$$

Moreover,

$$J_{\hat{t}}(\widehat{u}) = J_t(u).$$

Thus

$$\begin{aligned} \mu(t) &= \inf\{J_t(u) \mid u \in U_t(0,0), \|u\| = 1\} \\ &\geq \inf\{J_{\hat{t}}(u) \mid u \in U_{\hat{t}}(0,0), \|u\| = 1\} = \mu(\hat{t}). \end{aligned}$$

Continuity of $\mu(t)$: we show separately continuity from the right and from the left.

Continuity from the right. Let $t_n \searrow t$. We can assume that $\mu(t_n) < \frac{1}{2}$ (otherwise $\mu(t_n) = \mu(t) = \frac{1}{2}$), thus minimum in (16.6) is attained:

$$\mu(t_n) = \frac{1}{2} + I_{t_n}(u_n), \qquad u_n \in U_{t_n}(0,0), \quad \|u_n\| = 1.$$

Extend the functions $u_n \in L_2^m[0, t_n]$ to the segment $[0, t_1]$ by zero. Choosing a weakly converging subsequence in the unit ball, we can assume that

$$u_n \to u \text{ weakly as } n \to \infty, \qquad u \in U_t(0,0), \quad \|u\| \leq 1,$$

thus

$$I_{t_n}(u_n) \to I_t(u) \geq \inf\{I_t(v) \mid v \in U_t(0,0), \quad \|v\| = 1\}, \quad t_n \searrow t.$$

Then

$$\mu(t) \leq \frac{1}{2} + \lim_{t_n \searrow t} I_{t_n}(u_n) = \lim_{t_n \searrow t} \mu(t_n).$$

By monotonicity of μ,

$$\mu(t) = \lim_{t_n \searrow t} \mu(t_n),$$

i.e., continuity from the right is proved.

Continuity from the left. We can assume that $\mu(t) < \frac{1}{2}$ (otherwise $\mu(\tau) = \mu(t) = \frac{1}{2}$ for $\tau < t$). Thus minimum in (16.6) is attained:

$$\mu(t) = \frac{1}{2} + I_t(\widehat{u}), \qquad \widehat{u} \in U_t(0,0), \quad \|\widehat{u}\| = 1.$$

For the trajectory

$$\widehat{x}(\tau) = x(\tau, \widehat{u}, 0),$$

we have

$$\widehat{x}(\tau) = \int_0^\tau e^{(\tau-\theta)A} B\widehat{u}(\theta)\, d\theta.$$

Denote

$$\alpha(\varepsilon) = \| \, \widehat{u}|_{[0,\varepsilon]} \, \|$$

and notice that

$$\alpha(\varepsilon) \to 0, \qquad \varepsilon \to 0.$$

Denote the ball

$$B_\delta = \{u \in L_2^m \mid \|u\| \leq \delta, \ u \in U(0,0)\}.$$

Obviously,

$$x(\varepsilon, B_{\alpha(\varepsilon)}, 0) \ni \widehat{x}(\varepsilon).$$

The mapping $u \mapsto x(\varepsilon, u(\cdot), 0)$ from L_2^m to $\mathbb{R}^n$ is linear, and the system $\dot{x} = Ax + Bu$ is controllable, thus $x(\varepsilon, B_{\alpha(\varepsilon)}, 0)$ is a convex full-dimensional set in $\mathbb{R}^n$ such that the positive cone generated by this set is the whole $\mathbb{R}^n$. That is why

$$x(\varepsilon, 2B_{\alpha(\varepsilon)}, 0) = 2x(\varepsilon, B_{\alpha(\varepsilon)}, 0) \supset O_{x(\varepsilon, B_{\alpha(\varepsilon)}, 0)}$$

for some neighborhood $O_{x(\varepsilon, B_{\alpha(\varepsilon)}, 0)}$ of the set $x(\varepsilon, B_{\alpha(\varepsilon)}, 0)$. Further, there exists an instant $t_\varepsilon > \varepsilon$ such that

$$\widehat{x}(t_\varepsilon) \in x(\varepsilon, 2B_{\alpha(\varepsilon)}, 0),$$

consequently,

$$\widehat{x}(t_\varepsilon) = x(\varepsilon, v_\varepsilon, 0), \qquad \|v_\varepsilon\| \leq 2\alpha(\varepsilon).$$

Consider the following family of controls that approximate $\widehat{u}$:

$$u_\varepsilon(\tau) = \begin{cases} v_\varepsilon(\tau), & 0 \leq \tau \leq t_\varepsilon, \\ \widehat{u}(\tau + t_\varepsilon - \varepsilon), & t_\varepsilon < \tau \leq t + \varepsilon - t_\varepsilon. \end{cases}$$

We have

$$\begin{aligned} &u_\varepsilon \in U_{t+\varepsilon-t_\varepsilon}(0,0), \\ &\|\widehat{u} - u_\varepsilon\| \to 0, \qquad \varepsilon \to 0. \end{aligned}$$

But $t + \varepsilon - t_\varepsilon < t$ and μ is nonincreasing, thus it is continuous from the left.

Continuity from the right was already proved, hence μ is continuous. □

Now we prove that the function μ can have not more than one root.

Proposition 16.3. *If $\mu(t) = 0$ for some $t > 0$, then $\mu(\tau) < 0$ for all $\tau > t$.*

Proof. Let $\mu(t) = 0$, $t > 0$. By Proposition 16.2, infimum in (16.6) is attained at some control $\widehat{u} \in U_t(0,0)$, $\|\widehat{u}\| = 1$:

$$\begin{aligned} \mu(t) &= \min\{J_t(u) \mid u \in U_t(0,0), \ \|u\| = 1\} \\ &= J_t(\widehat{u}) = 0. \end{aligned}$$

Then

$$J_t(u) \geq J_t(\widehat{u}) = 0 \quad \forall u \in U_t(0,0),$$

i.e., the control $\widehat{u}$ is optimal, thus it satisfies PMP. There exists a solution $(\xi(\tau), x(\tau))$, $\tau \in [0,t]$, of the Hamiltonian system

$$\begin{cases} \dot{\xi} = x^* Q - \xi A, \\ \dot{x} = Ax + BB^*\xi, \end{cases}$$

with the boundary conditions

$$x(0) = x(t) = 0,$$

and

$$u(\tau) = B^* \xi^*(\tau), \qquad \tau \in [0,t].$$

We proved that for any root t of the function μ, any control $u \in U_t(0,0)$, $\|u\| = 1$, with $J_t(u) = 0$ satisfies PMP.

Now we prove that $\mu(\tau) < 0$ for all $\tau > t$. By contradiction, suppose that the function μ vanishes at some instant $t' > t$. Since μ is monotone, then

$$\mu|_{[t,t']} \equiv 0.$$

Consequently, the control

$$u'(\tau) = \begin{cases} \widehat{u}(\tau), & \tau \leq t, \\ 0, & \tau \in [t, t'], \end{cases}$$

satisfies the conditions:

$$\begin{aligned} &u' \in U_{t'}(0,0), \quad \|u'\| = 1, \\ &J_{t'}(u') = 0. \end{aligned}$$

Thus u' satisfies PMP, i.e.,

$$u'(\tau)B^*{\xi^*}'(\tau), \qquad \tau \in [0, t'],$$

is an analytic function. But $u'|_{[t,t']} \equiv 0$, thus $u' \equiv 0$, a contradiction with $\|u'\| = 1$. □

It would be nice to have a way to solve the equation $\mu(t) = 0$ without performing the minimization procedure in (16.6). This can be done in terms of the following notion.

Definition 16.1. A point $t > 0$ is *conjugate* to 0 for the linear-quadratic problem in question if there exists a nontrivial solution $(\xi(\tau), x(\tau))$ of the Hamiltonian system

$$\begin{cases} \dot{\xi} = x^*Q - \xi A, \\ \dot{x} = Ax + BB^*\xi \end{cases}$$

such that $x(0) = x(t) = 0$.

Proposition 16.4. *The function μ vanishes at a point $t > 0$ if and only if t is the closest to 0 conjugate point.*

Proof. Let $\mu(t) = 0$, $t > 0$. First of all, t is conjugate to 0, we showed this in the proof of Proposition 16.3.

Suppose that $t' > 0$ is conjugate to 0. Compute the functional $J_{t'}$ on the corresponding control $u(\tau) = B^*\xi^*(\tau)$, $\tau \in [0, t']$:

$$\begin{aligned} J_{t'}(u) &= \frac{1}{2}\int_0^{t'} \langle B^*\xi^*(\tau), B^*\xi^*(\tau)\rangle + \langle Qx(\tau), x(\tau)\rangle \, d\tau \\ &= \frac{1}{2}\int_0^{t'} \langle BB^*\xi^*(\tau), \xi^*(\tau)\rangle + \langle Qx(\tau), x(\tau)\rangle \, d\tau \\ &= \frac{1}{2}\int_0^{t'} \xi(\tau)(\dot{x}(\tau) - Ax(\tau)) + x^*(\tau)Qx(\tau) \, d\tau \\ &= \frac{1}{2}\int_0^{t'} (\xi\dot{x} + \dot{\xi}x) \, d\tau \\ &= \frac{1}{2}(\xi(t')x(t') - \xi(0)x(0)) = 0. \end{aligned}$$

Thus $\mu(t') \leq J_{t'}\left(\frac{u}{\|u\|}\right) = 0$. Now the result follows since μ is nonincreasing. □

The first (closest to zero) conjugate point determines existence and uniqueness properties of optimal control in linear-quadratic problems.

Before the first conjugate point, optimal control exists and is unique for any boundary conditions (if there are two optimal controls, then their difference gives rise to a conjugate point).

At the first conjugate point, there is existence and nonuniqueness for some boundary conditions, and nonexistence for other boundary conditions.

And after the first conjugate point, the problem has no optimal solutions for any boundary conditions.

Chapter 17

Sufficient optimality conditions, Hamilton-Jacobi equation, and Dynamic Programming

17.1 Sufficient optimality conditions

Pontryagin Maximum Principle is a universal and powerful necessary optimality condition, but the theory of sufficient optimality conditions is not so complete. In this section we consider an approach to sufficient optimality conditions that generalizes fields of extremals of the Classical Calculus of Variations.

Consider the following optimal control problem:

$$\dot{q} = f_u(q), \qquad q \in M, \ u \in U, \tag{17.1}$$

$$q(0) = q_0, \ q(t_1) = q_1, \quad q_0, q_1, t_1 \text{ fixed}, \tag{17.2}$$

$$\int_0^{t_1} \varphi(q(t), u(t))\, dt \to \min. \tag{17.3}$$

The control-dependent Hamiltonian of PMP corresponding to the normal case is

$$h_u(\lambda) = \langle \lambda, f_u(q) \rangle - \varphi(q, u), \qquad \lambda \in T^*M, \ q = \pi(\lambda) \in M, \ u \in U.$$

Assume that the maximized Hamiltonian

$$H(\lambda) = \max_{u \in U} h_u(\lambda) \tag{17.4}$$

is defined and smooth on T^*M. We can assume smoothness of H on an open domain $O \subset T^*M$ and modify respectively the subsequent results. But for simplicity of exposition we prefer to take $O = T^*M$. Then trajectories of the Hamiltonian system

$$\dot{\lambda} = \vec{H}(\lambda)$$

are extremals of problem (17.1)–(17.3). We assume that the Hamiltonian vector field $\vec{H}$ is complete.

17.1.1 Integral invariant

First we consider a general construction that will play a key role in the proof of sufficient optimality conditions.

Fix an arbitrary smooth function

$$a \in C^\infty(M).$$

Then the graph of differential da is a smooth submanifold in T^*M:

$$\mathcal{L}_0 = \{d_q a \mid q \in M\} \subset T^*M,$$
$$\dim \mathcal{L}_0 = \dim M = n.$$

Translations of $\mathcal{L}_0$ by the flow of the Hamiltonian vector field

$$\mathcal{L}_t = e^{t\vec{H}}(\mathcal{L}_0)$$

are smooth n-dimensional submanifolds in T^*M, and the graph of the mapping $t \mapsto \mathcal{L}_t$,

$$\mathcal{L} = \{(\lambda, t) \mid \lambda \in \mathcal{L}_t,\ 0 \le t \le t_1\} \subset T^*M \times \mathbb{R}$$

is a smooth $(n+1)$-dimensional submanifold in $T^*M \times \mathbb{R}$.

Consider the 1-form

$$s - H\, dt \in \Lambda^1(T^*M \times \mathbb{R}).$$

Recall that s is the tautological 1-form on T^*M, $s_\lambda = \lambda \circ \pi_*$, and its differential is the canonical symplectic structure on T^*M, $ds = \sigma$. In mechanics, the form $s - H\, dt = p\, dq - H\, dt$ is called the *integral invariant of Poincaré-Cartan* on the extended phase space $T^*M \times \mathbb{R}$.

Proposition 17.1. *The form* $(s - H\, dt)|_{\mathcal{L}}$ *is exact.*

Proof. First we prove that the form is closed:

$$0 = d(s - H\, dt)|_{\mathcal{L}} = (\sigma - dH \wedge dt)|_{\mathcal{L}}. \tag{17.5}$$

(1) Fix $\mathcal{L}_t = \mathcal{L} \cap \{t = \text{const}\}$ and consider restriction of the form $\sigma - dH \wedge dt$ to $\mathcal{L}_t$. We have

$$(\sigma - dH \wedge dt)|_{\mathcal{L}_t} = \sigma|_{\mathcal{L}_t}$$

since $dt|_{\mathcal{L}_t} = 0$. Recall that $\widehat{e^{t\vec{H}}}\,\sigma = \sigma$, thus

$$\sigma|_{\mathcal{L}_t} = \left(\widehat{e^{t\vec{H}}}\,\sigma\right)\Big|_{\mathcal{L}_0} = \sigma|_{\mathcal{L}_0} = ds|_{\mathcal{L}_0}.$$

But $s|_{\mathcal{L}_0} = d(a \circ \pi)|_{\mathcal{L}_0}$, hence

$$ds|_{\mathcal{L}_0} = d \circ d(a \circ \pi)|_{\mathcal{L}_0} = 0.$$

We proved that $(\sigma - dH \wedge dt)_{\mathcal{L}_t} = 0$.

(2) The manifold $\mathcal{L}$ is the image of the smooth mapping

$$(\lambda, t) \mapsto \left(e^{t\vec{H}}\lambda, t\right),$$

thus the tangent vector to $\mathcal{L}$ transversal to $\mathcal{L}_t$ is

$$\vec{H}(\lambda) + \frac{\partial}{\partial t} \in T_{(\lambda,t)}\mathcal{L}.$$

So

$$T_{(\lambda,t)}\mathcal{L} = T_{(\lambda,t)}\mathcal{L}_t \oplus \mathbb{R}\left(\vec{H}(\lambda) + \frac{\partial}{\partial t}\right).$$

To complete the proof, we substitute the vector $\vec{H}(\lambda) + \frac{\partial}{\partial t}$ as the first argument to $\sigma - dH \wedge dt$ and show that the result is equal to zero. We have:

$$\begin{aligned}
&i_{\vec{H}}\sigma = -dH, \quad i_{\frac{\partial}{\partial t}}\sigma = 0,\\
&i_{\vec{H}}(dH \wedge dt) = \underbrace{(i_{\vec{H}}dH)}_{=0} \wedge dt - dH \wedge \underbrace{(i_{\vec{H}}dt)}_{=0} = 0,\\
&i_{\frac{\partial}{\partial t}}(dH \wedge dt) = \underbrace{\left(i_{\frac{\partial}{\partial t}}dH\right)}_{=0} \wedge dt - dH \wedge \underbrace{\left(i_{\frac{\partial}{\partial t}}dt\right)}_{=1} = -dH,
\end{aligned}$$

consequently,

$$i_{\vec{H}+\frac{\partial}{\partial t}}(\sigma - dH \wedge dt) = -dH + dH = 0.$$

We proved that the form $(s - H\,dt)|_{\mathcal{L}}$ is closed.

(3) Now we show that this form is exact, i.e.,

$$\int_\gamma s - H\,dt = 0 \tag{17.6}$$

for any closed curve

$$\gamma \;:\; \tau \mapsto (\lambda(\tau), t(\tau)) \in \mathcal{L}, \qquad \tau \in [0,1].$$

The curve γ is homotopic to the curve

$$\gamma_0 \;:\; \tau \mapsto (\lambda(\tau), 0) \in \mathcal{L}_0, \qquad \tau \in [0,1].$$

Since the form $(s - H\,dt)|_{\mathcal{L}}$ is closed, Stokes' theorem yields that

$$\int_{\gamma} s - H\,dt = \int_{\gamma_0} s - H\,dt.$$

But the integral over the closed curve $\gamma_0 \subset \mathcal{L}_0$ is easily computed:

$$\int_{\gamma_0} s - H\,dt = \int_{\gamma_0} s = \int_{\gamma_0} da = 0.$$

Equality (17.6) follows, i.e., the form $(s - H\,dt)|_{\mathcal{L}}$ is exact. □

17.1.2 Problem with fixed time

Now we prove sufficient optimality conditions for problem (17.1)–(17.3).

Theorem 17.1. *Assume that the restriction of projection $\pi|_{\mathcal{L}_t}$ is a diffeomorphism for any $t \in [0, t_1]$. Then for any $\lambda_0 \in \mathcal{L}_0$, the normal extremal trajectory*

$$\widehat{q}(t) = \pi \circ e^{t\vec{H}}(\lambda_0), \qquad 0 \le t \le t_1,$$

realizes a strict minimum of the cost functional $\int_0^{t_1} \varphi(q(t), u(t))\,dt$ among all admissible trajectories $q(t)$, $0 \le t \le t_1$, of system (17.1) with the same boundary conditions:

$$q(0) = \widehat{q}(0), \quad q(t_1) = \widehat{q}(t_1). \tag{17.7}$$

Remarks. (1) Under the hypotheses of this theorem, no check of existence of optimal control is required.

(2) If all assumptions (smoothness of H, extendibility of trajectories of $\vec{H}$ to the time segment $[0, t_1]$, diffeomorphic property of $\pi|_{\mathcal{L}_t}$) hold in a proper open domain $O \subset T^*M$, then the statement can be modified to give local optimality of $\widehat{q}(\cdot)$ in $\pi(O)$. These modifications are left to the reader.

Now we prove Theorem 17.1.

Proof. The curve $\widehat{q}(t)$ is projection of the normal extremal

$$\widehat{\lambda}_t = e^{t\vec{H}}(\lambda_0).$$

Let $\widehat{u}(t)$ be an admissible control that maximizes the Hamiltonian along this extremal:

$$H(\widehat{\lambda}_t) = h_{\widehat{u}(t)}(\widehat{\lambda}_t).$$

On the other hand, let $q(t)$ be an admissible trajectory of system (17.1) generated by a control $u(t)$ and satisfying the boundary conditions (17.7). We compare costs of the pairs $(\widehat{q}, \widehat{u})$ and (q, u).

Since $\pi : \mathcal{L}_t \to M$ is a diffeomorphism, the trajectory $\{q(t) \mid 0 \le t \le t_1\} \subset M$ can be lifted to a smooth curve $\{\lambda(t) \mid 0 \le t \le t_1\} \subset T^*M$:

$$\forall t \in [0, t_1] \quad \exists!\ \lambda(t) \in \mathcal{L}_t \quad \text{such that} \quad \pi(\lambda(t)) = q(t).$$

Then

$$\begin{aligned}
\int_0^{t_1} \varphi(q(t), u(t))\, dt &= \int_0^{t_1} \langle \lambda(t), f_{u(t)}(q(t))\rangle - h_{u(t)}(\lambda(t))\, dt \\
&\geq \int_0^{t_1} \langle \lambda(t), \dot{q}(t)\rangle - H(\lambda(t))\, dt \qquad (17.8)\\
&= \int_0^{t_1} \langle s_{\lambda(t)}, \dot{\lambda}(t)\rangle - H(\lambda(t))\, dt \\
&= \int_\gamma s - H\, dt,
\end{aligned}$$

where

$$\gamma \;:\; t \mapsto (\lambda(t), t) \in \mathcal{L}, \qquad t \in [0, t_1].$$

By Proposition 17.1, the form $(s - H\, dt)|_{\mathcal{L}}$ is exact. Then integral of the form $(s - H\, dt)|_{\mathcal{L}}$ along a curve depends only upon endpoints of the curve. The curves γ and

$$\widehat{\gamma} \;:\; t \mapsto (\widehat{\lambda}_t, t) \in \mathcal{L}, \quad \widehat{\lambda}_t = e^{t\vec{H}}(\lambda_0), \qquad t \in [0, t_1],$$

have the same endpoints, thus

$$\begin{aligned}
\int_\gamma s - H\, dt &= \int_{\widehat{\gamma}} s - H\, dt = \int_0^{t_1} \langle \widehat{\lambda}_t, \dot{\widehat{q}}(t)\rangle - H(\widehat{\lambda}_t)\, dt \\
&= \int_0^{t_1} \langle \widehat{\lambda}_t, f_{\widehat{u}(t)}(\widehat{q}(t))\rangle - h_{\widehat{u}(t)}(\widehat{\lambda}_t)\, dt \\
&= \int_0^{t_1} \varphi(\widehat{q}(t), \widehat{u}(t))\, dt.
\end{aligned}$$

So

$$\int_0^{t_1} \varphi(q(t), u(t))\, dt \geq \int_0^{t_1} \varphi(\widehat{q}(t), \widehat{u}(t))\, dt, \tag{17.9}$$

i.e., the trajectory $\widehat{q}(t)$ is optimal.

It remains to prove that the minimum of the pair $(\widehat{q}(t), \widehat{u}(t))$ is strict, i.e, that inequality (17.9) is strict.

For a fixed point $q \in M$, write cotangent vectors as $\lambda = (p, q)$, where p are coordinates of a covector λ in $T_q^* M$. The control- dependent Hamiltonians $h_u(p, q)$ are affine w.r.t. p, thus their maximum $H(p, q)$ is convex w.r.t. p. Any vector $\xi \in T_q M$ such that

$$\langle p, \xi\rangle = \max_{u \in U} \langle p, f_u(q)\rangle$$

defines a hyperplane of support to the epigraph of the mapping $p \mapsto H(p, q)$. Since H is smooth in p, such a hyperplane of support is unique and maximum in (17.4) is attained at a unique velocity vector:

$$H(p, q) = h_u(p, q) \text{ at a unique vector } \dot{\widehat{q}} = f_u(q).$$

If $q(t) \not\equiv \widehat{q}(t)$, then inequality (17.8) becomes strict, as well as inequality (17.9) (for details on convex functions, see e.g. [19]). The theorem is proved. $\square$

Sufficient optimality condition of Theorem 17.1 is given in terms of the manifolds $\mathcal{L}_t$, which are in their turn defined by a function a and the Hamiltonian flow of $\vec{H}$. One can prove optimality of a normal extremal trajectory $\widehat{q}(t)$, $t \in [0, t_1]$, if one succeeds to find an appropriate function $a \in C^\infty(M)$ for which the projections $\pi : \mathcal{L}_t \to M$, $t \in [0, t_1]$, are diffeomorphisms.

For $t = 0$ the projection $\pi : \mathcal{L}_0 \to M$ is a diffeomorphism. So for small $t > 0$ any function $a \in C^\infty(M)$ provides manifolds $\mathcal{L}_t$ well projected to M, at least if we restrict ourselves by a compact $K \Subset M$. Thus the sufficient optimality condition for small pieces of extremal trajectories follows.

Corollary 17.1. *For any compact $K \Subset M$ that contains a normal extremal trajectory*

$$\widehat{q}(t) = \pi \circ e^{t\vec{H}}(\lambda_0), \qquad 0 \le t \le t_1,$$

there exists $t_1' \in (0, t_1]$ such that the piece

$$\widehat{q}(t), \qquad 0 \le t \le t_1',$$

is optimal w.r.t. all trajectories contained in K.

In many problems, one can choose a sufficiently large compact $K \supset \widehat{q}$ such that the functional J is separated from below from zero on all trajectories leaving K (this is the case, e.g., if $\varphi(q, u) > 0$). Then small pieces of $\widehat{q}$ are globally optimal.

17.1.3 Problem with free time

For problems with integral cost and free terminal time t_1, a sufficient optimality condition similar to Theorem 17.1 is valid, see Theorem 17.2 below.

Recall that all normal extremals of the free time problem lie in the zero level $H^{-1}(0)$ of the maximized Hamiltonian H.

Proposition 17.2. *Assume that 0 is a regular value of the restriction $H|_{\mathcal{L}_0}$, i.e. $d_\lambda H|_{T_\lambda \mathcal{L}_0} \neq 0$ for all $\lambda \in \mathcal{L}_0 \cap H^{-1}(0)$. Then the mapping*

$$\Phi : \mathcal{L}_0 \cap H^{-1}(0) \times \mathbb{R} \to T^*M, \qquad \Phi(\lambda_0, t) = e^{t\vec{H}}(\lambda_0),$$

is an immersion and $\widehat{\Phi}s$ is an exact form.

Proof. First of all, regularity of the value 0 for $H|_{\mathcal{L}_0}$ implies that $\mathcal{L}_0 \cap H^{-1}(0)$ is a smooth manifold. Then, the exactness of $\widehat{\Phi}s$ easily follows from Proposition 17.1. To prove that Φ is an immersion, it is enough to show that the vector $\frac{\partial \Phi}{\partial t}(\lambda_0, t) = \vec{H}(\lambda_t)$, $\lambda_t = \Phi(\lambda_0, t)$, is not tangent to the image of $\mathcal{L}_0 \cap H^{-1}(0)$ under the diffeomorphism $e^{t\vec{H}} : T^*M \to T^*M$ for all $\lambda_0 \in \mathcal{L}_0 \cap H^{-1}(0)$. Note that

$e^{t\vec{H}}(\mathcal{L}_0 \cap H^{-1}(0)) = \mathcal{L}_t \cap H^{-1}(0)$. We are going to prove a little bit more than we need, namely, that $\vec{H}(\lambda_t)$ is not tangent to $\mathcal{L}_t$.

Indeed, Proposition 17.1 implies that $\sigma|_{\mathcal{L}_t} = ds|_{\mathcal{L}_t} = 0$. Hence it is enough to show that the form $(i_{\vec{H}}\sigma)|_{\mathcal{L}_t}$ does not vanish at the point λ_t. Recall that the Hamiltonian flow $e^{t\vec{H}}$ preserves both σ and $\vec{H}$. In particular,

$$(i_{\vec{H}}\sigma)|_{\mathcal{L}_t} = \widehat{e^{t\vec{H}}}\left((i_{\vec{H}}\sigma)|_{\mathcal{L}_0}\right) = -\widehat{e^{t\vec{H}}}\left(dH|_{\mathcal{L}_0}\right).$$

The mapping $\widehat{e^{t\vec{H}}}$ is invertible. So it is enough to prove that $dH|_{\mathcal{L}_0}$ does not vanish at λ_0. But the last statement is our assumption! □

Now we obtain a sufficient optimality condition for the problem with free time.

Theorem 17.2. *Let W be a domain in $\mathcal{L}_0 \cap H^{-1}(0) \times \mathbb{R}$ such that*

$$\pi \circ \Phi|_W : W \to M$$

is a diffeomorphism of W onto a domain in M, and let

$$\widehat{\lambda}_t = e^{t\vec{H}}(\widehat{\lambda}_0), \qquad t \in [0, t_1],$$

be a normal extremal such that $(\widehat{\lambda}_0, t) \in W$ for all $t \in [0, t_1]$. Then the extremal trajectory $\widehat{q}(t) = \pi(\widehat{\lambda}_t)$ (with the corresponding control $\widehat{u}(t)$) realizes a strict minimum of the cost $\int_0^\tau \varphi(q(t), u(t))\, dt$ among all admissible trajectories such that $q(t) \in \pi \circ \Phi(W)$ for all $t \in [0, \tau]$, $q(0) = \widehat{q}(0)$, $q(\tau) = \widehat{q}(t_1)$, $\tau > 0$.

Proof. Set $\mathcal{L} = \Phi(W)$, then $\pi : \mathcal{L} \to \pi(\mathcal{L})$ is a diffeomorphism and $s|_{\mathcal{L}}$ is an exact form. Let $q(t)$, $t \in [0, \tau]$, be an admissible trajectory generated by a control $u(t)$ and contained in $\pi(\mathcal{L})$, with the boundary conditions $q(0) = \widehat{q}(0)$, $q(\tau) = \widehat{q}(t_1)$. Then $q(t) = \pi(\lambda(t))$, $0 \leq t \leq \tau$, where $t \mapsto \lambda(t)$ is a smooth curve in $\mathcal{L}$ such that $\lambda(0) = \widehat{\lambda}_0$, $\lambda(\tau) = \widehat{\lambda}_{t_1}$.

We have $\int\limits_{\lambda(\cdot)} s = \int\limits_{\widehat{\lambda}_\cdot} s$. Further,

$$\int\limits_{\widehat{\lambda}_\cdot} s = \int\limits_0^{t_1} \left\langle \widehat{\lambda}_t, \dot{\widehat{q}}(t) \right\rangle dt = \int\limits_0^{t_1} \left\langle \widehat{\lambda}_t, f_{\widehat{u}(t)}(\widehat{q}(t)) \right\rangle dt = \int\limits_0^{t_1} \varphi_{\widehat{u}(t)}(\widehat{q}(t))\, dt.$$

The last equality follows from the fact that

$$\left\langle \widehat{\lambda}(t), f_{\widehat{u}(t)}(\widehat{q}(t)) \right\rangle - \varphi(\widehat{q}(t), \widehat{u}(t)) = H(\widehat{\lambda}(t)) = 0.$$

On the other hand,

$$\begin{aligned} \int\limits_{\lambda(\cdot)} s &= \int\limits_0^{\tau} \langle \lambda(t), \dot{q}(t) \rangle\, dt = \int\limits_0^{\tau} h_{u(t)}(\lambda(t))\, dt + \int\limits_0^{\tau} \varphi(q(t), u(t))\, dt \\ &\leq \int\limits_0^{\tau} \varphi(q(t), u(t))\, dt. \end{aligned}$$

The last inequality follows since $\max\limits_{u\in U} h_u(\lambda(t)) = H(\lambda(t)) = 0$. Moreover, the inequality is strict if $t \mapsto \lambda(t)$ is not an extremal, i.e., if it does not coincide with $\widehat{\lambda}(t)$. Summing up,

$$\int\limits_0^{t_1} \varphi(\widehat{q}(t), \widehat{u}(t))\, dt \le \int\limits_0^{\tau} \varphi(q(t), u(t))\, dt$$

and the inequality is strict if q differs from $\widehat{q}$. □

17.2 Hamilton-Jacobi equation

Suppose that conditions of Theorem 17.1 are satisfied. As we showed in the proof of this theorem, the form $(s - H\, dt)|_{\mathcal{L}}$ is exact, thus it coincides with differential of some function:

$$(s - H\, dt)|_{\mathcal{L}} = dg, \qquad g : \mathcal{L} \to \mathbb{R}. \tag{17.10}$$

Since the projection $\pi : \mathcal{L}_t \to M$ is one-to-one, we can identify $(\lambda, t) \in \mathcal{L}_t \times \mathbb{R} \subset \mathcal{L}$ with $(q, t) \in M \times \mathbb{R}$ and define g as a function on $M \times \mathbb{R}$:

$$g = g(q, t).$$

In order to understand the meaning of the function g for our optimal control problem, consider an extremal

$$\widehat{\lambda}_t = e^{t\vec{H}}(\lambda_0)$$

and the curve

$$\widehat{\gamma} \subset \mathcal{L}, \quad \widehat{\gamma} : t \mapsto (\widehat{\lambda}_t, t),$$

as in the proof of Theorem 17.1. Then

$$\int_{\widehat{\gamma}} s - H\, dt = \int_0^{t_1} \varphi(\widehat{q}(\tau), \widehat{u}(\tau))\, d\tau, \tag{17.11}$$

where $\widehat{q}(t) = \pi(\widehat{\lambda}_t)$ is an extremal trajectory and $\widehat{u}(t)$ is the control that maximizes the Hamiltonian $h_u(\lambda)$ along $\widehat{\lambda}_t$. Equalities (17.10) and (17.11) mean that

$$g(\widehat{q}(t), t) = g(q_0, 0) + \int_0^t \varphi(\widehat{q}(\tau), \widehat{u}(\tau))\, d\tau,$$

i.e., $g(q, t) - g(q_0, 0)$ is the optimal cost of motion between points q_0 and q for the time t. Initial value for g can be chosen of the form

$$g(q_0, 0) = a(q_0), \qquad q_0 \in M. \tag{17.12}$$

Indeed, at $t = 0$ definition (17.11) of the function g reads

$$dg|_{t=0} = (s - H\,dt)|_{\mathcal{L}_0} = s|_{\mathcal{L}_0} = da,$$

which is compatible with (17.12).

We can rewrite equation (17.10) as a partial differential equation on g. In local coordinates on M and T^*M, we have

$$q = x \in M, \quad \lambda = (\xi, x) \in T^*M, \quad g = g(x, t).$$

Then equation (17.10) reads

$$(\xi\,dx - H(\xi, x)\,dt)|_{\mathcal{L}} = dg(x, t),$$

i.e.,

$$\begin{cases} \frac{\partial g}{\partial x} = \xi, \\ \frac{\partial g}{\partial t} = -H(\xi, x). \end{cases}$$

This system can be rewritten as a single first order nonlinear PDE:

$$\frac{\partial g}{\partial t} + H\left(\frac{\partial g}{\partial x}, x\right) = 0, \tag{17.13}$$

which is called *Hamilton-Jacobi equation*. We showed that the optimal cost $g(x, t)$ satisfies Hamilton-Jacobi equation (17.13) with initial condition (17.12).

Characteristic equations of PDE (17.13) have the form

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial \xi}, \\ \dot{\xi} = -\frac{\partial H}{\partial x}, \\ \dfrac{d}{dt} g(x(t), t) = \xi \dot{x} - H. \end{cases}$$

The first two equations form the Hamiltonian system $\dot{\lambda} = \vec{H}(\lambda)$ for normal extremals. Thus solving our optimal control problem (17.1)–(17.3) leads to the method of characteristics for the Hamilton-Jacobi equation for optimal cost.

17.3 Dynamic programming

One can derive the Hamilton-Jacobi equation for optimal cost directly, without Pontryagin Maximum Principle, due to an idea going back to Huygens and constituting a basis for Bellman's method of *Dynamic Programming*, see [7]. For this, it is necessary to assume that the optimal cost $g(q, t)$ exists and is C^1-smooth.

Let an optimal trajectory steer a point q_0 to a point q for a time t. Apply a constant control u on a time segment $[t, t + \delta t]$ and denote the trajectory starting at the point q by $q_u(\tau)$, $\tau \in [t, t + \delta t]$. Since $q_u(t + \delta t)$ is the endpoint of an admissible trajectory starting at q_0, the following inequality for optimal cost holds:

$$g(q_u(t+\delta t), t+\delta t) \leq g(q,t) + \int_t^{t+\delta t} \varphi(q_u(\tau), u)\, d\tau.$$

Divide by δt:

$$\frac{1}{\delta t}(g(q_u(t+\delta t), t+\delta t) - g(q,t)) \leq \frac{1}{\delta t}\int_t^{t+\delta t} \varphi(q_u(\tau), u)\, d\tau$$

and pass to the limit as $\delta t \to 0$:

$$\left\langle \frac{\partial g}{\partial q}, f_u(q) \right\rangle + \frac{\partial g}{\partial t} \leq \varphi(q,u).$$

So we obtain the inequality

$$\frac{\partial g}{\partial t} + h_u\left(\frac{\partial g}{\partial q}, q\right) \leq 0, \qquad u \in U. \tag{17.14}$$

Now let $(\tilde{q}(t), \tilde{u}(t))$ be an optimal pair. Let $t > 0$ be a Lebesgue point of the control $\tilde{u}$. Take any $\delta t \in (0,t)$. A piece of an optimal trajectory is optimal, thus $\tilde{q}(t-\delta t)$ is the endpoint of an optimal trajectory, as well as $\tilde{q}(t)$. So the optimal cost g satisfies the equality:

$$g(\tilde{q}(t), t) = g(\tilde{q}(t-\delta t), t-\delta t) + \int_{t-\delta t}^{t} \varphi(\tilde{q}(\tau), \tilde{u}(\tau))\, d\tau.$$

We repeat the above argument:

$$\frac{1}{\delta t}(g(\tilde{q}(t), t) - g(\tilde{q}(t-\delta t), t-\delta t)) = \frac{1}{\delta t}\int_{t-\delta t}^{t} \varphi(\tilde{q}(\tau), \tilde{u}(\tau))\, d\tau,$$

take the limit $\delta t \to 0$:

$$\frac{\partial g}{\partial t} + h_{\tilde{u}(t)}\left(\frac{\partial g}{\partial q}, q\right) = 0. \tag{17.15}$$

This equality together with inequality (17.14) means that

$$h_{\tilde{u}(t)}\left(\frac{\partial g}{\partial q}, q\right) = \max_{u \in U} h_u\left(\frac{\partial g}{\partial q}, q\right).$$

We denote

$$H(\xi, q) = \max_{u \in U} h_u(\xi, q)$$

and write (17.15) as Hamilton-Jacobi equation:

$$\frac{\partial g}{\partial t} + H\left(\frac{\partial g}{\partial q}, q\right) = 0.$$

Thus derivative of the optimal cost $\dfrac{\partial g}{\partial q}$ is equal to the impulse ξ along the optimal trajectory $\tilde{q}(t)$.

We do not touch here a huge theory on nonsmooth generalized solutions of Hamilton-Jacobi equation for smooth and nonsmooth Hamiltonians.

Chapter 18

Hamiltonian systems for geometric optimal control problems

18.1 Hamiltonian systems on trivialized cotangent bundle

18.1.1 Motivation

Consider a control system described by a family of vector fields on a manifold M:

$$\dot{q} = f_u(q), \qquad u \in \{1, \dots, k\}, \quad q \in M. \tag{18.1}$$

We construct a parametrization of the cotangent bundle T^*M adapted to this system. First, choose a basis in tangent spaces T_qM of the fields $f_u(q)$ and their iterated Lie brackets:

$$T_qM = \operatorname{span}(f_1(q), \dots, f_n(q)),$$

we assume that the system is bracket-generating. Then we have special coordinates in the tangent spaces:

$$\forall\ v \in T_qM \qquad v = \sum_{i=1}^{n} \xi_i f_i(q),$$
$$(\xi_1, \dots, \xi_n) \in \mathbb{R}^n.$$

Thus any tangent vector to M can be represented as an $(n+1)$-tuple

$$(\xi_1, \dots, \xi_n; q), \qquad (\xi_1, \dots, \xi_n) \in \mathbb{R}^n, \quad q \in M,$$

i.e., we obtain a kind of parametrization of the tangent bundle $TM = \cup_{q \in M} T_qM$ (one can construct coordinates on TM by choosing local coordinates in M, but such a choice is extraneous to our system, and we stay without any coordinates in M).

Having in mind the Hamiltonian system of PMP, we pass to the cotangent bundle. Construct the dual basis in T^*M: choose differential forms

$$\omega_1, \dots, \omega_n \in \Lambda^1 M$$

such that

$$\langle \omega_i, f_j \rangle = \delta_{ij}, \qquad i, j = 1, \dots, n.$$

Then the cotangent spaces become endowed with special coordinates:

$$\forall\, \lambda \in T_q^* M \qquad \lambda = \sum_{i=1}^{n} \eta_i \omega_{iq},$$
$$(\eta_1, \dots, \eta_n) \in \mathbb{R}^n.$$

So we obtain a kind of parametrization of the cotangent bundle:

$$\lambda \mapsto (\eta_1, \dots, \eta_n; q), \qquad (\eta_1, \dots, \eta_n) \in \mathbb{R}^n, \quad q \in M.$$

In notation of Sec. 11.5,

$$\eta_i = f_i^*(\lambda) = \langle \lambda, f_i(q) \rangle$$

is the linear on fibers Hamiltonian corresponding to the field f_i. Canonical coordinates on T^*M arise in a similar way from commuting vector fields $f_i = \frac{\partial}{\partial x_i}$, $i = 1, \dots, n$, corresponding to local coordinates $(x_1, \dots, x_n)$ on M. Consequently, in the (only interesting in control theory) case where the fields f_i do not commute, the "coordinates" $(\eta_1, \dots, \eta_n; q)$ on T^*M are not canonical.

Now our aim is to write Hamiltonian system in these nonstandard coordinates on T^*M, or in other natural coordinates adapted to the control system in question.

18.1.2 Trivialization of T^*M

Let M be a smooth manifold of dimension n, and let E be an n-dimensional vector space. Suppose that we have a *trivialization* of the cotangent bundle T^*M, i.e., a diffeomorphism

$$\Phi : E \times M \to T^*M$$

such that:

(1) the diagram

$$\begin{array}{ccc} E \times M & \xrightarrow{\Phi} & T^*M \\ \downarrow & & \downarrow \pi \\ M & = & M \end{array}$$

is commutative, i.e.,

$$\pi \circ \Phi(e, q) = q, \qquad e \in E, \quad q \in M,$$

(2) for any $q \in M$ the mapping

$$e \mapsto \Phi(e, q), \qquad e \in E,$$

is a linear isomorphism of vector spaces:

$$\Phi(\,\cdot\,, q) \,:\, E \to T_q^* M.$$

So the space E is identified with any vertical fiber $T_q^* M$, it is a typical fiber of the cotangent bundle $T^* M$.

For a fixed vector $e \in E$, we obtain a differential form on M:

$$\Phi_e \stackrel{\text{def}}{=} \Phi(e, \cdot\,) \in \Lambda^1 M.$$

In the previous section we had

$$E = \{(\eta_1, \dots, \eta_n)\} = \mathbb{R}^n,$$
$$\Phi(e, q) = \sum_{i=1}^{n} \eta_i \omega_{iq},$$

but now we do not fix any basis in E.

18.1.3 Symplectic form on $E \times M$

In order to write a Hamiltonian system on $E \times M \cong T^* M$, we compute the symplectic form $\widehat{\Phi}\sigma$ on $E \times M$. We start from the Liouville form

$$s \in \Lambda^1(T^* M)$$

and evaluate its pull-back

$$\widehat{\Phi} s \in \Lambda^1(E \times M).$$

The tangent and cotangent spaces are naturally identified with the direct products:

$$T_{(e,q)}(E \times M) \cong T_e E \oplus T_q M \cong E \oplus T_q M,$$
$$T^*_{(e,q)}(E \times M) \cong T_e^* E \oplus T_q^* M \cong E^* \oplus T_q^* M.$$

Any vector field $V \in \operatorname{Vec}(E \times M)$ is a sum of its vertical and horizontal parts:

$$V = V_v + V_h, \qquad V_v(e, q) \in E, \quad V_h(e, q) \in T_q M.$$

Similarly, any differential form

$$\omega \in \Lambda^1(E \times M)$$

is decomposed into its vertical and horizontal parts:

$$\omega = \omega_v + \omega_h, \qquad \omega_{v\,(e,q)} \in E^*, \quad \omega_{h\,(e,q)} \in T_q^* M.$$

The vertical part ω_v vanishes on horizontal tangent vectors, while the horizontal one ω_h vanishes on vertical tangent vectors.

In particular, vector fields and differential forms on M (possibly depending on $e \in E$) can be considered as horizontal vector fields and differential forms on $E \times M$:

$$T_q M = 0 \oplus T_q M \subset T_{(e,q)}(E \times M),$$
$$T_q^* M = 0 \oplus T_q^* M \subset T_{(e,q)}^*(E \times M).$$

Compute the action of the form $\widehat{\Phi}s$ on a tangent vector $(\xi, v) \in T_e E \oplus T_q M$:

$$\langle \widehat{\Phi}s, (\xi, v)\rangle = \langle s_{\Phi(e,q)}, \Phi_*(\xi, v)\rangle = \langle s_{\Phi(e,q)}, (\Phi_*\xi, v)\rangle = \langle \Phi(e, q), v\rangle.$$

Thus

$$(\widehat{\Phi}s)_{(e,q)} = \Phi(e, q), \tag{18.2}$$

where Φ in the right-hand side of (18.2) is considered as a horizontal form on $E \times M$.

We go on and compute the pull-back of the standard symplectic form:

$$\widehat{\Phi}\sigma = \widehat{\Phi}ds = d\widehat{\Phi}s = d\Phi.$$

Recall that differential of a form $\Lambda^1(N)$ can be evaluated by formula (11.15):

$$d\omega(W_1, W_2) = W_1\langle \omega, W_2\rangle - W_2\langle \omega, W_1\rangle - \langle \omega, [W_1, W_2]\rangle, \qquad W_1,\ W_2 \in \operatorname{Vec} N. \tag{18.3}$$

In our case $N = E \times M$ we take test vector fields of the form

$$W_i = (\xi_i, V_i) \in \operatorname{Vec}(E \times M), \qquad i = 1,\ 2,$$

where $\xi_i = \text{const} \in E$ are constant vertical vector fields and $V_i \in \operatorname{Vec} M$ are horizontal vector fields. By (18.3),

$$\begin{aligned} &d\Phi((\xi_1, V_1), (\xi_2, V_2)) \\ &\qquad = (\xi_1, V_1)\langle \Phi(\,\cdot\,,\cdot), V_2\rangle - (\xi_2, V_2)\langle \Phi(\,\cdot\,,\cdot), V_1\rangle - \langle \Phi(\,\cdot\,,\cdot), [V_1, V_2]\rangle \end{aligned}$$

since $[(\xi_1, V_1), (\xi_2, V_2)] = [V_1, V_2]$. Further,

$$\left((\xi_1, V_1)\langle \Phi(\,\cdot\,,\cdot), V_2\rangle\right)_{(e,\cdot)} = \left(\xi_1\langle \Phi(\,\cdot\,,\cdot), V_2\rangle + V_1\langle \Phi(\,\cdot\,,\cdot), V_2\rangle\right)_{(e,\cdot)}$$

and taking into account that Φ is linear w.r.t. e

$$= \langle \Phi_{\xi_1}, V_2\rangle + V_1\langle \Phi_e, V_2\rangle.$$

Consequently,

$$\begin{aligned} &d\Phi((\xi_1, V_1), (\xi_2, V_2))_{(e,\cdot)} = \\ &\qquad \langle \Phi_{\xi_1}, V_2\rangle - \langle \Phi_{\xi_2}, V_1\rangle + V_1\langle \Phi_e, V_2\rangle - V_2\langle \Phi_e, V_1\rangle - \langle \Phi_e, [V_1, V_2]\rangle. \end{aligned}$$

We denote the first two terms

$$\widetilde{\Phi}((\xi_1, V_1), (\xi_2, V_2)) = \langle \Phi_{\xi_1}, V_2 \rangle - \langle \Phi_{\xi_2}, V_1 \rangle,$$

and apply formula (18.3) to the horizontal form Φ_e:

$$d\Phi_e(V_1, V_2) = V_1 \langle \Phi_e, V_2 \rangle - V_2 \langle \Phi_e, V_1 \rangle - \langle \Phi_e, [V_1, V_2] \rangle.$$

Finally, we obtain the expression for pull-back of the symplectic form:

$$\widehat{\Phi}\sigma_{(e,\cdot)}((\xi_1, V_1), (\xi_2, V_2)) = \widetilde{\Phi}((\xi_1, V_1), (\xi_2, V_2)) + d\Phi_e(V_1, V_2), \tag{18.4}$$

i.e.,

$$\widehat{\Phi}\sigma_{(e,\cdot)} = \widetilde{\Phi} + d\Phi_e.$$

Remark. In the case of canonical coordinates we can take test vector fields $V_i = \frac{\partial}{\partial x_i}$, then it follows that $d\Phi_e = 0$.

18.1.4 Hamiltonian system on $E \times M$

Formula (18.4) describes the symplectic structure $\widehat{\Phi}\sigma$ on $E \times M$. Now we compute the Hamiltonian vector field corresponding to a Hamiltonian function

$$h \in C^\infty(E \times M).$$

One can consider h as a family of functions on M parametrized by vectors from E:

$$h_e = h(e, \cdot) \in C^\infty(M), \qquad e \in E.$$

Decompose the required Hamiltonian vector field into the sum of its vertical and horizontal parts:

$$\begin{aligned} \vec{h} &= X + Y, \\ X &= X(e, q) \in E, \\ Y &= Y(e, q) \in T_q M. \end{aligned}$$

By definition of a Hamiltonian field,

$$i_{X+Y} \widehat{\Phi}\sigma = -dh. \tag{18.5}$$

Transform the both sides of this equality:

$$\begin{aligned} -dh &= \underbrace{-\frac{\partial h}{\partial e}}_{\in E^*} - \underbrace{dh_e}_{\in T^*M}, \\ i_{X+Y}\big|_e \widehat{\Phi}\sigma &= i_{X+Y}\big|_e (\widetilde{\Phi} + d\Phi_e) = i_{(X,Y)}\big|_e \widetilde{\Phi} + i_{(X,Y)}\big|_e d\Phi_e \\ &= \underbrace{\langle \Phi_X, \cdot \rangle}_{\in T^*M} - \underbrace{\langle \Phi_\cdot, Y \rangle}_{\in E^*} + \underbrace{i_Y d\Phi_e}_{\in T^*M}. \end{aligned}$$

Now we equate the vertical parts of (18.5):

$$\langle \Phi \cdot, Y \rangle = \frac{\partial h}{\partial e}, \tag{18.6}$$

from this equation we can find the horizontal part of Y of the Hamiltonian field $\vec{h}$. Indeed, the linear isomorphism

$$\Phi(\,\cdot\,, q) \,:\, E \to T_q^* M$$

has a dual mapping

$$\Phi^*(\,\cdot\,, q) \,:\, T_q M \to E^*.$$

Then equation (18.6) can be written as

$$\Phi^*(\,\cdot\,, q) Y = \frac{\partial h}{\partial e}(e, q)$$

and then solved w.r.t. Y:

$$Y = \Phi^{*-1} \frac{\partial h}{\partial e}.$$

To find the vertical part X of the field $\vec{h}$, we equate the horizontal parts of (18.5):

$$\Phi_X + i_Y d\Phi_e = -dh_e,$$

rewrite as

$$\Phi_X = -i_Y d\Phi_e - dh_e,$$

and solve this equation w.r.t. X:

$$X = -\Phi^{-1}(i_Y d\Phi_e + dh_e).$$

Thus the Hamiltonian system on $E \times M$ corresponding to a Hamiltonian h has the form:

$$\begin{cases} \dot{q} = \Phi^{*-1} \dfrac{\partial h}{\partial e}, \\ \dot{e} = -\Phi^{-1}(i_{\dot{q}} d\Phi_e + dh_e). \end{cases} \tag{18.7}$$

Now we write this system using coordinates in the cotangent and tangent spaces (we do not require any coordinates on M).

Choose a basis in E:

$$E = \operatorname{span}(e_1, \dots, e_n),$$

so that vectors $u \in E$ are decomposed as

$$u = \sum_{i=1}^{n} u_i e_i.$$

Then

$$\Phi(u, \cdot) = \Phi\left(\sum_{i=1}^{n} u_i e_i, \cdot\right) = \sum_{i=1}^{n} u_i \omega_i,$$

where

$$\omega_i = \Phi_{e_i} \in \Lambda^1(M), \qquad i = 1, \dots, n,$$

are basis 1-forms on M. Further, the wedge products

$$\omega_i \wedge \omega_j \in \Lambda^2(M), \qquad 1 \le i < j \le n,$$

form a basis in the space $\Lambda^2(M)$ of 2-forms on M. Decompose the differentials in this basis:

$$d\omega_k = \sum_{1 \le i < j \le n} c_{ij}^k \, \omega_i \wedge \omega_j = \sum_{i,j=1}^{n} \frac{1}{2} c_{ij}^k \, \omega_i \wedge \omega_j,$$

where coefficients are smooth functions

$$c_{ij}^k \in C^\infty(M), \qquad i, \ j, \ k = 1, \dots, n,$$

skew-symmetric w.r.t. lower indices:

$$c_{ij}^k = -c_{ji}^k.$$

The coefficients c_{ij}^k are called structural constants (although, in general, they are not constant). We explain the name and give a simple recipe for computing them below in Proposition 18.1.

Choose a frame in $T_q M$ dual to the frame $\omega_1, \dots, \omega_n$:

$$V_1, \dots, V_n \in \operatorname{Vec} M,$$
$$\langle \omega_i, V_j \rangle = \delta_{ij}, \qquad i, \ j = 1, \dots, n.$$

Now we compute our Hamiltonian system (18.7) in the coordinates introduced. The Hamiltonian function has the form

$$h \in C^\infty(\mathbb{R}^n \times M),$$
$$h = h(u_1, \dots, u_n, q), \qquad (u_1, \dots, u_n) \in \mathbb{R}^n, \quad q \in M.$$

We have

$$\langle \Phi^*(V_i), e_j \rangle = \langle \Phi_{e_j}, V_i \rangle = \langle \omega_j, V_i \rangle = \delta_{ij},$$

thus

$$\Phi^*(V_i) = e_i^* = \begin{pmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{pmatrix},$$

the only unit is the i-th component. Consequently, the horizontal part of the field $\vec{h}$ decomposes along the basis horizontal fields as follows:

$$Y = \sum_{i=1}^{n} \frac{\partial h}{\partial u_i} V_i.$$

Consider the vertical part of $\vec{h}$:

$$X = -\Phi^{-1}(i_Y \, d\Phi_u + dh_u).$$

The second term is easily computed since

$$dh_u = \sum_{i=1}^{n} (V_i h_u) \omega_i,$$

this decomposition is immediately checked on basis vector fields V_i. And the first term has the form

$$-\Phi^{-1} i_Y \, d\Phi_u = \sum_{i,j,k=1}^{n} \frac{1}{2} u_k c_{ij}^k \left(\frac{\partial h}{\partial u_j} \frac{\partial}{\partial u_i} - \frac{\partial h}{\partial u_i} \frac{\partial}{\partial u_j} \right),$$

we leave this as an exercise for the reader.

Finally, the Hamiltonian system in the moving frames $(V_1, \dots, V_n)$ and $(\omega_1, \dots, \omega_n)$ reads:

$$\begin{cases} \dot{q} = \displaystyle\sum_{i=1}^{n} \frac{\partial h}{\partial u_i} V_i, \\ \dot{u}_i = -V_i h_u + \displaystyle\sum_{j,k=1}^{n} u_k c_{ij}^k \frac{\partial h}{\partial u_j}, \qquad i = 1, \dots, n. \end{cases}$$

Remark. This system becomes especially simple (triangular) when the Hamiltonian does not depend upon the point in the base:

$$\frac{\partial h}{\partial q} = 0.$$

The vertical subsystem simplifies even more when

$$c_{ij}^k = \text{const}, \qquad i, \ j, \ k = 1, \dots, n.$$

Both these conditions are satisfied for invariant problems on Lie groups discussed in subsequent sections.

The structural constants c_{ij}^k can easily be expressed in terms of Lie brackets of basis vector fields.

Proposition 18.1. *Let the frame of vector fields* $V_1, \dots, V_n \in \operatorname{Vec} M$ *be dual to the frame of 1-forms* $\omega_1, \dots, \omega_n \in \Lambda^1(M)$*:*

$$\langle \omega_i, V_j \rangle = \delta_{ij}, \qquad i, \ j = 1, \dots, n.$$

Then

$$d\omega_k = \sum_{i,j=1}^{n} \frac{1}{2} c_{ij}^k \, \omega_i \wedge \omega_j, \qquad k = 1, \dots, n,$$

if and only if

$$[V_i, V_j] = - \sum_{k=1}^{n} c_{ij}^k \, X_k, \qquad i, \ j = 1, \dots, n.$$

Proof. The equality for $d\omega_k$ can be written as

$$\langle d\omega_k, (V_i, V_j) \rangle = c_{ij}^k, \qquad i, \ j, \ k = 1, \dots, n.$$

The left-hand side is computed by formula (18.3):

$$\langle d\omega_k, (V_i, V_j) \rangle = \underbrace{V_i \langle \omega_k, V_j \rangle}_{=0} - \underbrace{V_j \langle \omega_k, V_i \rangle}_{=0} - \langle \omega_k, [V_i, V_j] \rangle,$$

and the statement follows. □

If the coefficients c_{ij}^k are constant, then the vector fields $V_1, \dots, V_n$ span a finite-dimensional Lie algebra, and the numbers c_{ij}^k are called *structural constants* of this Lie algebra. As we mentioned above, for general vector fields $c_{ij}^k \not\equiv \text{const}$.

18.2 Lie groups

State spaces for many interesting problems in geometry, mechanics, and applications are often not just smooth manifolds but Lie groups, in particular, groups of transformations. A manifold with a group structure is called a *Lie group* if the group operations are smooth. The cotangent bundle of a Lie group has a natural trivialization. We develop an approach of the previous sections and study optimal control problems on Lie groups.

18.2.1 Examples of Lie groups

The most important examples of Lie groups are given by groups of linear transformations of finite-dimensional vector spaces.

The group of all nondegenerate linear transformations of $\mathbb{R}^n$ is called the general linear group:

$$\mathrm{GL}(n) = \{ A : \mathbb{R}^n \to \mathbb{R}^n \mid \det A \neq 0 \}.$$

Linear volume-preserving transformations of $\mathbb{R}^n$ form the special linear group:

$$\mathrm{SL}(n) = \{A : \mathbb{R}^n \to \mathbb{R}^n \mid \det A = 1\}.$$

Another notation for these groups is respectively $\mathrm{GL}(\mathbb{R}^n)$ and $\mathrm{SL}(\mathbb{R}^n)$. Orthogonal group is formed by linear transformations preserving Euclidean structure:

$$\mathrm{O}(n) = \{A : \mathbb{R}^n \to \mathbb{R}^n \mid A^* A = \mathrm{Id}\},$$

and orthogonal orientation-preserving transformations form the special orthogonal group:

$$\mathrm{SO}(n) = \{A : \mathbb{R}^n \to \mathbb{R}^n \mid A^* A = \mathrm{Id},\ \det A = 1\}.$$

One can also consider the complex and Hermitian versions of these groups:

$$\mathrm{GL}(\mathbb{C}^n), \quad \mathrm{SL}(\mathbb{C}^n), \quad \mathrm{U}(n), \quad \mathrm{SU}(n).$$

Each of these groups can be realized as a subgroup of the corresponding real group commuting with multiplication by the imaginary unit:

$$\mathrm{GL}(\mathbb{C}^n) \subset \mathrm{GL}(\mathbb{R}^{2n}), \quad \mathrm{SL}(\mathbb{C}^n) \subset \mathrm{SL}(\mathbb{R}^{2n}), \quad \mathrm{U}(n) \subset \mathrm{O}(2n), \quad \mathrm{SU}(n) \subset \mathrm{SO}(2n).$$

Lie groups of linear transformations are called linear Lie groups. These groups often appear as a state space of a control system: e.g., $\mathrm{SO}(n)$ arises in the study of rotating configurations. For such systems, one can consider, as usual, the problems of controllability and optimal control.

18.2.2 Lie's theorem for linear Lie groups

Consider a control system of the form

$$\dot{X} = XA, \qquad X \in M = \mathrm{GL}(N), \quad A \in \mathcal{A} \subset \mathrm{gl}(N), \tag{18.8}$$

where $\mathcal{A}$ is an arbitrary subset of $\mathrm{gl}(N)$, the space of all real $N \times N$ matrices. We compute orbits of this system. Systems of the form (18.8) are called *left-invariant*: they are preserved by multiplication from the left by any constant matrix $Y \in \mathrm{GL}(N)$.

Notice that the ODE with a constant matrix A

$$\dot{X} = XA$$

is solved by the matrix exponential:

$$X(t) = X(0)e^{tA}.$$

Lie bracket of left-invariant vector fields is left-invariant as well:

$$[XA, XB] = X[A, B], \tag{18.9}$$

this follows easily from the coordinate expression for commutator (exercise).

Remark. Instead of left-invariant systems $\dot{X} = XA$, we can consider *right-invariant* ones: $\dot{X} = CX$. These forms are equivalent and transformed one into another by the inverse of matrix. Although, the Lie bracket for right-invariant vector fields is

$$[CX, DX] = [D, C]X,$$

which is less convenient than (18.9).

Return to control system (18.8). By the Orbit Theorem, the orbit through identity $\mathcal{O}_{\mathrm{Id}}(\mathcal{A})$ is an immersed submanifold of $\mathrm{GL}(N)$. Moreover, by definition, the orbit admits the representation via composition of flows:

$$\mathcal{O}_{\mathrm{Id}}(\mathcal{A}) = \{\mathrm{Id} \circ e^{t_1 A_1} \circ \cdots \circ e^{t_k A_k} \mid t_i \in \mathbb{R},\ A_i \in \mathcal{A},\ k \in \mathbb{N}\}$$

thus via products of matrix exponentials

$$= \{e^{t_1 A_1} \cdot \cdots \cdot e^{t_k A_k} \mid t_i \in \mathbb{R},\ A_i \in \mathcal{A},\ k \in \mathbb{N}\}.$$

Consequently, the orbit $\mathcal{O}_{\mathrm{Id}}(\mathcal{A})$ is a subgroup of $\mathrm{GL}(N)$. Further, in the proof of the Orbit Theorem we showed that the point $q \circ e^{t_1 A_1} \circ \cdots \circ e^{t_k A_k}$ depends continuously on $(t_1, \dots, t_k)$ in the "strong" topology of the orbit, thus it depends smoothly.

To summarize, we showed that the orbit through identity has the following properties:

(1) $\mathcal{O}_{\mathrm{Id}}(\mathcal{A})$ is an immersed submanifold of $\mathrm{GL}(N)$,

(2) $\mathcal{O}_{\mathrm{Id}}(\mathcal{A})$ is a subgroup of $\mathrm{GL}(N)$,

(3) the group operations $(X, Y) \mapsto XY$, $X \mapsto X^{-1}$ in $\mathcal{O}_{\mathrm{Id}}(\mathcal{A})$ are smooth.

In other words, the orbit $\mathcal{O}_{\mathrm{Id}}(\mathcal{A})$ is a *Lie subgroup* of $\mathrm{GL}(N)$.

The tangent spaces to the orbit are easily computed via the analytic version of the Orbit theorem (system (18.8) is real analytic):

$$\begin{aligned} T_{\mathrm{Id}}\mathcal{O}_{\mathrm{Id}}(\mathcal{A}) &= \mathrm{Lie}(\mathcal{A}), \\ T_X \mathcal{O}_{\mathrm{Id}}(\mathcal{A}) &= X\, \mathrm{Lie}(\mathcal{A}). \end{aligned} \tag{18.10}$$

The orbit of the left-invariant system (18.8) through any point $X \in \mathrm{GL}(N)$ is obtained by left translation of the orbit through identity:

$$\mathcal{O}_X(\mathcal{A}) = \{X e^{t_1 A_1} \cdots e^{t_k A_k} \mid t_i \in \mathbb{R},\ A_i \in \mathcal{A},\ k \in \mathbb{N}\} = X\mathcal{O}_{\mathrm{Id}}(\mathcal{A}).$$

We considered before system (18.8) defined by an arbitrary subset $\mathcal{A} \subset \mathrm{gl}(N)$. Restricting to Lie subalgebras

$$\mathcal{A} = \mathrm{Lie}(\mathcal{A}) \subset \mathrm{gl}(N),$$

we see that the following proposition was proved: to any Lie subalgebra $\mathcal{A} \subset \mathrm{gl}(N)$, there corresponds a connected Lie subgroup $M \subset \mathrm{GL}(N)$ such that $T_{\mathrm{Id}} M = \mathcal{A}$. Here $M = \mathcal{O}_{\mathrm{Id}}\mathcal{A}$. Now we show that this correspondence is invertible.

Let M be a connected Lie subgroup of $\mathrm{GL}(N)$, i.e.:

(1) M is an immersed connected submanifold of $\mathrm{GL}(N)$,

(2) M is a group w.r.t. matrix product,

(3) the group operations $(X, Y) \mapsto XY$, $X \mapsto X^{-1}$ in M are smooth mappings.

Then $\mathrm{Id} \in M$. Consider the tangent space

$$T_{\mathrm{Id}} M = \left\{ A = \left. \frac{d}{dt} \right|_{t=0} \Gamma_t \mid \Gamma_t \in M, \ \Gamma_t \text{ smooth}, \ \Gamma_0 = \mathrm{Id} \right\}.$$

Since $M \subset \mathrm{GL}(N) \subset \mathrm{gl}(N)$, then

$$T_{\mathrm{Id}} M \subset \mathrm{gl}(N).$$

Further,

$$A \in T_{\mathrm{Id}} M, \ X \in M \quad \Rightarrow \quad XA \in T_X M$$

since

$$XA = \left. \frac{d}{dt} \right|_{t=0} X\Gamma_t,$$

the velocity of the curve $X\Gamma_t$, where $A = \dot{\Gamma}_0$. Consequently, for any $A \in T_{\mathrm{Id}} M$ the vector field XA is identically tangent to M. So the following control system is well-defined on M:

$$\dot{X} = XA, \qquad X \in M, \quad A \in T_{\mathrm{Id}} M.$$

This system has a full rank. Since the state space M is connected, it coincides with the orbit $\mathcal{O}_{\mathrm{Id}}$ of this system through identity. We have already computed the tangent space to the orbit of a left-invariant system, see (18.10), thus

$$T_{\mathrm{Id}} M = T_{\mathrm{Id}}(\mathcal{O}_{\mathrm{Id}}) = \mathrm{Lie}(T_{\mathrm{Id}} M).$$

That is, $T_{\mathrm{Id}} M$ is a Lie subalgebra of $\mathrm{gl}(N)$. We proved the following classical proposition.

Theorem 18.1 (Lie). *There exists a one-to-one correspondence between Lie subalgebras $\mathcal{A} \subset \mathrm{gl}(N)$ and connected Lie subgroups $M \subset \mathrm{GL}(N)$ such that $T_{\mathrm{Id}} M = \mathcal{A}$.*

We showed that Lie's theorem for linear Lie algebras and Lie groups follows from the Orbit Theorem: connected Lie subgroups are orbits of left-invariant systems defined by Lie subalgebras, and Lie subalgebras are tangent spaces to Lie subgroups at identity.

18.2.3 Abstract Lie groups

An abstract Lie group is an abstract smooth manifold (not considered embedded into any ambient space) which is simultaneously a group, with smooth group operations. There holds Ado's theorem [13] stating that any finite-dimensional Lie algebra is isomorphic to a Lie subalgebra of $\mathrm{gl}(N)$. A similar statement for Lie groups is not true: a Lie group can be represented as a Lie subgroup of $\mathrm{GL}(N)$ only locally, but, in general, not globally. Although, the major part of properties of linear Lie groups can be generalized for abstract Lie groups.

In particular, let M be a Lie group. For any point $q \in M$, the left product by q:

$$\bar{q} : M \to M, \qquad \bar{q}(x) = qx, \quad x \in M,$$

is a diffeomorphism of M. Any tangent vector

$$v \in T_{\mathrm{Id}} M$$

can be shifted to any point $q \in M$ by the left translation $\bar{q}$:

$$V(q) = \bar{q}_* v \in T_q M, \qquad q \in M,$$

thus giving rise to a left-invariant vector field on M:

$$V \in \mathrm{Vec}\, M, \quad \bar{q}_* V = V, \qquad q \in M.$$

There is a one-to-one correspondence between left-invariant vector fields on M and tangent vectors to M at identity:

$$V \mapsto V(\mathrm{Id}) = v.$$

Left translations in M preserve flows of left-invariant vector fields on M, thus flows of their commutators. Consequently, left-invariant vector fields on a Lie group M form a Lie algebra, called the Lie algebra of the Lie group M. The tangent space $T_{\mathrm{Id}} M$ is thus also a Lie algebra.

Then, similar to the linear case, one can prove Lie's theorem on one-to-one correspondence between Lie subgroups of a Lie group M and Lie subalgebras of its Lie algebra $\mathcal{A}$.

18.3 Hamiltonian systems on Lie groups

18.3.1 Trivialization of cotangent bundle of a Lie group

Let $M \subset \mathrm{GL}(N)$ be a Lie subgroup. Denote by $\mathcal{M}$ the corresponding Lie subalgebra $T_{\mathrm{Id}} M \subset \mathrm{gl}(N)$. The cotangent bundle of M admits a trivialization of the form

$$\Phi : \mathcal{M}^* \times M \to T^* M,$$

where $\mathcal{M}^*$ is the dual space to the Lie algebra $\mathcal{M}$. We start from describing the dual mapping

$$\Phi^* : TM \to \mathcal{M} \times M.$$

Recall that $T_q M = q T_{\mathrm{Id}} M = q\mathcal{M}$ for any $q \in M$. We set

$$\Phi^* \ : \ qa \mapsto (a, q), \qquad a \in \mathcal{M}, \ q \in M, \ qa \in T_q M. \tag{18.11}$$

I.e., the value of a left-invariant vector field qa at a point q is mapped to the pair consisting of the value of this field at identity and the point q. Then the trivialization Φ has the form:

$$\Phi \ : \ (x, q) \mapsto \bar{x}_q, \qquad x \in \mathcal{M}^*, \ q \in M, \ \bar{x}_q \in T_q^* M, \tag{18.12}$$

where $\bar{x}$ is the left-invariant 1-form on M coinciding with x at identity:

$$\langle \bar{x}_q, qa \rangle \ \stackrel{\text{def}}{=} \ \langle x, a \rangle.$$

18.3.2 Hamiltonian system on $\mathcal{M}^* \times M$

The Hamiltonian system corresponding to a Hamiltonian

$$h = h(x, q) \in C^\infty(\mathcal{M}^* \times M)$$

was computed in Section 18.1, see (18.7):

$$\begin{cases} \dot{q} = \Phi^{-1*} \dfrac{\partial h}{\partial x}, \\ \dot{x} = -\Phi^{-1}(dh_x + i_{\dot{q}} d\Phi_x). \end{cases} \tag{18.13}$$

Taking into account definition (18.11) of Φ^*, we can write the first equation as follows:

$$\dot{q} = q \frac{\partial h}{\partial x}.$$

Here $\frac{\partial h}{\partial x}$ is the vertical part of $dh \in \Lambda^1(\mathcal{M}^* \times M)$, i.e.,

$$\frac{\partial h}{\partial x}(x, q) \in (\mathcal{M}^*)^* = \mathcal{M}, \qquad (x, q) \in \mathcal{M}^* \times M.$$

In order to find $\dot{x}$, compute the action of the differential $d\bar{x} = d\Phi_x$ on left-invariant vector fields by formula (11.15):

$$d\bar{x}(qa, qb) = (qa) \underbrace{\langle x, b \rangle}_{=\text{const}} - (qb) \underbrace{\langle x, a \rangle}_{=\text{const}} - \langle x, [a, b] \rangle = -\langle x, [a, b] \rangle.$$

Then

$$\begin{aligned} \Phi^{-1} i_{\dot{q}} d\Phi_x &= \Phi^{-1} i_{q \frac{\partial h}{\partial x}} d\bar{x} = -\left\langle x, \left[\frac{\partial h}{\partial x}, \ \cdot \ \right] \right\rangle = -\left\langle x, \left(\operatorname{ad} \frac{\partial h}{\partial x} \right) \cdot \ \right\rangle \\ &= -\left\langle \left(\operatorname{ad} \frac{\partial h}{\partial x} \right)^* x, \ \cdot \ \right\rangle = -\left(\operatorname{ad} \frac{\partial h}{\partial x} \right)^* x. \end{aligned}$$

So Hamiltonian system (18.13) takes the form:

$$\begin{cases} \dot{q} = q \dfrac{\partial h}{\partial x}, \\ \dot{x} = \left(\operatorname{ad} \dfrac{\partial h}{\partial x} \right)^* x - \Phi^{-1} dh_x. \end{cases} \tag{18.14}$$

Recall that dh_x is the horizontal part of dh, thus

$$(dh_x)_q \in T_q^* M, \qquad (x, q) \in \mathcal{M}^* \times M,$$

and

$$\Phi^{-1} dh_x \in \mathcal{M}^*.$$

System (18.14) describes the Hamiltonian system for an arbitrary Lie group and any Hamiltonian function h.

In the case of commutative Lie groups (which arise in trivialization of T^*M generated by local coordinates in M), the first term in the second equation (18.14) vanishes, and we obtain the usual form of Hamiltonian equations in canonical coordinates:

$$\begin{cases} \dot{q} = q \dfrac{\partial h}{\partial x}, \\ \dot{x} = -\Phi^{-1} dh_x. \end{cases}$$

On the contrary, if the Hamiltonian is left-invariant:

$$h = h(x),$$

then Hamiltonian system (18.14) becomes triangular:

$$\begin{cases} \dot{q} = q \dfrac{\partial h}{\partial x}, \\ \dot{x} = \left(\operatorname{ad} \dfrac{\partial h}{\partial x} \right)^* x. \end{cases} \tag{18.15}$$

Here the second equation does not contain q. So in left-invariant control problems, where the Hamiltonian h of PMP is left-invariant, one can solve the equation for vertical coordinates x independently, and then pass to the horizontal equation for q.

18.3.3 Compact Lie groups

The Hamiltonian system (18.15) simplifies even more in the case of compact Lie groups.

Let M be a compact Lie subgroup of $\mathrm{GL}(N)$. Then M can be considered as a Lie subgroup of $\mathrm{O}(N)$. Indeed, one can choose a Euclidean structure $\langle \cdot , \cdot \rangle$ in $\mathbb{R}^N$ invariant w.r.t. all transformations from M:

$$\langle Av, Aw \rangle = \langle v, w \rangle, \qquad v,\ w \in \mathbb{R}^N, \quad A \in M \subset \mathrm{GL}(N),$$

Such a structure can be obtained from any Euclidean structure $g(\,\cdot\,,\cdot\,)$ on $\mathbb{R}^N$ by averaging over $A \in M$ using a volume form $\omega_1 \wedge \ldots \wedge \omega_n$, where ω_i are basis left-invariant forms on M:

$$\langle v, w\rangle = \int_M \gamma_{v,w}\ \omega_1 \wedge \ldots \wedge \omega_n,$$
$$\gamma_{v,w}(A) = g(Av, Aw), \qquad A \in M.$$

So we will assume that elements of M are orthogonal $N \times N$ matrices, and the tangent space to M at identity consists of skew-symmetric matrices:

$$\mathcal{M} = T_{\mathrm{Id}} M \subset T_{\mathrm{Id}}\,\mathrm{O}(N) = \mathrm{so}(N) = \{a : \mathbb{R}^N \to \mathbb{R}^N \mid a^* + a = 0\}.$$

There is an invariant scalar product on $\mathrm{so}(N)$ defined as follows:

$$\langle a, b\rangle = -\operatorname{tr} ab, \qquad a,\ b \in \mathrm{so}(N).$$

This product is invariant in the sense that

$$\langle e^{t\,\mathrm{ad}\,c} a, e^{t\,\mathrm{ad}\,c} b\rangle = \langle a, b\rangle, \qquad a,\ b,\ c \in \mathrm{so}(N), \quad t \in \mathbb{R}, \tag{18.16}$$

i.e., the operator

$$\mathrm{Ad}\, e^{tc} = e^{t\,\mathrm{ad}\,c} : \mathrm{so}(N) \to \mathrm{so}(N)$$

is orthogonal w.r.t. this product. Equality (18.16) is a corollary of the invariance of trace:

$$\begin{aligned}\langle e^{t\,\mathrm{ad}\,c} a, e^{t\,\mathrm{ad}\,c} b\rangle &= \langle (\mathrm{Ad}\, e^{tc})a, (\mathrm{Ad}\, e^{tc})b\rangle = \langle e^{tc} a e^{-tc}, e^{tc} b e^{-tc}\rangle \\ &= -\operatorname{tr}(e^{tc} a e^{-tc} e^{tc} b e^{-tc}) = -\operatorname{tr}(e^{tc} ab e^{-tc}) = -\operatorname{tr}(ab) \\ &= \langle a, b\rangle.\end{aligned}$$

The sign minus in the definition of the invariant scalar product on $\mathrm{so}(N)$ provides positive-definiteness of the product. This can be easily seen in coordinates: if

$$a = (a_{ij}), \qquad b = (b_{ij}) \in \mathrm{so}(N),$$
$$a_{ij} = -a_{ji}, \qquad b_{ij} = -b_{ji}, \qquad i,\ j = 1, \ldots, N,$$

then

$$-\operatorname{tr}(ab) = -\sum_{i,j=1}^{N} a_{ij} b_{ji} = \sum_{i,j=1}^{N} a_{ij} b_{ij}.$$

The norm on $\mathrm{so}(N)$ is naturally defined:

$$|a| = \sqrt{\langle a, a\rangle}, \qquad a \in \mathrm{so}(N).$$

The infinitesimal version of the invariance property (18.16) is easily obtained by differentiation at $t = 0$:

$$\langle [c, a], b\rangle + \langle a, [c, b]\rangle = 0, \qquad a,\ b,\ c \in \mathrm{so}(N). \tag{18.17}$$

That is, all operators

$$\operatorname{ad} c \,:\, \mathrm{so}(N) \to \mathrm{so}(N), \qquad c \in \mathrm{so}(N),$$

are skew-symmetric w.r.t. the invariant scalar product. Equality (18.17) is a multidimensional generalization of a property of vector and scalar products in $\mathbb{R}^3 \cong \mathrm{so}(3)$.

Since $\mathcal{M} \subset \mathrm{so}(N)$, there is an invariant scalar product in the Lie algebra $\mathcal{M}$. Then the dual space $\mathcal{M}^*$ can be identified with the Lie algebra $\mathcal{M}$ via the scalar product $\langle \cdot , \cdot \rangle$:

$$\mathcal{M} \to \mathcal{M}^*, \qquad a \mapsto \langle a, \cdot \rangle.$$

In terms of this identification, the operator $(\operatorname{ad} a)^*$, $a \in \mathcal{M}$, takes the form:

$$(\operatorname{ad} a)^* \,:\, \mathcal{M} \to \mathcal{M}, \qquad (\operatorname{ad} a)^* = -\operatorname{ad} a.$$

In the case of a compact Lie group M, Hamiltonian system (18.15) for an invariant Hamiltonian $h = h(a)$ becomes defined on $\mathcal{M} \times M$ and reads

$$\begin{cases} \dot q = q \dfrac{\partial h}{\partial a}, \\ \dot a = \left[a, \dfrac{\partial h}{\partial a} \right]. \end{cases} \tag{18.18}$$

Now we apply this formula for solving several geometric optimal control problems.

Chapter 19

Examples of optimal control problems on compact Lie groups

19.1 Riemannian problem

Let M be a compact Lie group. The invariant scalar product $\langle \cdot, \cdot \rangle$ in the Lie algebra $\mathcal{M} = T_{\mathrm{Id}} M$ defines a left-invariant Riemannian structure on M:

$$\langle qu, qv \rangle_q \stackrel{\text{def}}{=} \langle u, v \rangle, \qquad u,\ v \in \mathcal{M}, \quad q \in M, \quad qu,\ qv \in T_q M.$$

So in every tangent space $T_q M$ there is a scalar product $\langle \cdot, \cdot \rangle_q$. For any Lipschitzian curve

$$q : [0,1] \to M$$

its Riemannian length is defined as integral of velocity:

$$l = \int_0^1 |\dot{q}(t)|\, dt, \qquad |\dot{q}| = \sqrt{\langle \dot{q}, \dot{q} \rangle}.$$

The problem is stated as follows: given any pair of points $q_0, q_1 \in M$, find the shortest curve in M that connects q_0 and q_1.

The corresponding optimal control problem is stated as follows:

$$\dot{q} = qu, \qquad q \in M, \quad u \in \mathcal{M}, \tag{19.1}$$

$$q(0) = q_0, \quad q(1) = q_1, \tag{19.2}$$

$$q_0,\ q_1 \in M \text{ fixed}, \tag{19.3}$$

$$l(u) = \int_0^1 |u(t)|\, dt \to \min.$$

First of all, we prove existence of optimal controls. Parametrizing trajectories of control system (19.1) by arc length, we see that the problem with unbounded admissible control $u \in \mathcal{M}$ on the fixed segment $t \in [0, 1]$ is equivalent to the problem with the compact space of control parameters $U = \{|u| = 1\}$ and free terminal time. Obviously, afterwards we can extend the set of control parameters to $U = \{|u| \leq 1\}$ so that the set of admissible velocities $f_U(q)$ become convex. Then Filippov's theorem implies existence of optimal controls in the problem obtained, thus in the initial one as well.

By Cauchy-Schwartz inequality,

$$(l(u))^2 = \left(\int_0^1 |u(t)|\, dt\right)^2 \leq \int_0^1 |u(t)|^2\, dt,$$

moreover, the equality occurs only if $|u(t)| \equiv \text{const}$. Consequently, the Riemannian problem $l \to \min$ is equivalent to the problem

$$J(u) = \frac{1}{2}\int_0^1 |u(t)|^2\, dt \to \min. \tag{19.4}$$

The functional J is more convenient than l since J is smooth and its extremals are automatically curves with constant velocity. In the sequel we consider the problem with the functional J: (19.1)–(19.4). The Hamiltonian of PMP for this problem has the form:

$$h_u^\nu(a, q) = \langle \bar{a}_q, qu\rangle + \frac{\nu}{2}|u|^2 = \langle a, u\rangle + \frac{\nu}{2}|u|^2.$$

The maximality condition of PMP is:

$$h_{u(t)}^\nu(a(t), q(t)) = \max_{v \in \mathcal{M}}(\langle a(t), v\rangle + \frac{\nu}{2}|v|^2), \qquad \nu \leq 0.$$

(1) Abnormal case: $\nu = 0$.

The maximality condition implies that $a(t) \equiv 0$. This contradicts PMP since the pair (ν, a) should be nonzero. So there are no abnormal extremals.

(2) Normal case: $\nu = -1$.

The maximality condition gives $u(t) \equiv a(t)$, thus the maximized Hamiltonian is smooth:

$$H(a) = \frac{1}{2}|a|^2.$$

Notice that the Hamiltonian H is invariant (does not depend on q), which is a corollary of left-invariance of the problem.

Optimal trajectories are projections of solutions of the Hamiltonian system corresponding to H. This Hamiltonian system has the form (see (18.18)):

$$\begin{cases} \dot{q} = qa, \\ \dot{a} = [a, a] = 0. \end{cases}$$

Thus optimal trajectories are one-parameter subgroups in M:

$$q(t) = q_0 e^{ta}, \qquad a \in \mathcal{M},$$

recall that an optimal solution exists. In particular, for the case $q_0 = \mathrm{Id}$, we obtain that any point $q_1 \in M$ can be represented in the form

$$q_1 = e^a, \qquad a \in \mathcal{M}.$$

That is, any element q_1 in a compact Lie group M has a logarithm a in the Lie algebra $\mathcal{M}$.

19.2 A sub-Riemannian problem

Now we modify the previous problem. As before, we should find the shortest path between fixed points q_0, q_1 in a compact Lie group M. But now admissible velocities $\dot{q}$ are not free: they should be tangent to a left-invariant distribution (of corank 1) on M. That is, we define a left-invariant field of tangent hyperplanes on M, and $\dot{q}(t)$ should belong to the hyperplane attached at the point $q(t)$. A problem of finding shortest curves tangent to a given distribution is called a *sub-Riemannian problem*.

To state the problem as an optimal control one, choose any element $b \in \mathcal{M}$, $|b| = 1$. Then the set of admissible velocities at identity is the hyperplane

$$U = b^{\perp} = \{u \in \mathcal{M} \mid \langle u, b \rangle = 0\}.$$

Remark. In the case $M = \mathrm{SO}(3)$, this restriction on velocities means that we fix an axis b in a rigid body and allow only rotations of the body around any axis u orthogonal to b.

The optimal control problem is stated as follows.

$$\begin{aligned}
&\dot{q} = qu, \qquad q \in M, \quad u \in U,\\
&q(0) = q_0, \quad q(1) = q_1,\\
&q_0,\ q_1 \in M \text{ fixed},\\
&l(u) = \int_0^1 |u(t)|\,dt \to \min .
\end{aligned}$$

Similarly to the Riemannian problem, Filippov's theorem guarantees existence of optimal controls, and the length minimization problem is equivalent to the problem

$$J(u) = \frac{1}{2}\int_0^1 |u(t)|^2\,dt \to \min .$$

The Hamiltonian of PMP is the same as in the previous problem:

$$h_u^{\nu}(a, q) = \langle a, u \rangle + \frac{\nu}{2}|u|^2,$$

but the maximality condition differs since now the set U is smaller:

$$h^{\nu}_{u(t)}(a(t), q(t)) = \max_{v \in b^{\perp}}(\langle a(t), v\rangle + \frac{\nu}{2}|v|^2).$$

Consider first the normal case: $\nu = -1$. Then the Lagrange multipliers rule implies that the maximum

$$\max_{v \in b^{\perp}} h^{-1}_v(a, q)$$

is attained at the vector

$$v_{\max} = a - \langle a, b\rangle b,$$

the orthogonal projection of a to $U = b^{\perp}$. The maximized Hamiltonian is smooth:

$$H(a) = \frac{1}{2}(|a|^2 - \langle a, b\rangle^2),$$

and the Hamiltonian system for normal extremals reads as follows:

$$\begin{cases} \dot{q} = q(a - \langle a, b\rangle b), \\ \dot{a} = \langle a, b\rangle [b, a]. \end{cases}$$

The second equation has an integral of the form

$$\langle a, b\rangle = \text{const},$$

this is easily verified by differentiation w.r.t. this equation:

$$\frac{d}{dt}\langle a, b\rangle = \langle a, b\rangle\langle [b, a], b\rangle$$

by invariance of the scalar product

$$= -\langle a, b\rangle\langle a, [b, b]\rangle = 0.$$

Consequently, the equation for a can be written as

$$\dot{a} = \langle a_0, b\rangle [b, a] = \operatorname{ad}(\langle a_0, b\rangle b)a,$$

where $a_0 = a(0)$. This linear ODE is easily solved:

$$a(t) = e^{t \operatorname{ad}(\langle a_0, b\rangle b)} a_0.$$

Now consider the equation for q:

$$\dot{q} = q\left(e^{t \operatorname{ad}(\langle a_0, b\rangle b)} a_0 - \langle a_0, b\rangle b\right)$$

since $e^{t \operatorname{ad}(\langle a_0, b\rangle b)} b = b$

$$= q e^{t \operatorname{ad}(\langle a_0, b\rangle b)}(a_0 - \langle a_0, b\rangle b). \tag{19.5}$$

This ODE can be solved with the help of the Variations formula. Indeed, we have (see (2.29)):

$$e^{t(f+g)} = \overrightarrow{\exp} \int_0^t e^{\tau \operatorname{ad} f} g \, d\tau \circ e^{tf},$$

i.e.,

$$\overrightarrow{\exp} \int_0^t e^{\tau \operatorname{ad} f} g \, d\tau = e^{t(f+g)} \circ e^{-tf} \tag{19.6}$$

for any vector fields f and g. Taking

$$f = \langle a_0, b \rangle b, \qquad g = a_0 - \langle a_0, b \rangle b,$$

we solve ODE (19.5):

$$q(t) = q_0 \, e^{t a_0} \, e^{-t \langle a_0, b \rangle b}. \tag{19.7}$$

Consequently, normal trajectories are products of two one-parameter subgroups.

Consider the abnormal case: $\nu = 0$. The Hamiltonian

$$h_u^0(a, q) = \langle a, u \rangle, \qquad u \perp b,$$

attains maximum only if

$$a(t) = \alpha(t) b, \qquad \alpha(t) \in \mathbb{R}. \tag{19.8}$$

But the second equation of the Hamiltonian system reads

$$\dot{a} = [a, u], \tag{19.9}$$

thus

$$\langle \dot{a}, a \rangle = \langle [a, u], a \rangle = -\langle u, [a, a] \rangle = 0.$$

That is, $\dot{a} \perp a$. In combination with (19.8) this means that

$$a(t) = \text{const} = \alpha b, \qquad \alpha \neq 0, \quad \alpha \in \mathbb{R}. \tag{19.10}$$

Notice that $\alpha \neq 0$ since the pair $(\nu, a(t))$ should be nonzero. Equalities (19.9) and (19.10) imply that abnormal extremal controls $u(t)$ satisfy the relation

$$[u(t), b] = 0.$$

That is, $u(t)$ belong to the Lie subalgebra

$$H_b = \{ c \in \mathcal{M} \mid [c, b] = 0 \} \subset \mathcal{M}.$$

For generic $b \in \mathcal{M}$ the subalgebra H_b is a Cartan subalgebra of $\mathcal{M}$, thus H_b is Abelian. In this case the first equation of the Hamiltonian system

$$\dot{q} = qu$$

contains only mutually commuting controls:

$$u(\tau) \in H_b \quad \Rightarrow \quad [u(\tau_1), u(\tau_2)] = 0,$$

and the equation is easily solved:

$$q(t) = q_0 e^{\int_0^t u(\tau)\, d\tau}. \tag{19.11}$$

Conversely, any trajectory of the form (19.11) with $u(\tau) \in H_b$, $\tau \in [0,t]$ is abnormal: it is a projection of abnormal extremal $(q(t), a(t))$ with $a(t) = \alpha b$ for any $\alpha \neq 0$.

We can give an elementary explanation of the argument on Cartan subalgebra in the case $\mathcal{M} = \mathrm{so}(n)$. Any skew-symmetric matrix $b \in \mathrm{so}(n)$ can be transformed by a change of coordinates to the diagonal form:

$$TbT^{-1} = \begin{pmatrix} i\alpha_1 & & & & \\ & -i\alpha_1 & & & \\ & & i\alpha_2 & & \\ & & & -i\alpha_2 & \\ & & & & \ddots \end{pmatrix} \tag{19.12}$$

for some $T \in \mathrm{GL}(n, \mathbb{C})$. But changes of coordinates (even complex) do not affect commutativity:

$$[c, b] = 0 \quad \Leftrightarrow \quad [TcT^{-1}, TbT^{-1}] = 0,$$

thus we can compute the subalgebra H_b using new coordinates:

$$H_b = T^{-1} H_{TbT^{-1}} T.$$

Generic skew-symmetric matrices $b \in \mathrm{so}(n)$ have distinct eigenvalues, thus for generic b the diagonal matrix (19.12) has distinct diagonal entries. For such b the Lie algebra $H_{TbT^{-1}}$ is easily found. Indeed, the commutator of a diagonal matrix

$$b = \begin{pmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_n \end{pmatrix}$$

with any matrix $c = (c_{ij})$ is computed as follows:

$$(\mathrm{ad}\, b)\; c = ((\beta_i - \beta_j) c_{ij}).$$

If a diagonal matrix b has simple spectrum:

$$\beta_i - \beta_j \neq 0, \qquad i \neq j,$$

then the Lie algebra H_b consists of diagonal matrices of the form (19.12), consequently H_b is Abelian.

So for a matrix $b \in \mathrm{so}(n)$ with mutually distinct eigenvalues (i.e., for generic $b \in \mathrm{so}(n)$) the Lie algebra $H_{TbT^{-1}}$ is Abelian, thus H_b is Abelian as well.

Returning to our sub-Riemannian problem, we conclude that we described all normal extremal curves (19.7), and described abnormal extremal curves (19.11) for generic $b \in \mathcal{M}$.

Exercise 19.1. Consider a more general sub-Riemannian problem stated in the same way as in this section, but with the space of control parameters $U \subset \mathcal{M}$ any linear subspace such that its orthogonal complement $U^\perp$ w.r.t. the invariant scalar product is a Lie subalgebra:

$$[U^\perp, U^\perp] \subset U^\perp. \tag{19.13}$$

Prove that normal extremals in this problem are products of two one-parameter groups (as in the corank one case considered above):

$$a_\perp = \mathrm{const}, \tag{19.14}$$

$$a_U(t) = e^{t \,\mathrm{ad}\, a_\perp}\, a_U^0, \qquad a_U^0 = a_U(0), \tag{19.15}$$

$$q(t) = q_0\, e^{ta}\, e^{-ta_\perp}, \tag{19.16}$$

where $a = a_U + a_\perp$ is the decomposition of a vector $a \in \mathcal{M}$ corresponding to the splitting $\mathcal{M} = U \oplus U^\perp$. We apply these results in the next problem.

19.3 Control of quantum systems

This section is based on the paper of U. Boscain, T. Chambrion, and J.-P. Gauthier [8].

Consider a three-level quantum system described by the Schrödinger equation (in a system of units such that $\hbar = 1$):

$$i\dot\psi = H\psi, \tag{19.17}$$

where $\psi : \mathbb{R} \to \mathbb{C}^3$, $\psi = (\psi_1, \psi_2, \psi_3)$, is a wave function and

$$H = \begin{pmatrix} E_1 & \Omega_1 & 0 \\ \overline{\Omega}_1 & E_2 & \Omega_2 \\ 0 & \overline{\Omega}_2 & E_3 \end{pmatrix} \tag{19.18}$$

is the Hamiltonian. Here $E_1 < E_2 < E_3$ are constant energy levels of the system and $\Omega_i : \mathbb{R} \to \mathbb{C}$ are controls describing the influence of the external pulsed field. The controls are connected to the physical parameters by $\Omega_j(t) = \mu_j \mathcal{F}_j(t)/2$, $j = 1, 2$, with $\mathcal{F}_j$ the external pulsed field and μ_j the couplings (intrinsic to the quantum system) that we have restricted to couple only levels j and $j+1$ by pairs.

This finite-dimensional problem can be thought as the reduction of an infinite-dimensional problem in the following way. We start with a Hamiltonian which is the sum of a drift-term H_0, plus a time dependent potential $V(t)$ (the control term, i.e., the lasers). The drift term is assumed to be diagonal, with eigenvalues (energy levels) $E_1 < E_2 < E_3 < \cdots$. Then in this spectral resolution of H_0, we assume the control term $V(t)$ to couple only the energy levels E_1, E_2 and E_2, E_3. The projected problem in the eigenspaces corresponding to E_1, E_2, E_3 is completely decoupled and is described by Hamiltonian (19.18).

The problem is stated as follows. Assume that at the initial instant $t = 0$ the state of the system lies in the eigenspace corresponding to the ground eigenvalue E_1. The goal is to determine controls Ω_1, Ω_2 that steer the system at the terminal instant $t = t_1$ to the eigenspace corresponding to E_3, requiring that these controls minimize the cost (energy in the following):

$$J = \int_0^T \left(|\Omega_1(t)|^2 + |\Omega_2(t)|^2\right) dt.$$

From the physical viewpoint, this problem may be considered either with arbitrary controls $\Omega_i(t) \in \mathbb{C}$, or with controls "in resonance":

$$\Omega_j(t) = u_j(t) e^{i(\omega_j t + \alpha_j)}, \qquad \omega_j = E_{j+1} - E_j, \tag{19.19}$$

$$u_j : \mathbb{R} \to \mathbb{R}, \qquad \alpha_j \in [-\pi, \pi], \quad j = 1,\ 2. \tag{19.20}$$

In the sequel we call this second problem of minimizing the energy J, which in this case reduces to

$$\int_0^{t_1} \left(u_1^2(t) + u_2^2(t)\right) dt, \tag{19.21}$$

the "real-resonant" problem. The first problem (with arbitrary complex controls) will be called the "general-complex" problem.

Since Hamiltonian (19.18) is self-adjoint: $H^* = H$, it follows that Schrödinger equation (19.17) is well-defined on the unit sphere

$$S_{\mathbb{C}} = S^5 = \left\{\psi = (\psi_1, \psi_2, \psi_3) \in \mathbb{C}^3 \mid |\psi|^2 = |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 = 1\right\}.$$

The source and the target, i.e., the initial and the terminal manifolds in the general-complex problem are respectively the circles

$$\mathcal{S}_{\mathbb{C}}^d = \{(e^{i\varphi}, 0, 0) \mid \varphi \in \mathbb{R}\}, \qquad \mathcal{T}_{\mathbb{C}}^d = \{(0, 0, e^{i\varphi}) \mid \varphi \in \mathbb{R}\}.$$

The meaning of the label $(^d)$ here will be clarified later.

Summing up, the general-complex problem is stated as follows:

$$\begin{aligned} &i\dot{\psi} = H\psi, \qquad \psi \in S^5, \quad \Omega_1,\ \Omega_2 \in \mathbb{C}, \\ &\psi(0) \in \mathcal{S}_{\mathbb{C}}^d, \qquad \psi(t_1) \in \mathcal{T}_{\mathbb{C}}^d, \\ &\int_0^{t_1} \left(|\Omega_1|^2 + |\Omega_2|^2\right) dt \to \min, \end{aligned}$$

with the Hamiltonian H defined by (19.18).

For the real-resonant case, the control system is (19.17) with Hamiltonian (19.18), admissible controls (19.19), (19.20), and cost (19.21). The natural state space, source, and target in this problem will be found later.

19.3.1 Elimination of the drift

We change variables in order to transfer the affine in control system (19.17), (19.18) to a system linear in control, both in the general-complex and real-resonant cases.

For $\Omega \in \mathbb{C}$, denote by $M_j(\Omega)$ and $N_j(\Omega)$ the $n \times n$ matrices:

$$
\begin{aligned}
M_j(\Omega)_{k,l} &= \delta_{j,k}\delta_{j+1,l}\Omega + \delta_{j+1,k}\delta_{j,l}\overline{\Omega} \\
N_j(\Omega)_{k,l} &= \delta_{j,k}\delta_{j+1,l}\Omega - \delta_{j+1,k},\delta_{j,l}\overline{\Omega}, \qquad j = 1,\ 2,
\end{aligned}
\tag{19.22}
$$

where δ is the Kronecker symbol: $\delta_{i,j} = 1$ if $i = j$, $\delta_{i,j} = 0$ if $i \neq j$. Let $\Delta = \operatorname{diag}(E_1, E_2, E_3)$, $\omega_j = E_{j+1} - E_j$, $j = 1, 2$. We will consider successively the general-complex problem:

$$
i\dot{\psi} = H\psi, \qquad H = \Delta + \sum_{j=1}^{2} M_j(\Omega_j), \quad \Omega_j \in \mathbb{C},
$$

and the real-resonant problem:

$$
i\dot{\psi} = H\psi, \qquad H = \Delta + \sum_{j=1}^{2} M_j(e^{i(\omega_j t + \alpha_j)} u_j), \qquad u_j, \alpha_j \in \mathbb{R}.
$$

In both cases, we first make the change of variable $\psi = e^{-it\Delta}\Lambda$ to get:

$$
i\dot{\Lambda} = \sum_{j=1}^{2} \left(\operatorname{Ad} e^{it\Delta} M_j(\Omega_j)\right)\Lambda = \sum_{j=1}^{2} M_j\left(e^{-it\omega_j}\Omega_j\right)\Lambda.
$$

The source $\mathcal{S}$ and the target $\mathcal{T}$ are preserved by this first change of coordinates.

The general-complex case.

In that case, we make the time-dependent cost preserving change of controls:

$$
e^{-it\omega_j}\Omega_j = i\tilde{\Omega}_j.
$$

Hence our problem becomes (after the change of notation $\Lambda \to \psi$, $\tilde{\Omega}_j \to u_j$):

$$
\dot{\psi} = \sum_{j=1}^{2} N_j(u_j)\psi = \widetilde{H}_{\mathbb{C}}\,\psi, \qquad u_j \in \mathbb{C}, \tag{19.23}
$$

$$
\int_0^{t_1} \left(|u_1|^2 + |u_2|^2\right)\, dt \to \min, \tag{19.24}
$$

$$
\psi(0) \in \mathcal{S}^d_{\mathbb{C}}, \qquad \psi(t_1) \in \mathcal{T}^d_{\mathbb{C}}, \tag{19.25}
$$

where

$$
\widetilde{H}_{\mathbb{C}} = \begin{pmatrix} 0 & u_1(t) & 0 \\ -\bar{u}_1(t) & 0 & u_2(t) \\ 0 & -\bar{u}_2(t) & 0 \end{pmatrix}. \tag{19.26}
$$

Notice that the matrices $N_j(1), N_j(i)$ generate su(3) as a Lie algebra. The cost and the relation between controls before and after elimination of the drift are:

$$J = \int_0^{t_1} \left(|u_1(t)|^2 + |u_2(t)|^2\right) dt, \tag{19.27}$$

$$\Omega_1(t) = u_1(t) e^{i[(E_2-E_1)t+\pi/2]}, \tag{19.28}$$

$$\Omega_2(t) = u_2(t) e^{i[(E_3-E_2)t+\pi/2]}. \tag{19.29}$$

The real-resonant case.

In this case $\Omega_j = u_j e^{i(\omega_j t + \alpha_j)}$, and we have:

$$i\dot{\Lambda} = \sum_{j=1}^{2} M_j \left(e^{i\alpha_j} u_j\right) \Lambda, \qquad u_j \in \mathbb{R}.$$

We make another diagonal, linear change of coordinates:

$$\Lambda = e^{iL}\phi, \qquad L = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3), \quad \lambda_j \in \mathbb{R},$$

which gives:

$$i\dot{\phi} = \sum_{j=1}^{2} M_j \left(e^{i(\alpha_j + \lambda_{j+1} - \lambda_j)} u_j\right) \phi.$$

Choosing the parameters λ_j such that $e^{i(\alpha_j+\lambda_{j+1}-\lambda_j)} = i$, we get:

$$\dot{\phi} = \sum_{j=1}^{2} N_j(u_j)\phi, \qquad u_j \in \mathbb{R}. \tag{19.30}$$

The source and the target are also preserved by this change of coordinates. Notice that the matrices $N_1(1)$, $N_2(1)$, in (19.30) generate so(3) as a Lie algebra. This means that the orbit of system (19.30) through the points $(\pm 1, 0, 0)$ is the real sphere S^2. Hence (by multiplication on the right by $e^{i\varphi}$), the orbit through the points $(\pm e^{i\varphi}, 0, 0)$ is the set $S^2 e^{i\varphi}$. Therefore (after the change of notation $\phi \to \psi$) the real-resonant problem is well-defined on the real sphere

$$S_{\mathbb{R}} = S^2 = \left\{\psi = (\psi_1, \psi_2, \psi_3) \in \mathbb{R}^3 \mid |\psi|^2 = \psi_1^2 + \psi_2^2 + \psi_3^2 = 1\right\},$$

as follows:

$$\dot{\psi} = \sum_{j=1}^{2} N_j(u_j)\psi = \tilde{H}_{\mathbb{R}}\,\psi, \qquad \psi \in S^2, \quad u_j \in \mathbb{R}, \tag{19.31}$$

$$\int_0^{t_1} \left(u_1^2 + u_2^2\right) dt \to \min, \tag{19.32}$$

$$\psi(0) \in \{(\pm 1, 0, 0)\}, \qquad \psi(t_1) \in \{(0, 0, \pm 1)\}, \tag{19.33}$$

where

$$\tilde{H}_{\mathbb{R}} = \begin{pmatrix} 0 & u_1(t) & 0 \\ -u_1(t) & 0 & u_2(t) \\ 0 & -u_2(t) & 0 \end{pmatrix}. \tag{19.34}$$

The cost is given again by formula (19.27) and the relation between controls before and after elimination of the drift is:

$$\begin{aligned} &\Omega_j(t) = u_j(t)\ e^{i(\omega_j t + \alpha_j)}, \qquad \omega_j = E_{j+1} - E_j, \\ &u_j : \mathbb{R} \to \mathbb{R}, \qquad \alpha_j \in [-\pi, \pi], \quad j = 1,\ 2. \end{aligned}$$

In the following we will use the labels $({}_{\mathbb{C}})$ and $({}_{\mathbb{R}})$ to indicate respectively the general-complex problem and the real-resonant one. When these labels are dropped in a formula, we mean that it is valid for both the real-resonant and the general-complex problem. With this notation:

$$\begin{aligned} \mathcal{S}^d_{\mathbb{C}} &= \{(e^{i\varphi}, 0, 0)\}, & \mathcal{T}^d_{\mathbb{C}} &= \{(0, 0, e^{i\varphi})\}, \\ \mathcal{S}^d_{\mathbb{R}} &= \{(\pm 1, 0, 0)\}, & \mathcal{T}^d_{\mathbb{R}} &= \{(0, 0, \pm 1)\}. \end{aligned}$$

19.3.2 Lifting of the problems

The problems (19.23)–(19.25) and (19.31)–(19.33) on the spheres $S_{\mathbb{C}} = S^5$ and $S_{\mathbb{R}} = S^2$ are naturally lifted to right-invariant problems on the Lie groups $M_{\mathbb{C}} = \mathrm{SU}(3)$ and $M_{\mathbb{R}} = \mathrm{SO}(3)$ respectively. The lifted systems read

$$\dot{q} = \widetilde{H} q, \qquad q \in M. \tag{19.35}$$

Denote the projections

$$\pi_{\mathbb{C}} : \mathrm{SU}(3) \to S^5, \qquad \pi_{\mathbb{R}} : \mathrm{SO}(3) \to S^2$$

both defined as

$$q \mapsto q \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

i.e., a matrix maps to its first column. We call problems (19.35) on the Lie groups M problems upstairs, and problems (19.23), (19.31) on the spheres S problems downstairs. We denote the problem upstairs by the label $({}^u)$ in parallel with the label $({}^d)$ for the problem downstairs.

Now we compute boundary conditions for the problems upstairs. Define the corresponding sources and targets:

$$\mathcal{S}^u = \pi^{-1}(\mathcal{S}^d), \qquad \mathcal{T}^u = \pi^{-1}(\mathcal{T}^d).$$

The source $\mathcal{S}^u_{\mathbb{C}}$ consists of all matrices $q \in \mathrm{SU}(3)$ with the first column in $\mathcal{S}^d_{\mathbb{C}}$:

$$q = \left(\begin{array}{c|c} \alpha & 0 \\ \hline 0 & A \end{array}\right), \qquad \alpha \in \mathrm{U}(1), \quad A \in \mathrm{U}(2), \quad \det q = 1.$$

We denote the subgroup of SU(3) consisting of such matrices by $\mathrm{S}(\mathrm{U}(1)\times\mathrm{U}(2))$. So the source upstairs in the general-complex problem is the subgroup

$$\mathcal{S}^u_{\mathbb{C}} = \mathrm{S}(\mathrm{U}(1)\times\mathrm{U}(2)).$$

Further, the matrix

$$\widehat{q} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

maps $\mathcal{S}^d_{\mathbb{C}}$ into $\mathcal{T}^d_{\mathbb{C}}$, thus

$$\mathcal{T}^u_{\mathbb{C}} = \widehat{q}\,\mathcal{S}^u_{\mathbb{C}} = \widehat{q}\ \mathrm{S}(\mathrm{U}(1)\times\mathrm{U}(2)).$$

Similarly, in the real case the source upstairs is

$$\mathcal{S}^u_{\mathbb{R}} = \mathrm{S}(\mathrm{O}(1)\times\mathrm{O}(2)),$$

the subgroup of SO(3) consisting of the matrices

$$q = \left(\begin{array}{c|c} \alpha & 0 \\ \hline 0 & A \end{array}\right), \qquad \alpha\in\mathrm{O}(1), \quad A\in\mathrm{O}(2), \quad \det q = 1,$$

and the target is

$$\mathcal{T}^u_{\mathbb{R}} = \widehat{q}\,\mathcal{S}^u_{\mathbb{R}} = \widehat{q}\ \mathrm{S}(\mathrm{O}(1)\times\mathrm{O}(2)).$$

Summing up, we state the lifted problems. The real problem upstairs reads:

$$\begin{aligned} &\dot q = \widetilde{H}_{\mathbb{R}}\, q = (u_1X_1 + u_2X_2)\, q, \qquad q\in\mathrm{SO}(3), \quad u_1,\ u_2\in\mathbb{R}, \\ &q(0)\in\mathcal{S}^u_{\mathbb{R}}, \qquad q(t_1)\in\mathcal{T}^u_{\mathbb{R}}, \\ &\int_0^{t_1}\left(u_1^2+u_2^2\right)\,dt\to\min, \end{aligned} \tag{19.36}$$

where

$$X_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad X_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}. \tag{19.37}$$

Notice that the real problem upstairs is a right-invariant sub-Riemannian problem on the compact Lie group SO(3) with a corank one set of control parameters $U\subset\mathrm{so}(3)$, i.e., a problem already considered in Section 19.2. We have

$$U = \mathrm{span}(X_1,X_2), \qquad U^\perp = \mathrm{span}(X_3), \qquad X_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Moreover, the frame (19.37) is orthonormal w.r.t. the invariant scalar product

$$\langle X, Y \rangle = -\frac{1}{2}\operatorname{tr}(XY), \qquad X,\ Y \in \mathrm{so}(3).$$

The complex problem upstairs is stated as follows:

$$\begin{aligned} &\dot q = \widetilde{H}_{\mathbb{C}}\, q = (u_1 X_1 + u_2 X_2 + u_3 Y_1 + u_4 Y_2)\, q, \qquad q \in \mathrm{SU}(3), \quad u_j \in \mathbb{R}, \\ &q(0) \in \mathcal{S}^u_{\mathbb{C}}, \qquad q(t_1) \in \mathcal{T}^u_{\mathbb{C}}, \\ &\int_0^{t_1} \left(u_1^2 + u_2^2 + u_3^2 + u_4^2\right) dt \to \min. \end{aligned} \tag{19.38}$$

Here X_1 and X_2 are given by (19.37) and

$$Y_1 = \begin{pmatrix} 0 & i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{pmatrix}.$$

The set of control parameters is

$$U = \operatorname{span}(X_1, X_2, Y_1, Y_2).$$

Notice that its orthogonal complement is

$$U^\perp = \operatorname{span}(Z_1, Z_2, Z_3, Z_4),$$

where

$$Z_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad Z_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix},$$

$$Z_3 = \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Z_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix},$$

and it is easy to check that $U^\perp$ is a Lie subalgebra. So the general-complex problem is of the form considered in Exercise 19.1. Again the distribution is right-invariant and the frame (X_1, X_2, Y_1, Y_2) is orthonormal for the metric

$$\langle X, Y \rangle = -\frac{1}{2}\operatorname{tr}(XY), \qquad X,\ Y \in \mathrm{su}(3).$$

The problems downstairs and upstairs are related as follows. For any trajectory upstairs $q(t) \in M$ satisfying the boundary conditions in M, its projection $\psi(t) = \pi(q(t)) \in S$ is a trajectory of the system downstairs satisfying the boundary conditions in S. And conversely, any trajectory downstairs $\psi(t)$ with the boundary conditions can be lifted to a trajectory upstairs $q(t)$ with the corresponding boundary conditions (such $q(t)$ is a matrix fundamental solution of the system downstairs). The cost for the problems downstairs and upstairs is the same. Thus solutions of the optimal control problems downstairs are projections of the solutions upstairs.

19.3.3 Controllability

The set of control parameters U in the both problems upstairs (19.38), (19.36) satisfies the property $[U,U] = U^\perp$, thus

$$U + [U,U] = \mathcal{M} = T_{\mathrm{Id}} M. \tag{19.39}$$

The systems upstairs have a full rank and are symmetric, thus they are completely controllable on the corresponding Lie groups SU(3), SO(3). Passing to the projections π, we obtain that the both systems downstairs (19.23), (19.31) are completely controllable on the corresponding spheres S^5, S^2.

19.3.4 Extremals

The problems upstairs are of the form considered in Section 19.2 and Exercise 19.1, but right-invariant not left-invariant ones. Thus normal extremals are given by formulas (19.14)–(19.16), where multiplication from the left is replaced by multiplication from the right:

$$\begin{aligned}
&a_\perp = \text{const},\\
&a_U(t) = e^{-t \, \mathrm{ad}\, a_\perp}\, a_U^0, \qquad a_U^0 = a_U(0),\\
&q(t) = e^{-t a_\perp}\, e^{t a}\, q_0,
\end{aligned} \tag{19.40}$$

for any $a_\perp \in U^\perp$, $a_U^0 \in U$. Geodesics are parametrized by arclength iff

$$\langle a_U^0, a_U^0 \rangle = 1. \tag{19.41}$$

Equality (19.39) means that in the problems upstairs, vector fields in the right-hand sides and their first order Lie brackets span the whole tangent space. Such control systems are called 2-generating. In Chapter 20 we prove that for such systems strictly abnormal geodesics (i.e., trajectories that are projections of abnormal extremals but not projections of normal ones) are not optimal, see the argument before Example 20.1. Thus we do not consider abnormal extremals in the sequel.

19.3.5 Transversality conditions

In order to select geodesics meeting the boundary conditions, we analyze transversality conditions upstairs.

Transversality conditions of PMP on T^*M corresponding to the boundary conditions

$$q(0) \in \mathcal{S}, \quad q(t_1) \in \mathcal{T}, \qquad \mathcal{S},\ \mathcal{T} \subset M,$$

read as follows:

$$\langle \lambda_0, T_{q(0)}\mathcal{S} \rangle = \langle \lambda_{t_1}, T_{q(t_1)}\mathcal{T} \rangle = 0. \tag{19.42}$$

Via trivialization (18.12) of T^*M, transversality conditions (19.42) are rewritten for the extremal $(x(t), q(t)) \in \mathcal{M}^* \times M$ in the form:

$$\left\langle x(0),\, q(0)^{-1}\, T_{q(0)}\mathcal{S} \right\rangle = \left\langle x(t_1),\, q(t_1)^{-1}\, T_{q(t_1)}\mathcal{T} \right\rangle = 0.$$

Here the brackets $\langle \cdot\,, \cdot \rangle$ denote action of a covector on a vector. The transversality conditions for the extremal $(a(t), q(t)) \in \mathcal{M} \times M$ read as follows:

$$\left\langle a(0),\, q(0)^{-1}\, T_{q(0)}\mathcal{S} \right\rangle = \left\langle a(t_1),\, q(t_1)^{-1}\, T_{q(t_1)}\mathcal{T} \right\rangle = 0,$$

where the brackets denote the invariant scalar product in $\mathcal{M}$.

For the right-invariant problem, transversality conditions are written in terms of right translations:

$$\left\langle a(0),\, (T_{q(0)}\mathcal{S})\, q(0)^{-1} \right\rangle = \left\langle a(t_1), (T_{q(t_1)}\mathcal{T})\, q(t_1)^{-1} \right\rangle = 0. \tag{19.43}$$

The following features of transversality conditions for our problems upstairs simplifies their analysis.

Lemma 19.1. (1) *Transversality conditions at the source are required only at the identity.*

(2) *Transversality conditions at the source imply transversality conditions at the target.*

Proof. Item (1) follows since the problem is right-invariant and the source $\mathcal{S}^u$ is a subgroup.

Item (2). Let $\lambda_t \in T^*_{q(t)}M$ be a normal extremal for the problem upstairs such that $q(0) = \mathrm{Id}$. We assume the transversality conditions at the source:

$$\langle \lambda_0, T_{\mathrm{Id}}\mathcal{S}^u \rangle = 0,$$

and prove the transversality conditions at the target:

$$\langle \lambda_{t_1}, T_{q(t_1)}\mathcal{T}^u \rangle = 0. \tag{19.44}$$

Notice first of all that since $q(t_1) \in \mathcal{T}^u = \widehat{q}\,\mathcal{S}^u$, then $\widehat{q}^{-1}q(t_1) \in \mathcal{S}^u$ and

$$\mathcal{T}^u = \widehat{q}\,\mathcal{S}^u = \widehat{q}\,(\widehat{q}^{-1}q(t_1))\mathcal{S}^u = q(t_1)\mathcal{S}^u.$$

Then transversality conditions at the target (19.44) read

$$\langle \lambda_{t_1}, T_{q(t_1)}\,(q(t_1)\mathcal{S}^u) \rangle = 0.$$

In order to complete the proof, we show that the function

$$I(t) = \langle \lambda_t, T_{q(t)}\,(q(t)\mathcal{S}^u) \rangle, \qquad t \in [0, t_1],$$

is constant. Denote the tangent space $S = T_{\mathrm{Id}}\mathcal{S}^u$. Then we have:

$$\begin{aligned} I(t) &= \langle \lambda_t, q(t)S \rangle = \left\langle x(t), q(t)Sq(t)^{-1} \right\rangle \\ &= \langle x(t), (\operatorname{Ad} q(t))S \rangle = \langle a(t), (\operatorname{Ad} q(t))S \rangle \\ &= \left\langle (\operatorname{Ad} e^{-ta_\perp})a(0), (\operatorname{Ad} e^{-ta_\perp})(\operatorname{Ad} e^{-ta(0)})S \right\rangle \end{aligned}$$

by invariance of the scalar product

$$
\begin{aligned}
&= \left\langle a(0), (\operatorname{Ad} e^{-ta(0)})S \right\rangle = \left\langle (\operatorname{Ad} e^{-ta(0)})a(0), S \right\rangle = \langle a(0), S\rangle \\
&= I(0).
\end{aligned}
$$

That is, $I(t) \equiv \text{const}$, and item (2) of this lemma follows. □

19.3.6 Optimal geodesics upstairs and downstairs

Similarly to N_1, N_2 (see formula (19.22)), let us define $N_{1,3}$ by:

$$
N_{1,3}(a_3 e^{i\theta_3}) = \begin{pmatrix} 0 & 0 & a_3 e^{i\theta_3} \\ 0 & 0 & 0 \\ -a_3 e^{-i\theta_3} & 0 & 0 \end{pmatrix}.
$$

Let us set, in the real-resonant case

$$
a_U^0 = a_1 N_1(1) + a_2 N_2(1), \qquad a_\perp = a_3 N_{1,3}(1).
$$

In the general complex case, set

$$
a_U^0 = N_1(a_1 e^{i\theta_1}) + N_2(a_2 e^{i\theta_2}), \qquad a_\perp = a_4 Z_3 + a_5 Z_4 + N_{1,3}(a_3 e^{i\theta_3}).
$$

Here $a_i \in \mathbb{R}$ and $\theta_i \in [-\pi, \pi]$.

The real-resonant case.

Proposition 19.1. *For the real-resonant problem, transversality condition at the identity in the source* $\langle a, T_{\mathrm{Id}}\mathcal{S}^u_{\mathbb{R}}\rangle = 0$ *means that* $a_2 = 0$.

Proof. We have

$$
T_{\mathrm{Id}}\mathcal{S}^u_{\mathbb{R}} = \left\{ \left(\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & 0 & -\beta \\ 0 & \beta & 0 \end{array} \right), \quad \beta \in \mathbb{R} \right\},
$$

thus the equation $\langle a, T_{\mathrm{Id}}\mathcal{S}^u_{\mathbb{R}}\rangle = 0$ is satisfied for every $\beta \in \mathbb{R}$ if and only if $a_2 = 0$. □

From Proposition 19.1 and condition (19.41), one gets the covectors to be used in formula (19.40):

$$
a^{\pm} = \begin{pmatrix} 0 & \pm 1 & a_3 \\ \mp 1 & 0 & 0 \\ -a_3 & 0 & 0 \end{pmatrix}. \tag{19.45}
$$

Proposition 19.2. *Geodesics* (19.40) *with the initial condition* $q(0) = \mathrm{Id}$ *and matrix* a *given by* (19.45) *reach the target* $\mathcal{T}^u_{\mathbb{R}}$ *for the smallest time (arclength)* $|t|$, *if and only if* $a_3 = \pm 1/\sqrt{3}$. *Moreover, the 4 geodesics (corresponding to* $a^{\pm}$ *and to the signs* $\pm$ *in* a_3*) have the same length and reach the target at the time*

$$t_1 = \frac{\sqrt{3}}{2}\pi.$$

Proof. Computing $q(t) = e^{-a_{\perp}t}e^{(a_{\perp}+a^0_U)t}$, with a given by formula (19.45), and recalling that

$$\psi(t) = q(t)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

one gets for the square of the third component of the wave function:

$$(\psi_3(t))^2 = \frac{\left(\cos(t\,a_3)\,\sin(t\,\gamma)\;a_3\,\gamma - \cos(t\,\gamma)\,\sin(t\,a_3)\,\gamma^2\right)^2}{\gamma^4}, \qquad (19.46)$$
$$\gamma = \sqrt{1+a_3^2}.$$

Then the following lemma completes the proof of this proposition. □

Lemma 19.2. *Set* $f_a = \cos(ta)\sin(t\sqrt{1+a^2})\frac{a}{\sqrt{1+a^2}} - \cos(t\sqrt{1+a^2})\sin(ta)$, *then* $|f_a| \leq 1$. *Moreover,* $|f_a| = 1$ *iff* $\frac{|a|}{\sqrt{1+a^2}} = \left|\frac{1}{2k} + \frac{k'}{k}\right| < 1$, $k \neq 0$ *and* $t = \frac{k\pi}{\sqrt{1+a^2}}$. *In particular, the smallest* $|t|$ *is obtained for* $k = \pm 1$, $a = \pm\frac{1}{\sqrt{3}}$, $t = \frac{\pm\pi\sqrt{3}}{2}$.

Proof. Set $\lambda = \frac{a}{\sqrt{1+a^2}}$, $\theta = t\sqrt{1+a^2}$, then:

$$\begin{aligned} f_a(t) &= \lambda\cos(\lambda\theta)\sin(\theta) - \cos(\theta)\sin(\lambda\theta) \\ &= \langle(\lambda\cos(\lambda\theta), \sin(\lambda\theta)), (\sin(\theta), -\cos(\theta))\rangle \\ &= \langle v_1, v_2\rangle. \end{aligned}$$

Both v_1, v_2 have norm ≤ 1 and $|f_a| \leq 1$. Hence, for $|f_a| = 1$, we must have $|v_1| = |v_2| = 1$, $v_1 = \pm v_2$. It follows that $\cos(\lambda\theta) = 0$ and $\cos(\theta) = \pm 1$. Hence $\theta = k\pi$, $\lambda\theta = \frac{\pi}{2} + k'\pi$, $\lambda = \frac{1}{2k} + \frac{k'}{k}$. Therefore, $\left|\frac{1}{2k} + \frac{k'}{k}\right| = \lambda < 1$. Conversely, choose k, k' meeting this condition, and $\theta = k\pi$. Then $\cos(\theta) = \pm 1$, $\sin(\lambda\theta) = \pm 1$, $f_a(t) = \pm 1$. Now, $|t| = \frac{k\pi}{\sqrt{1+a^2}}$, and the smallest $|t|$ is obtained for $k = \pm 1$ (if $k = 0$, $\theta = 0$ and $f_a(t) = 0$). Moreover, $\left|\frac{1}{2k} + \frac{k'}{k}\right| < 1$ is possible only for $(k, k') = (1, 0)$ or $(1, -1)$ or $(-1, 0)$ or $(-1, -1)$. In all cases, $|\lambda| = \frac{1}{2}$, $a = \pm\frac{1}{\sqrt{3}}$, and $t = \pm\frac{\pi\sqrt{3}}{2}$. □

Let us fix for instance the sign $-$ in (19.45) and $a_3 = +1/\sqrt{3}$. The expressions of the three components of the wave function are:

$$\psi_1(t) = \cos\left(\frac{t}{\sqrt{3}}\right)^3,$$
$$\psi_2(t) = \frac{\sqrt{3}}{2}\sin\left(\frac{2\,t}{\sqrt{3}}\right),$$
$$\psi_3(t) = -\sin\left(\frac{t}{\sqrt{3}}\right)^3.$$

Notice that this curve is not a circle on S^2.

Controls can be obtained with the following expressions:

$$u_1 = (\dot{q}q^{-1})_{1,2}, \qquad u_2 = (\dot{q}q^{-1})_{2,3}.$$

We get:

$$u_1(t) = -\cos\left(\frac{t}{\sqrt{3}}\right),$$
$$u_2(t) = \sin\left(\frac{t}{\sqrt{3}}\right).$$

Using conditions (19.19)–(19.20) (resonance hypothesis), we get for the external fields:

$$\Omega_1(t) = -\cos\left(t/\sqrt{3}\right)e^{i(\omega_1 t+\alpha_1)},$$
$$\Omega_2(t) = \sin\left(t/\sqrt{3}\right)e^{i(\omega_2 t+\alpha_2)}.$$

Notice that the phases $\alpha_1,\ \alpha_2$ are arbitrary.

The general-complex case.

Proposition 19.3. *For the general-complex problem, transversality condition at the identity in the source* $\langle a, T_{\mathrm{Id}}\mathcal{S}^u_{\mathbb{C}}\rangle = 0$ *means that* $a_2 = a_4 = a_5 = 0$.

Proof. We have:

$$T_{\mathrm{Id}}\mathcal{S}^u_{\mathbb{C}} = \left\{\left(\begin{array}{c|cc} i\alpha_1 & 0 & 0 \\ \hline 0 & i(\alpha_2-\alpha_1) & \beta_1+i\beta_2 \\ 0 & -\beta_1+i\beta_2 & -i\alpha_2 \end{array}\right),\quad \alpha_1,\ \alpha_2,\ \beta_1,\ \beta_2\in\mathbb{R}\right\}.$$

The equation $\langle a, T_{\mathrm{Id}}\mathcal{S}^u_{\mathbb{C}}\rangle = 0$ is satisfied for every $\alpha_1,\alpha_2,\beta_1,\beta_2\in\mathbb{R}$ if and only if $a_2 = a_4 = a_5 = 0$. $\square$

The covector to be used in formula (19.40) is then:

$$a^{(\theta_1,\theta_3)} = \begin{pmatrix} 0 & e^{i\theta_1} & a_3e^{i\theta_3} \\ -e^{-i\theta_1} & 0 & 0 \\ -a_3e^{-i\theta_3} & 0 & 0 \end{pmatrix}. \tag{19.47}$$

Proposition 19.4. *The geodesics* (19.40)*, with a given by formula* (19.47) *(for which* $q(0) = \mathrm{Id}$*), reach the target* $\mathcal{T}_{\mathbb{C}}^{u}$ *for the smallest time (arclength)* $|t|$*, if and only of* $a_3 = \pm 1/\sqrt{3}$*. Moreover, all the geodesics of the two parameter family corresponding to* $\theta_1, \theta_3 \in [-\pi, \pi]$*, have the same length:*

$$t_1 = \frac{\sqrt{3}}{2}\pi.$$

Proof. The explicit expression for $|\psi_3|^2$ is given by the right-hand side of formula (19.46). The conclusion follows as in the proof of Proposition 19.2. □

The expressions of the three components of the wave function and of optimal controls are:

$$\psi_1(t) = \cos\left(\frac{t}{\sqrt{3}}\right)^3,$$
$$\psi_2(t) = -\frac{\sqrt{3}}{2}\sin\left(\frac{2t}{\sqrt{3}}\right)e^{-i\theta_1},$$
$$\psi_3(t) = -\sin\left(\frac{t}{\sqrt{3}}\right)^3 e^{-i\theta_3},$$

and

$$u_1(t) = \cos\left(t/\sqrt{3}\right)e^{i\theta_1},$$
$$u_2(t) = -\sin\left(t/\sqrt{3}\right)e^{i(\theta_3-\theta_1)}.$$

Notice that all the geodesics of the family described by Proposition 19.4 have the same length as the 4 geodesics described by Proposition 19.2. This proves that the use of the complex Hamiltonian (19.26) instead of the real one (19.34) does not allow to reduce the cost (19.27). We obtain the following statement.

Proposition 19.5. *For the three-level problem with complex controls, optimality implies resonance. More precisely, controls* Ω_1*,* Ω_2 *are optimal if and only if they have the following form:*

$$\Omega_1(t) = \cos(t/\sqrt{3})e^{i[(E_2-E_1)t+\varphi_1]},$$
$$\Omega_2(t) = \sin(t/\sqrt{3})e^{i[(E_3-E_2)t+\varphi_2]}.$$

where φ_1, φ_2 *are two arbitrary phases. Here the final time* t_1 *is fixed in such a way sub-Riemannian geodesics are parametrized by arclength, and it is given by* $t_1 = \frac{\sqrt{3}}{2}\pi$*.*

19.4 A time-optimal problem on SO(3)

Consider a rigid body in $\mathbb{R}^3$. Assume that the body can rotate in a certain direction around some axis fixed in the body. At each instant of time, orientation of the body in $\mathbb{R}^3$ defines an orthogonal transformation $q \in \mathrm{SO}(3)$. We are interested in the length of the curve in $\mathrm{SO}(3)$ corresponding to the motion of the body. Choose a natural parameter (arc length) t, then the curve $q = q(t)$ satisfies the ODE

$$\dot{q} = qf,$$

where

$$f \in \mathrm{so}(3), \qquad |f| = 1,$$

is the unit vector of angular velocity corresponding to the fixed axis of rotation in the body. The curve is a one-parameter subgroup in $\mathrm{SO}(3)$:

$$q(t) = q(0)e^{tf},$$

and we obviously have no controllability on $\mathrm{SO}(3)$.

In order to extend possibilities of motion in $\mathrm{SO}(3)$, assume now that there are two linearly independent axes in the body:

$$f,\ g \in \mathrm{so}(3), \qquad |f| = |g| = 1, \qquad f \wedge g \neq 0,$$

and we can rotate the body around these axes in certain directions. Now we have a control system

$$\dot{q} = \begin{cases} qf \\ qg \end{cases},$$

which is controllable on $\mathrm{SO}(3)$:

$$\mathrm{Lie}(qf, qg) = \mathrm{span}(qf, qg, q[f, g]) = q\,\mathrm{so}(3) = T_q\,\mathrm{SO}(3).$$

In order to simplify notation, choose vectors

$$a,\ b \in \mathrm{so}(3)$$

such that

$$f = a + b, \qquad g = a - b.$$

Then the control system reads

$$\dot{q} = q(a \pm b).$$

We are interested in the shortest rotation of the body steering an initial orientation q_0 to a terminal orientation q_1. The corresponding optimal control problem is

$$q(0) = q(0), \qquad q(t_1) = q_1,$$
$$l = \int_0^{t_1} |\dot{q}|\, dt \to \min.$$

Since $|\dot{q}| = |a \pm b| = 1$, this problem is equivalent to the time-optimal problem:

$$t_1 \to \min.$$

Notice that

$$\langle a, b\rangle = \langle (f+g)/2, (f-g)/2\rangle = 0. \tag{19.48}$$

Moreover, by rescaling time we can assume that

$$|a| = 1. \tag{19.49}$$

Passing to convexification, we obtain the following final form of the problem:

$$\begin{aligned}
&\dot q = q(a+ub), \qquad u \in [-1,1], \quad q \in \mathrm{SO}(3),\\
&q(0) = q_0, \qquad q(t_1) = q_1,\\
&t_1 \to \min,
\end{aligned}$$

where $a, b \in \mathrm{so}(3)$ are given vectors that satisfy equalities (19.48), (19.49). Now we study this time-optimal problem.

By PMP, if a pair $(u(\cdot), q(\cdot))$ is optimal, then there exists a Lipschitzian curve $x(t) \in \mathrm{so}(3)$ such that:

$$\begin{aligned}
&\begin{cases} \dot q = q(a+u(t)b),\\ \dot x = [x, a+u(t)b],\end{cases}\\
&h_{u(t)}(x(t)) = \langle x(t), a+u(t)b\rangle = \max_{|v|\le 1}\langle x(t), a+vb\rangle \ge 0,
\end{aligned}$$

moreover,

$$h_{u(t)}(x(t)) = \mathrm{const}\,.$$

The maximality condition for the function

$$v \mapsto \langle x(t), a+vb\rangle = \langle x(t), a\rangle + v\langle x(t), b\rangle, \qquad v \in [-1,1],$$

is easily resolved if the *switching function*

$$x \mapsto \langle x, b\rangle, \qquad x \in \mathcal{M},$$

does not vanish at $x(t)$. Indeed, in this case optimal control can take only extremal values ± 1:

$$\langle x(t), b\rangle \ne 0 \quad \Rightarrow \quad u(t) = \mathrm{sgn}\langle x(t), b\rangle.$$

If the switching function has only isolated roots on some real segment, then the corresponding control $u(t)$ takes on this segment only extremal values. Moreover, the instants where $u(t)$ switches from one extremal value to another are isolated. Such a control is called bang-bang.

Now we study the structure of optimal controls. Take an arbitrary extremal with the curve $x(t)$ satisfying the initial condition

$$\langle x(0), b\rangle \ne 0.$$

Then the ODE

$$\dot{x} = [x, a \pm b], \qquad \pm = \operatorname{sgn}\langle x(0), b\rangle$$

is satisfied for $t > 0$ until the switching function $\langle x(t), b\rangle$ remains nonzero. Thus at such a segment of time

$$x(t) = e^{-t \operatorname{ad}(a \pm b)} x(0).$$

We study the switching function $\langle x(t), b\rangle$. Notice that its derivative does not depend upon control:

$$\frac{d}{dt}\langle x(t), b\rangle = \langle [x(t), a + u(t)b], b\rangle = -\langle x(t), [a, b]\rangle.$$

If the switching function vanishes:

$$\langle x(t), b\rangle = 0$$

at a point where

$$\langle x(t), [a, b]\rangle \neq 0,$$

then the corresponding control switches, i.e., changes its value from $+1$ to -1 or from -1 to $+1$. In order to study, what sequences of switchings of optimal controls are possible, it is convenient to introduce coordinates in the Lie algebra $\mathcal{M}$.

In view of equalities (19.48), (19.49), the Lie bracket $[a, b]$ satisfies the conditions

$$[a, b] \perp a, \qquad [a, b] \perp b, \qquad |[a, b]| = |b|,$$

this follows easily from properties of cross-product in $\mathbb{R}^3$. Thus we can choose an orthonormal basis:

$$\operatorname{so}(3) = \operatorname{span}(e_1, e_2, e_3)$$

such that

$$a = e_2, \qquad b = \nu e_3, \qquad [a, b] = \nu e_1, \qquad \nu > 0.$$

In this basis, switching points belong to the horizontal plane $\operatorname{span}(e_1, e_2)$.

Let $x(\tau_0)$ be a switching point, i.e., $t = \tau_0$ is a positive root of $\langle x(t), b\rangle$. Assume that at this point control switches from $+1$ to -1 (the case of switching from -1 to $+1$ is completely similar, we show this later). Then

$$\langle \dot{x}(\tau_0), b\rangle = -\langle \dot{x}(\tau_0), [a, b]\rangle \leq 0,$$

thus

$$\langle x(\tau_0), e_1\rangle \geq 0.$$

Further, since the Hamiltonian of PMP is nonnegative, then

$$h_{u(\tau_0)}(x(\tau_0)) = \langle x(\tau_0), a\rangle = \langle x(\tau_0), e_2\rangle \geq 0.$$

So the point $x(\tau_0)$ lies in the first quadrant of the plane $\operatorname{span}(e_1, e_2)$:

$$x(\tau_0) \in \operatorname{cone}(e_1, e_2).$$

Let $x(\tau_1)$ be the next switching point after τ_0. The control has the form

$$u(t) = \begin{cases} 1, & t \in [\tau_0 - \varepsilon, \tau_0], \\ -1, & t \in [\tau_0, \tau_1], \end{cases}$$

and the curve $x(t)$ between the switchings is an arc of the circle obtained by rotation of the point $x(\tau_0)$ around the vector $a - b = e_2 - \nu e_3$:

$$x(t) = e^{-t \operatorname{ad}(a-b)} x(\tau_0), \qquad t \in [\tau_0, \tau_1].$$

The switching points $x(\tau_0)$, $x(\tau_1)$ satisfy the equalities:

$$\begin{aligned} &\langle x(\tau_0), e_3 \rangle = \langle x(\tau_1), e_3 \rangle = 0, \\ &\langle x(\tau_0), e_2 \rangle = \langle x(\tau_1), e_2 \rangle = h_{u(\tau_0 - \varepsilon)}(x(\tau_0 - \varepsilon)), \\ &|x(\tau_0) = |x(\tau_1)|. \end{aligned}$$

Consequently,

$$\langle x(\tau_0), e_1 \rangle = -\langle x(\tau_1), e_1 \rangle,$$

i.e., $x(\tau_1)$ is the reflection of $x(\tau_0)$ w.r.t. the plane $\operatorname{span}(e_2, e_3)$. Geometrically it is easy to see that the angle of rotation θ from $x(\tau_0)$ to $x(\tau_1)$ around $a - b$ is bounded as follows:

$$\theta \in [\pi, 2\pi].$$

The extremal values of θ are attained when the point $x(\tau_0)$ is on the boundary of $\operatorname{cone}(e_1, e_2)$:

$$\begin{aligned} x(\tau_0) \in \mathbb{R}_+ e_1 \quad &\Rightarrow \quad \theta = \pi, \\ x(\tau_0) \in \mathbb{R}_+ e_2 \quad &\Rightarrow \quad \theta = 2\pi. \end{aligned}$$

In the second case the point $x(t)$, as well as the point $q(t)$, makes a complete revolution at the angle 2π. Such an arc cannot be a part of an optimal trajectory: it can be eliminated with decrease of the terminal time t_1. Consequently, the angle between two switchings is

$$\theta \in [\pi, 2\pi).$$

Let $x(\tau_2)$ be the next switching after $x(\tau_1)$. The behavior of control after the switching $x(\tau_1)$ from -1 to $+1$ is similar to the behavior after $x(\tau_0)$. Indeed, our time-optimal problem admits the symmetry

$$b \mapsto -b.$$

After the change of basis

$$e_3 \mapsto -e_3, \qquad e_1 \mapsto -e_1, \qquad e_2 \mapsto e_2$$

the curve $x(t)$ is preserved, but now it switches at $x(\tau_1)$ from $+1$ to -1. This case was already studied, thus the angle of rotation from $x(\tau_1)$ to $x(\tau_2)$ is again

θ, moreover, $x(\tau_2) = x(\tau_0)$. The next switching point is $x(\tau_3) = x(\tau_1)$, and so on.

Thus the structure of bang-bang optimal trajectories is quite simple. Such trajectories contain a certain number of switching points. Between these switching points the vector $x(t)$ rotates alternately around the vectors $a+b$ and $a-b$ at an angle $\theta \in [\pi, 2\pi)$ constant along each bang-bang trajectory. Before the first switching and after the last switching the vector $x(t)$ can rotate at angles θ_0 and θ_1 respectively, $0 < \theta_0,\ \theta_1 \leq \theta$. The system of all optimal bang-bang trajectories is parametrized by 3 continuous parameters θ_0, θ, θ_1, and 2 discrete parameters: the number of switchings and the initial control $\operatorname{sgn}\langle x(0), b\rangle$.

An optimal trajectory can be not bang-bang only if the point $x(\tau_0)$ corresponding to the first nonnegative root of the equation $\langle x(t), b\rangle = 0$ satisfies the equalities

$$\langle x(\tau_0), b\rangle = \langle x(\tau_0), [a,b]\rangle = 0.$$

Then

$$x(\tau_0) = \mu e_2, \qquad \mu \neq 0.$$

There can be two possibilities:

(1) either the switching function $\langle x(t), b\rangle$ takes nonzero values for some $t > \tau_0$ and arbitrarily close to τ_0,

(2) or

$$\langle x(t), b\rangle \equiv 0, \qquad t \in [\tau_0, \tau_0 + \varepsilon], \tag{19.50}$$

for some $\varepsilon > 0$.

We start from the first alternative. From the analysis of bang-bang trajectories it follows that switching times cannot accumulate to τ_0 from the right: the angle of rotation between two consecutive switchings $\theta \geq \pi$. Thus in case (1) we have

$$\langle x(t), b\rangle > 0, \qquad t \in [\tau_0, \tau_0 + \delta],$$

for some $\delta > 0$. That is, τ_0 is a switching time. Since $x(\tau_0) \in \mathbb{R}e_1$, then the angle of rotation until the next switching point is $\theta = 2\pi$, which is not optimal. So case (1) cannot occur for an optimal trajectory.

Consider case (2). We differentiate identity (19.50) twice w.r.t. t:

$$\begin{aligned}
\frac{d}{dt}\langle x(t), b\rangle &= -\langle x(t), [a,b]\rangle \equiv 0,\\
\frac{d}{dt}\langle x(t), [a,b]\rangle &= \langle [x(t), a + u(t)b, [a,b]\rangle = u(t)\langle [x(t), b], [a,b]\rangle\\
&= 0.
\end{aligned}$$

Then $x(t) = \mu(t)e_2$, $t \in [\tau_0, \tau_0 + \varepsilon]$, thus

$$u(t)\langle [a,b], [a,b]\rangle = 0,$$

i.e.,

$$u(t) \equiv 0, \qquad t \in [\tau_0, \tau_0 + \varepsilon].$$

This control is not determined directly from PMP (we found it with the help of differentiation). Such a control is called singular.

Optimal trajectories containing a singular part (corresponding to the control $u(t) \equiv 0$) can have an arc with $u \equiv \pm 1$ before the singular part, with the angle of rotation around $a \pm b$ less then 2π; such an arc can also be after the singular one. So there can be 4 types of optimal trajectories containing a singular arc:

$$+\,0\,+, \qquad +\,0\,-, \qquad -\,0\,+, \qquad -\,0\,-.$$

The family of such trajectories is parametrized by 3 continuous parameters (angles of rotation at the corresponding arcs) and by 2 discrete parameters (signs at the initial and final segments).

So we described the structure of all possible optimal trajectories: the bang-bang one, and the strategy with a singular part. The domains of points in SO(3) attained via these strategies are 3-dimensional, and the union of these domains covers the whole group SO(3). But it is easy to see that a sufficiently long trajectory following any of the two strategies is not optimal: the two domains in SO(3) overlap. Moreover, each of the strategies overlaps with itself. In order to know optimal trajectory for any point in SO(3), one should study the interaction of the two strategies and intersections of trajectories that follow the same strategy. This interesting problem remains open.

Notice that the structure of optimal trajectories in this left-invariant time-optimal problem on SO(3) is similar to the structure of optimal trajectories for Dubins car (Sec. 13.5). This resemblance is not accidental: the problem on Dubins car can be formulated as a left-invariant time-optimal problem on the group of isometries of the plane.

Chapter 20

Second order optimality conditions

20.1 Hessian

In this chapter we obtain second order necessary optimality conditions for control problems. As we know, geometrically the study of optimality reduces to the study of boundary of attainable sets (see Section 10.2). Consider a control system

$$\dot{q} = f_u(q), \qquad q \in M, \quad u \in U = \operatorname{int} U \subset \mathbb{R}^m, \tag{20.1}$$

where the state space M is, as usual, a smooth manifold, and the space of control parameters U is open (essentially, this means that we study optimal controls that do not come to the boundary of U, although a similar theory for bang-bang controls can also be constructed). The attainable set $\mathcal{A}_{q_0}(t_1)$ of system (20.1) is the image of the *endpoint mapping*

$$F_{t_1} \ : \ u(\cdot) \mapsto q_0 \circ \overrightarrow{\exp} \int_0^{t_1} f_{u(t)} \, dt.$$

We say that a trajectory $q(t)$, $t \in [0, t_1]$, is *geometrically optimal* for system (20.1) if it comes to the boundary of the attainable set for the terminal time t_1:

$$q(t_1) \in \partial \mathcal{A}_{q_0}(t_1).$$

Necessary conditions for this inclusion are given by Pontryagin Maximum Principle. A part of the statements of PMP can be viewed as the first order optimality condition (we see this later). Now we seek for optimality conditions of the second order.

Consider the problem in a general setting. Let

$$F \, : \mathcal{U} \to M$$

be a smooth mapping, where $\mathcal{U}$ is an open subset in a Banach space and M is a smooth n-dimensional manifold (usually in the sequel $\mathcal{U}$ is the space of admissible controls $L_\infty([0,t_1],U)$ and $F = F_{t_1}$ is the endpoint mapping of a control system). The first differential

$$D_u F : T_u \mathcal{U} \to T_{F(u)} M$$

is well defined independently on coordinates. This is not the case for the second differential. Indeed, consider the case where u is a regular point for F, i.e., the differential $D_u F$ is surjective. By implicit function theorem, the mapping F becomes linear in suitably chosen local coordinates in $\mathcal{U}$ and M, thus it has no intrinsic second differential. In the general case, well defined independently of coordinates is only a certain part of the second differential.

The differential of a smooth mapping $F : \mathcal{U} \to M$ can be defined via the first order derivative

$$D_u F\, v = \left.\frac{d}{d\,\varepsilon}\right|_{\varepsilon=0} F(\varphi(\varepsilon)) \tag{20.2}$$

along a curve $\varphi : (-\varepsilon_0, \varepsilon_0) \to \mathcal{U}$ with the initial conditions

$$\varphi(0) = u \in \mathcal{U}, \qquad \dot\varphi(0) = v \in T_u \mathcal{U}.$$

In local coordinates, this derivative is computed as

$$\frac{d\,F}{d\,u}\dot\varphi, \qquad \dot\varphi = \dot\varphi(0).$$

In other coordinates $\tilde q$ in M, derivative (20.2) is evaluated as

$$\frac{d\,\widetilde F}{d\,u}\dot\varphi = \frac{d\,\tilde q}{d\,q}\frac{d\,F}{d\,u}\dot\varphi.$$

Coordinate representation of the first order derivative (20.2) transforms under changes of coordinates as a tangent vector to M — it is multiplied by the Jacobian matrix $\dfrac{d\,\tilde q}{d\,q}$.

The second order derivative

$$\left.\frac{d^2}{d\,\varepsilon^2}\right|_{\varepsilon=0} F(\varphi(\varepsilon)), \tag{20.3}$$
$$\varphi(0) = u \in \mathcal{U}, \qquad \dot\varphi(0) = v \in T_u \mathcal{U},$$

is evaluated in coordinates as

$$\frac{d^2\,F}{d\,u^2}(\dot\varphi,\dot\varphi) + \frac{d\,F}{d\,u}\ddot\varphi.$$

Transformation rule for the second order directional derivative under changes of coordinates has the form:

$$\frac{d^2\widetilde{F}}{du^2}(\dot\varphi,\dot\varphi)+\frac{d\widetilde{F}}{du}\ddot\varphi=\frac{d\tilde q}{dq}\left[\frac{d^2F}{du^2}(\dot\varphi,\dot\varphi)+\frac{dF}{du}\ddot\varphi\right]$$
$$+\frac{d^2\tilde q}{dq^2}\left(\frac{dF}{du}\dot\varphi,\frac{dF}{du}\dot\varphi\right). \tag{20.4}$$

The second order derivative (20.3) transforms as a tangent vector in $T_{F(u)}M$ only if $\dot\varphi=v\in\operatorname{Ker}D_uF$, i.e., if term (20.4) vanishes. Moreover, it is determined by u and v only modulo the subspace $\operatorname{Im}D_uF$, which is spanned by the term $\dfrac{dF}{du}\ddot\varphi$.

Thus intrinsically defined is the quadratic mapping

$$\operatorname{Ker}D_uF\to T_{F(u)}M/\operatorname{Im}D_uF,$$
$$v\mapsto \left.\frac{d^2}{d\varepsilon^2}\right|_{\varepsilon=0}F(\varphi(\varepsilon)) \mod \operatorname{Im}D_uF. \tag{20.5}$$

After this preliminary discussion, we turn to formal definitions.

The *Hessian* of a smooth mapping $F:\mathcal{U}\to M$ at a point $u\in\mathcal{U}$ is a symmetric bilinear mapping

$$\operatorname{Hess}_u F:\operatorname{Ker}D_uF\times\operatorname{Ker}D_uF\to\operatorname{Coker}D_uF=T_{F(u)}M/\operatorname{Im}D_uF. \tag{20.6}$$

In particular, at a regular point $\operatorname{Coker}D_uF=0$, thus $\operatorname{Hess}_u F=0$. Hessian is defined as follows. Let

$$v,\ w\in\operatorname{Ker}D_uF$$

and

$$\lambda\in(\operatorname{Im}D_uF)^\perp\subset T^*_{F(u)}M.$$

In order to define the value

$$\lambda\operatorname{Hess}_u F(v,w),$$

take vector fields

$$V,\ W\in\operatorname{Vec}\mathcal{U},\qquad V(u)=v,\quad W(u)=w,$$

and a function

$$a\in C^\infty(M),\qquad d_{F(u)}a=\lambda.$$

Then

$$\lambda\operatorname{Hess}_u F(v,w)\stackrel{\text{def}}{=}\left.V\circ W\,(a\circ F)\right|_u. \tag{20.7}$$

We show now that the right-hand side does not depend upon the choice of V, W, and a. The first Lie derivative is

$$W(a\circ F)=\langle d_{F(\cdot)}a,\ F_*W(\cdot)\rangle,$$

and the second Lie derivative $V \circ W\,(a \circ F)|_u$ does not depend on second derivatives of a since $F_* W(u) = 0$. Moreover, the second Lie derivative obviously depends only on the value of V at u but not on derivatives of V at u. In order to show the same for the field W, we prove that the right-hand side of the definition of Hessian is symmetric w.r.t. V and W:

$$(W \circ V(a \circ F) - V \circ W(a \circ F))|_u = [W, V]\,(a \circ F)|_u = \underbrace{d_{F(u)}a}_{=\lambda} \circ D_u F\,[W, V](u)$$

$$= 0$$

since $\lambda \perp \operatorname{Im} D_u F$. We showed that the mapping $\operatorname{Hess}_u F$ given by (20.7) is intrinsically defined independently of coordinates as in (20.6).

Exercise 20.1. Show that the quadratic mapping (20.5) defined via the second order directional derivative coincides with $\operatorname{Hess}_u F(v, v)$.

If we admit only linear changes of variables in $\mathcal{U}$, then we can correctly define the full *second differential*

$$D_u^2 F \,:\, \operatorname{Ker} D_u F \times \operatorname{Ker} D_u F \to T_{F(u)} M$$

in the same way as Hessian (20.7), but the covector is arbitrary:

$$\lambda \in T^*_{F(u)} M,$$

and the vector fields are constant:

$$V \equiv v, \qquad W \equiv w.$$

The Hessian is the part of the second differential independent on the choice of linear structure in the preimage.

Exercise 20.2. Compute the Hessian of the restriction $F|_{f^{-1}(0)}$ of a smooth mapping F to a level set of a smooth function f. Consider the restriction of a smooth mapping $F : \mathcal{U} \to M$ to a smooth hypersurface $S = f^{-1}(0)$, $f : \mathcal{U} \to \mathbb{R}$, $df \neq 0$, and let $u \in S$ be a regular point of F. Prove that the Hessian of the restriction is computed as follows:

$$\lambda \operatorname{Hess}_u (F|_S) = \lambda D_u^2 F - d_u^2 f, \qquad \lambda \perp \operatorname{Im} D_u\ F|_S\,, \quad \lambda \in T^*_{F(u)} M \setminus \{0\},$$

and the covector λ is normalized so that

$$\lambda D_u F = d_u f.$$

20.2 Local openness of mappings

A mapping $F : \mathcal{U} \to M$ is called *locally open* at a point $u \in \mathcal{U}$ if

$$F(u) \in \operatorname{int} F(O_u)$$

for any neighborhood $O_u \subset \mathcal{U}$ of u. In the opposite case, i.e., when

$$F(u) \in \partial F(O_u)$$

for some neighborhood O_u, the point u is called *locally geometrically optimal* for F.

A point $u \in \mathcal{U}$ is called *locally finite-dimensionally optimal* for a mapping F if for any finite-dimensional smooth submanifold $S \subset \mathcal{U}$, $u \in S$, the point u is locally geometrically optimal for the restriction $F|_S$.

20.2.1 Critical points of corank one

Corank of a critical point u of a smooth mapping F is by definition equal to corank of the differential $D_u F$:

$$\operatorname{corank} D_u F = \operatorname{codim} \operatorname{Im} D_u F.$$

In the sequel we will often consider critical points of corank one. In this case the Lagrange multiplier

$$\lambda \in (\operatorname{Im} D_u F)^\perp, \qquad \lambda \neq 0,$$

is defined uniquely up to a nonzero factor, and

$$\lambda \operatorname{Hess}_u F \;:\; \operatorname{Ker} D_u F \times \operatorname{Ker} D_u F \to \mathbb{R}$$

is just a quadratic form (in the case $\operatorname{corank} D_u F > 1$, we should consider a family of quadratic forms).

Now we give conditions of local openness of a mapping F at a corank one critical point u in terms of the quadratic form $\lambda \operatorname{Hess}_u F$.

Theorem 20.1. *Let $F : \mathcal{U} \to M$ be a continuous mapping having smooth restrictions to finite-dimensional submanifolds of $\mathcal{U}$. Let $u \in \mathcal{U}$ be a corank one critical point of F, and let $\lambda \in (\operatorname{Im} D_u F)^\perp$, $\lambda \neq 0$.*

(1) *If the quadratic form $\lambda \operatorname{Hess}_u F$ is sign-indefinite, then F is locally open at u.*

(2) *If the form $\lambda \operatorname{Hess}_u F$ is negative (or positive), then u is locally finite-dimensionally optimal for F.*

Remark. A quadratic form is locally open at the origin iff it is sign-indefinite.

Proof. The statements of the theorem are local, so we fix local coordinates in $\mathcal{U}$ and M centered at u and $F(u)$ respectively, and assume that $\mathcal{U}$ is a Banach space and $M = \mathbb{R}^n$.

(1) Consider the splitting into direct sum in the preimage:

$$T_u \mathcal{U} = E \oplus \operatorname{Ker} D_u F, \qquad \dim E = n - 1, \tag{20.8}$$

and the corresponding splitting in the image:

$$T_{F(u)}M = \operatorname{Im} D_u F \oplus V, \qquad \dim V = 1. \tag{20.9}$$

The quadratic form $\lambda \operatorname{Hess}_u F$ is sign-indefinite, i.e., it takes values of both signs on $\operatorname{Ker} D_u F$. Thus we can choose vectors

$$v, \; w \in \operatorname{Ker} D_u F$$

such that

$$\lambda F''_u(v,v) = 0, \qquad \lambda F''_u(v,w) \neq 0,$$

we denote by F', F'' derivatives of the vector function F in local coordinates. Since the first differential is an isomorphism:

$$D_u F = F'_u \, : \, E \to \operatorname{Im} D_u F = \lambda^{\perp},$$

there exists a vector $x_0 \in E$ such that

$$F'_u x_0 = -\frac{1}{2} F''_u(v,v).$$

Introduce the following family of mappings:

$$\begin{aligned}
&\Phi_\varepsilon \, : \, E \times \mathbb{R} \to M, \qquad \varepsilon \in \mathbb{R}, \\
&\Phi_\varepsilon(x,y) = F(\varepsilon^2 v + \varepsilon^3 y w + \varepsilon^4 x_0 + \varepsilon^5 x), \qquad x \in E, \; y \in \mathbb{R},
\end{aligned}$$

notice that

$$\operatorname{Im} \Phi_\varepsilon \subset \operatorname{Im} F$$

for small ε. Thus it is sufficient to show that Φ_ε is open. The Taylor expansion

$$\Phi_\varepsilon(x,y) = \varepsilon^5 (F'_u x + y F''_u(v,w)) + O(\varepsilon^6), \qquad \varepsilon \to 0,$$

implies that the family $\frac{1}{\varepsilon^5}\Phi_\varepsilon$ is smooth w.r.t. parameter ε at $\varepsilon = 0$. For $\varepsilon = 0$ this family gives a surjective linear mapping. By implicit function theorem, the mappings $\frac{1}{\varepsilon^5}\Phi_\varepsilon$ are submersions, thus are locally open for small $\varepsilon > 0$. Thus the mapping F is also locally open at u.

(2) Take any smooth finite-dimensional submanifold $S \subset \mathcal{U}$, $u \in S$. Similarly to (20.8), (20.9), consider the splittings in the preimage:

$$S \cong T_u S = L \oplus \operatorname{Ker} D_u \left. F \right|_S,$$

and in the image:

$$\begin{aligned}
&M \cong T_{F(u)} M = \operatorname{Im} D_u \left. F \right|_S \oplus W, \\
&\dim W = k = \operatorname{corank} D_u \left. F \right|_S \geq 1.
\end{aligned}$$

Since the differential $D_u F : E \to \operatorname{Im} D_u F$ is an isomorphism, we can choose, by implicit function theorem, coordinates (x, y) in S and coordinates in M such that the mapping F takes the form

$$F(x, y) = \begin{pmatrix} x \\ \varphi(x, y) \end{pmatrix}, \qquad x \in L, \quad y \in \operatorname{Ker} D_u \left. F\right|_S .$$

Further, we can choose coordinates $\varphi = (\varphi_1, \dots, \varphi_k)$ in W such that

$$\lambda F(x, y) = \varphi_1(x, y).$$

Now we write down hypotheses of the theorem in these coordinates. Since $\operatorname{Im} D_u \left. F\right|_S \cap W = \{0\}$, then

$$D_{(0,0)}\varphi_1 = 0.$$

Further, the hypothesis that the form $\lambda \operatorname{Hess}_u F$ is negative reads

$$\left. \frac{\partial^2 \varphi_1}{\partial y^2} \right|_{(0,0)} < 0.$$

Then the function

$$\varphi_1(0, y) < 0 \quad \text{for small } y.$$

Thus the mapping $\left. F\right|_S$ is not locally open at u. □

There holds the following statement, which is much stronger than the previous one.

Theorem 20.2 (Generalized Morse's lemma). *Suppose that $u \in \mathcal{U}$ is a corank one critical point of a smooth mapping $F : \mathcal{U} \to M$ such that* $\operatorname{Hess}_u F$ *is a nondegenerate quadratic form. Then there exist local coordinates in $\mathcal{U}$ and M in which F has only terms of the first and second orders:*

$$F(x, v) = D_u F\, x + \frac{1}{2} \operatorname{Hess}_u F(v, v),$$
$$(x, v) \in \mathcal{U} \cong E \oplus \operatorname{Ker} D_u F.$$

We do not prove this theorem since it will not be used in the sequel.

20.2.2 Critical points of arbitrary corank

The necessary condition of local openness given by item (1) of Theorem 20.1 can be generalized for critical points of arbitrary corank.

Recall that *positive (negative) index* of a quadratic form Q is the maximal dimension of a positive (negative) subspace of Q:

$$\operatorname{ind}_+ Q = \max\left\{ \dim L \mid \left. Q\right|_{L\setminus\{0\}} > 0 \right\},$$
$$\operatorname{ind}_- Q = \max\left\{ \dim L \mid \left. Q\right|_{L\setminus\{0\}} < 0 \right\}.$$

Theorem 20.3. *Let $F : \mathcal{U} \to M$ be a continuous mapping having smooth restrictions to finite-dimensional submanifolds. Let $u \in \mathcal{U}$ be a critical point of F of corank m. If*

$$\operatorname{ind}_- \lambda \operatorname{Hess}_u F \geq m \qquad \forall\, \lambda \perp \operatorname{Im} D_u F,\ \lambda \neq 0,$$

then the mapping F is locally open at the point u.

Proof. First of all, the statement is local, so we can choose local coordinates and assume that $\mathcal{U}$ is a Banach space and $u = 0$, and $M = \mathbb{R}^n$ with $F(0) = 0$.

Moreover, we can assume that the space $\mathcal{U}$ is finite-dimensional, now we prove this. For any $\lambda \perp \operatorname{Im} D_u F$, $\lambda \neq 0$, there exists a subspace

$$E_\lambda \subset \mathcal{U}, \qquad \dim E_\lambda = m,$$

such that

$$\lambda \operatorname{Hess}_u F|_{E_\lambda \setminus \{0\}} < 0.$$

We take λ from the unit sphere

$$S^{m-1} = \left\{ \lambda \in (\operatorname{Im} D_u F)^\perp \mid |\lambda| = 1 \right\}.$$

For any $\lambda \in S^{m-1}$, there exists a neighborhood $O_\lambda \subset S^{m-1}$, $\lambda \in O_\lambda$, such that $E_{\lambda'} = E_\lambda$ for any $\lambda' \in O_\lambda$, this easily follows from continuity of the form $\lambda' \operatorname{Hess}_u F$ on the unit sphere in E_λ. Choose a finite covering:

$$S^{m-1} = \bigcup_{i=1}^{N} O_{\lambda_i}.$$

Then restriction of F to the finite-dimensional subspace $\sum_{i=1}^N E_{\lambda_i}$ satisfies the hypothesis of the theorem. Thus we can assume that $\mathcal{U}$ is finite-dimensional. Then the theorem is a consequence of the following Lemmas 20.1 and 20.2. □

Lemma 20.1. *Let $F : \mathbb{R}^N \to \mathbb{R}^n$ be a smooth mapping, and let $F(0) = 0$. Assume that the quadratic mapping*

$$Q = \operatorname{Hess}_0 F : \operatorname{Ker} D_0 F \to \operatorname{Coker} D_0 F$$

has a regular zero:

$$\exists\ v \in \operatorname{Ker} D_0 F \text{ s.t. } Q(v) = 0, \quad D_v Q \text{ surjective.}$$

Then the mapping F has regular zeros arbitrarily close to the origin in $\mathbb{R}^N$.

Proof. We modify slightly the argument used in the proof of item (1) of Theorem 20.1. Decompose preimage of the first differential:

$$\mathbb{R}^N = E \oplus \operatorname{Ker} D_0 F, \qquad \dim E = n - m,$$

then the restriction

$$D_0F \,:\, E \to \operatorname{Im} D_0F$$

is one-to-one. The equality $Q(v) = \operatorname{Hess}_0 F(v) = 0$ means that

$$F_0''(v,v) \in \operatorname{Im} D_0F.$$

Then there exists $x_0 \in E$ such that

$$F_0'x_0 = -\frac{1}{2}F_0''(v,v).$$

Define the family of mappings

$$\Phi_\varepsilon(x,y) = F(\varepsilon^2 v + \varepsilon^3 y + \varepsilon^4 x_0 + \varepsilon^5 x), \qquad x \in E, \quad y \in \operatorname{Ker} D_0F.$$

The first four derivatives of Φ_ε vanish at $\varepsilon = 0$, and we obtain the Taylor expansion

$$\frac{1}{\varepsilon^5}\Phi_\varepsilon(x,y) = F_0'x + F_0''(v,y) + O(\varepsilon), \quad \varepsilon \to 0.$$

Then we argue as in the proof of Theorem 20.1. The family $\frac{1}{\varepsilon^5}\Phi_\varepsilon$ is smooth and linear surjective at $\varepsilon = 0$. By implicit function theorem, the mappings $\frac{1}{\varepsilon^5}\Phi_\varepsilon$ are submersions for small $\varepsilon > 0$, thus they have regular zeros in any neighborhood of the origin in $\mathbb{R}^N$. Consequently, the mapping F also has regular zeros arbitrarily close to the origin in $\mathbb{R}^N$. □

Lemma 20.2. *Let $Q : \mathbb{R}^N \to \mathbb{R}^m$ be a quadratic mapping such that*

$$\operatorname{ind}_- \lambda Q \geq m \qquad \forall\ \lambda \in \mathbb{R}^{m*},\ \lambda \neq 0.$$

Then the mapping Q has a regular zero.

Proof. We can assume that the quadratic form Q has no kernel:

$$Q(v,\cdot) \neq 0 \qquad \forall\ v \neq 0. \tag{20.10}$$

If this is not the case, we factorize by kernel of Q. Since $D_vQ = 2Q(v,\cdot)$, condition (20.10) means that $D_vQ \neq 0$ for $v \neq 0$.

Now we prove the lemma by induction on m.

In the case $m = 1$ the statement is obvious: a sign-indefinite quadratic form has a regular zero.

Induction step: we prove the statement of the lemma for any $m > 1$ under the assumption that it is proved for all values less than m.

(1) Suppose first that $Q^{-1}(0) \neq 0$. Take any $v \neq 0$ such that $Q(v) = 0$. If v is a regular point of Q, then the statement of this lemma follows. Thus we assume that v is a critical point of Q. Since $D_vQ \neq 0$, then

$$\operatorname{rank} D_vQ = k, \qquad 0 < k < m.$$

Consider Hessian of the mapping Q:

$$\operatorname{Hess}_v Q : \operatorname{Ker} D_v Q \to \mathbb{R}^{m-k}.$$

The second differential of a quadratic mapping is the doubled mapping itself, thus

$$\lambda \operatorname{Hess}_v Q = 2\, \lambda Q|_{\operatorname{Ker} D_v Q}.$$

Further, since $\operatorname{ind}_- \lambda Q \geq m$ and $\operatorname{codim} \operatorname{Ker} D_v Q = k$, then

$$\operatorname{ind}_- \lambda \operatorname{Hess}_v Q = \operatorname{ind}_- \lambda Q|_{\operatorname{Ker} D_v Q} \geq m - k.$$

By the induction assumption, the quadratic mapping $\operatorname{Hess}_v Q$ has a regular zero. Then Lemma 20.1 applied to the mapping Q yields that Q has a regular zero as well. The statement of this lemma in case (1) follows.

(2) Consider now the second case: $Q^{-1}(0) = 0$.

(2.a) It is obvious that $\operatorname{Im} Q$ is a closed cone.

(2.b) Moreover, we can assume that $\operatorname{Im} Q \setminus \{0\}$ is open. Indeed, suppose that there exists

$$x = Q(v) \in \partial \operatorname{Im} Q, \qquad x \neq 0.$$

Then v is a critical point of Q, and in the same way as in case (1) the induction assumption for $\operatorname{Hess}_v Q$ yields that $\operatorname{Hess}_v Q$ has a regular zero. By Lemma 20.1, Q is locally open at v and $Q(v) \in \operatorname{int} \operatorname{Im} Q$. Thus we assume in the sequel that $\operatorname{Im} Q \setminus \{0\}$ is open. Combined with item (a), this means that Q is surjective.

(2.c) We show now that this property leads to a contradiction which proves the lemma.

The smooth mapping

$$\frac{Q}{|Q|} : S^{N-1} \to S^{m-1}, \qquad v \mapsto \frac{Q(v)}{|Q(v)|}, \quad v \in S^{N-1},$$

is surjective. By Sard's theorem, it has a regular value. Let $x \in S^{m-1}$ be a regular value of the mapping $Q/|Q|$.

Now we proceed as follows. We find the minimal $a > 0$ such that

$$Q(v) = ax, \qquad v \in S^{N-1},$$

and apply optimality conditions at the solution v_0 to show that $\operatorname{ind}_- \lambda Q \leq m-1$, a contradiction.

So consider the following finite-dimensional optimization problem with constraints:

$$a \to \min, \quad Q(v) = ax, \quad a > 0, \quad v \in S^{N-1}. \tag{20.11}$$

This problem obviously has a solution, let a pair (v_0, a_0) realize minimum. We write down first- and second-order optimality conditions for problem (20.11). There exist Lagrange multipliers

$$(\nu, \lambda) \neq 0, \qquad \nu \in \mathbb{R}, \quad \lambda \in T^*_{a_0 x}\mathbb{R}^m,$$

such that the Lagrange function

$$\mathcal{L}(\nu, \lambda, a, v) = \nu a + \lambda(Q(v) - ax)$$

satisfies the stationarity conditions:

$$\frac{\partial \mathcal{L}}{\partial a} = \nu - \lambda x = 0, \tag{20.12}$$
$$\left.\frac{\partial \mathcal{L}}{\partial v}\right|_{(v_0,a_0)} = \lambda D_{v_0} Q|_{S^{N-1}} = 0.$$

Since v_0 is a regular point of the mapping $Q/|Q|$, then $\nu \neq 0$, thus we can set

$$\nu = 1.$$

Then second-order necessary optimality condition for problem (20.11) reads

$$\lambda \operatorname{Hess}_{v_0} Q|_{S^{N-1}} \geq 0. \tag{20.13}$$

Recall that Hessian of restriction of a mapping is not equal to restriction of Hessian of this mapping, see Exercise 20.2 above.

Exercise 20.3. Prove that

$$\lambda \left(\operatorname{Hess}_v Q|_{S^{N-1}}\right)(u) = 2(\lambda Q(u) - |u|^2 \lambda Q(v)),$$
$$v \in S^{N-1}, \quad u \in \operatorname{Ker} D_v\, Q|_{S^{N-1}}.$$

That is, inequality (20.13) yields

$$\lambda Q(u) - |u|^2 \lambda Q(v_0) \geq 0, \qquad u \in \operatorname{Ker} D_{v_0}\, Q|_{S^{N-1}},$$

thus

$$\lambda Q(u) \geq |u|^2 \lambda Q(v_0) = |u|^2 a_0 \lambda x = |u|^2 a_0 \nu = |u|^2 a_0 > 0,$$

i.e.,

$$\lambda Q(u) \geq 0, \qquad u \in \operatorname{Ker} D_{v_0}\, Q|_{S^{N-1}}.$$

Moreover, since $v_0 \notin T_{v_0} S^{N-1}$, then

$$\lambda Q|_L \geq 0, \qquad L = \operatorname{Ker} D_{v_0}\, Q|_{S^{N-1}} \oplus \mathbb{R} v_0.$$

Now we compute dimension of the nonnegative subspace L of the quadratic form λQ. Since v_0 is a regular value of $\frac{Q}{|Q|}$, then

$$\dim \operatorname{Im} D_{v_0} \frac{Q}{|Q|} = m - 1.$$

Thus $\operatorname{Im} D_{v_0}\, Q|_{S^{N-1}}$ can have dimension m or $m-1$. But v_0 is a critical point of $Q|_{S^{N-1}}$, thus

$$\dim \operatorname{Im} D_{v_0}\, Q|_{S^{N-1}} = m - 1$$

and

$$\dim \operatorname{Ker} D_{v_0} Q|_{S^{N-1}} = N - 1 - (m-1) = N - m.$$

Consequently, $\dim L = N - m + 1$, thus $\operatorname{ind}_- \lambda Q \leq m - 1$, which contradicts the hypothesis of this lemma.

So case (c) is impossible, and the induction step in this lemma is proved. □

Theorem 20.3 is completely proved.

20.3 Differentiation of the endpoint mapping

In this section we compute differential and Hessian of the endpoint mapping for a control system

$$\begin{aligned}
&\dot q = f_u(q), \qquad u \in U \subset \mathbb{R}^m, \quad U = \operatorname{int} U, \quad q \in M, && (20.14)\\
&q(0) = q_0,\\
&u(\cdot) \in \mathcal{U} = L_\infty([0, t_1], U),
\end{aligned}$$

with the right-hand side $f_u(q)$ smooth in (u, q). We study the endpoint mapping

$$\begin{aligned}
&F_{t_1} : \mathcal{U} \to M,\\
&F_{t_1} \ : \ u(\cdot) \mapsto q_0 \circ \overrightarrow{\exp} \int_0^{t_1} f_{u(t)}\, dt
\end{aligned}$$

in the neighborhood of a fixed admissible control

$$\tilde u = \tilde u(\cdot) \in \mathcal{U}.$$

In the same way as in the proof of PMP (see Section 12.2), the Variations formula yields a decomposition of the flow:

$$F_{t_1}(u) = q_0 \circ \overrightarrow{\exp} \int_0^{t_1} g_{t,u(t)}\, dt \circ P_{t_1},$$

where

$$\begin{aligned}
P_t &= \overrightarrow{\exp} \int_0^t f_{\tilde u(\tau)}\, d\tau,\\
g_{t,u} &= P_{t*}^{-1}(f_u - f_{\tilde u(t)}).
\end{aligned}$$

Further, introduce an intermediate mapping

$$\begin{aligned}
&G_{t_1} : \mathcal{U} \to M,\\
&G_{t_1} \ : \ u \mapsto q_0 \circ \overrightarrow{\exp} \int_0^{t_1} g_{t,u(t)}\, dt.
\end{aligned}$$

Then

$$F_{t_1}(u) = P_{t_1}(G_{t_1}(u)),$$

consequently,

$$\begin{aligned} D_{\tilde u} F_{t_1} &= P_{t_1 *} D_{\tilde u} G_{t_1}, \\ \operatorname{Hess}_{\tilde u} F_{t_1} &= P_{t_1 *} \operatorname{Hess}_{\tilde u} G_{t_1}, \end{aligned}$$

so differentiation of F_{t_1} reduces to differentiation of G_{t_1}. We compute derivatives of the mapping G_{t_1} using the asymptotic expansion of the chronological exponential:

$$\begin{aligned} &a(G_{t_1}(u)) \\ &\quad = q_0 \circ \left(\mathrm{Id} + \int_0^{t_1} g_{\tau, u(\tau)}\, d\tau + \iint\limits_{0 \le \tau_2 \le \tau_1 \le t_1} g_{\tau_2, u(\tau_2)} \circ g_{\tau_1, u(\tau_1)}\, d\tau_1\, d\tau_2 \right) a \\ &\qquad\qquad + O\left(\|u - \tilde u\|^3_{L_\infty} \right). \qquad (20.15) \end{aligned}$$

Introduce some more notations:

$$\begin{aligned} g'_\tau &= \left. \frac{\partial}{\partial u} \right|_{\tilde u(\tau)} g_{\tau, u}, \\ g''_\tau &= \left. \frac{\partial^2}{\partial u^2} \right|_{\tilde u(\tau)} g_{\tau, u}, \\ h_u(\lambda) &= \langle \lambda, f_u(q) \rangle, \qquad \lambda \in T^*_q M, \\ h'_\tau &= \left. \frac{\partial}{\partial u} \right|_{\tilde u(\tau)} h_u, \\ h''_\tau &= \left. \frac{\partial^2}{\partial u^2} \right|_{\tilde u(\tau)} h_u. \end{aligned}$$

Then differential (the first variation) of the mapping G_{t_1} has the form:

$$(D_{\tilde u} G_{t_1}) v = q_0 \circ \int_0^{t_1} g'_t v(t)\, dt, \qquad v = v(\cdot) \in T_{\tilde u} \mathcal{U}.$$

The control $\tilde u$ is a critical point of F_{t_1} (or, which is equivalent, of G_{t_1}) if and only if there exists a Lagrange multiplier

$$\lambda_0 \in T^*_{q_0} M, \qquad \lambda_0 \neq 0,$$

such that

$$\lambda_0 (D_{\tilde u} G_{t_1}) v = 0 \qquad \forall\ v \in T_{\tilde u} \mathcal{U},$$

i.e.,

$$\lambda_0 g'_t(q_0) = 0, \qquad t \in [0, t_1].$$

Translate the covector λ_0 along the reference trajectory

$$q(t) = q_0 \circ P_t,$$

we obtain the covector curve

$$\lambda_t = P_t^{*-1} \lambda_0 = \lambda_0 P_{t*}^{-1} \in T^*_{q(t)} M,$$

which is a trajectory of the Hamiltonian system

$$\dot{\lambda}_t = \vec{h}_{\tilde{u}(t)}(\lambda_t), \qquad t \in [0, t_1],$$

see Proposition 11.3. Then

$$\lambda_0 g'_t(q_0) = \lambda_0 P_{t*}^{-1} \left. \frac{\partial}{\partial u} \right|_{\tilde{u}(t)} f_u(q(t)) = h'_t(\lambda_t).$$

We showed that $\tilde{u}$ is a critical point of the endpoint mapping F_{t_1} if and only if there exists a covector curve

$$\lambda_t \in T^*_{q(t)} M, \quad \lambda_t \neq 0, \qquad t \in [0, t_1],$$

such that

$$\dot{\lambda}_t = \vec{h}_{\tilde{u}(t)}(\lambda_t), \tag{20.16}$$

$$\left. \frac{\partial h_u}{\partial u} \right|_{\tilde{u}(t)} (\lambda_t) = 0, \qquad t \in [0, t_1]. \tag{20.17}$$

In particular, any Pontryagin extremal is a critical point of the endpoint mapping. Pontryagin Maximum Principle implies first order necessary optimality conditions (20.16), (20.17). Notice that PMP contains more than these conditions: by PMP, the Hamiltonian $h_u(\lambda_t)$ is not only critical, as in (20.17), but attains maximum along the optimal $\tilde{u}(t)$. We go further to second order conditions.

Asymptotic expansion (20.15) yields the expression for the second differential:

$$D^2_{\tilde{u}} G_{t_1}(v, w)\, a$$
$$= q_0 \circ \left(\int_0^{t_1} g''_\tau(v(\tau), w(\tau))\, d\tau + 2 \iint\limits_{0 \le \tau_2 \le \tau_1 \le t_1} (g'_{\tau_2} v(\tau_2)) \circ g'_{\tau_1} w(\tau_1)\, d\tau_1 d\tau_2 \right) a,$$

where $a \in C^\infty(M)$ and

$$v, w \in \operatorname{Ker} D_{\tilde{u}} G_{t_1} = \operatorname{Ker} D_{\tilde{u}} F_{t_1},$$

i.e.,

$$q_0 \circ \int_0^{t_1} g'_t v(t)\, dt = q_0 \circ \int_0^{t_1} g'_t w(t)\, dt = 0.$$

Now we transform the formula for the second variation via the following decomposition into symmetric and antisymmetric parts.

Exercise 20.4. Let X_τ be a nonautonomous vector field on M. Then

$$\iint\limits_{0\le\tau_2\le\tau_1\le t} X_{\tau_2}\circ X_{\tau_1}\,d\tau_1 d\tau_2$$
$$= \frac{1}{2}\int_0^t X_\tau\,d\tau\circ\int_0^t X_\tau\,d\tau + \frac{1}{2}\iint\limits_{0\le\tau_2\le\tau_1\le t}[X_{\tau_2},X_{\tau_1}]\,d\tau_1 d\tau_2.$$

Choosing $X_t = g'_t v(t)$ and taking into account that $q_0\circ\int_0^{t_1} g'_t v(t)\,dt = 0$, we obtain:

$$q_0\circ\iint\limits_{0\le\tau_2\le\tau_1\le t_1} X_{\tau_2}\circ X_{\tau_1}\,d\tau_1 d\tau_2 = \frac{1}{2}\,q_0\circ\iint\limits_{0\le\tau_2\le\tau_1\le t_1}[X_{\tau_2},X_{\tau_1}]\,d\tau_1 d\tau_2,$$

thus

$$\begin{aligned} &D^2_{\tilde u}G_{t_1}(v,w)a \\ &\quad= q_0\circ\left(\int_0^{t_1} g''_\tau(v(\tau),w(\tau))\,d\tau + \iint\limits_{0\le\tau_2\le\tau_1\le t_1}[g'_{\tau_2}v(\tau_2),g'_{\tau_1}w(\tau_1)]\,d\tau_1\,d\tau_2\right)a \\ &\quad= q_0\circ\left(\int_0^{t_1} g''_\tau(v(\tau),w(\tau))\,d\tau + \int_0^{t_1}\left[\int_0^{\tau_1} g'_{\tau_2}v(\tau_2)\,d\tau_2, g'_{\tau_1}w(\tau_1)\right]\,d\tau_1\right)a.\end{aligned}$$

The first term can conveniently be expressed in Hamiltonian terms since

$$\lambda_0 g_{\tau,u} = \lambda_0 P^{-1}_{\tau *}(f_u - f_{\tilde u(\tau)}) = h_u(\lambda_\tau) - h_{\tilde u(\tau)}(\lambda_\tau).$$

Then

$$\begin{aligned}&\lambda_{t_1}D^2_{\tilde u}F_{t_1}(v,w) = \lambda_0 D^2_{\tilde u}G_{t_1}(v,w)\\ &\quad= \int_0^{t_1} h''_\tau(\lambda_\tau)(v(\tau),w(\tau))\,d\tau + \int_0^{t_1}\lambda_0\left[\int_0^{\tau_1} g'_{\tau_2}v(\tau_2)\,d\tau_2, g'_{\tau_1}w(\tau_1)\right]\,d\tau_1.\end{aligned} \tag{20.18}$$

In order to write also the second term in this expression in the Hamiltonian form, compute the linear on fibers Hamiltonian corresponding to the vector field $g'_\tau v$:

$$\begin{aligned}\lambda_0 g'_\tau v &= \left\langle \lambda_0, P^{-1}_{\tau *}\frac{\partial}{\partial u}f_u v\right\rangle = \left\langle P^{*-1}_\tau\lambda_0, \frac{\partial}{\partial u}f_u v\right\rangle \\ &= \frac{\partial}{\partial u}\left\langle P^{*-1}_\tau\lambda_0, f_u\right\rangle v = \frac{\partial}{\partial u}h_u\circ P^{*-1}_\tau(\lambda_0)v,\end{aligned}$$

where derivatives w.r.t. u are taken at $u = \tilde u(\tau)$. Introducing the Hamiltonian

$$h_{u,\tau}(\lambda) = h_u(P^{*-1}_\tau(\lambda)),$$

we can write the second term in expression (20.18) for the second variation as follows:

$$
\begin{aligned}
&\int_0^{t_1}\int_0^{\tau_1} \lambda_0\left[g'_{\tau_2}v(\tau_2), g'_{\tau_1}w(\tau_1)\right]\, d\tau_2 d\tau_1 \\
&\quad = \int_0^{t_1}\int_0^{\tau_1} \left\{\frac{\partial}{\partial u}h_{u,\tau_2}\, v(\tau_2), \frac{\partial}{\partial u}h_{u,\tau_1}\, w(\tau_1)\right\}(\lambda_0)\, d\tau_2\, d\tau_1 \\
&\quad = \int_0^{t_1}\int_0^{\tau_1} \sigma_{\lambda_0}\left(\frac{\partial}{\partial u}\overrightarrow{h_{u,\tau_2}}\, v(\tau_2), \frac{\partial}{\partial u}\overrightarrow{h_{u,\tau_1}}\, w(\tau_1)\right) d\tau_2\, d\tau_1. \qquad (20.19)
\end{aligned}
$$

Here the derivatives $\dfrac{\partial}{\partial u}h_{u,\tau_i}$ and $\dfrac{\partial}{\partial u}\overrightarrow{h_{u,\tau_i}}$ are evaluated at $u = \tilde{u}(\tau_i)$.

20.4 Necessary optimality conditions

Now we apply our results on second variation and obtain necessary conditions for geometric optimality of an extremal trajectory of system (20.1).

20.4.1 Legendre condition

Fix an admissible control $\tilde{u}$ which is a corank $m \geq 1$ critical point of the endpoint mapping F_{t_1}. (For simplicity, we will suppose that $\tilde{u}(\cdot)$ is piecewise smooth, in concrete problems this is usually the case.) Take any Lagrange multiplier

$$\lambda_0 \in (\operatorname{Im} D_{\tilde{u}}F_{t_1})^{\perp} \setminus \{0\},$$

then

$$\lambda_t = P_t^{*-1}\lambda_0 = \lambda_0 \circ \overrightarrow{\exp}\int_0^t \vec{h}_{\tilde{u}(\tau)}\, d\tau, \qquad t \in [0, t_1],$$

is a trajectory of the Hamiltonian system of PMP. Denote the corresponding quadratic form that evaluates Hessian of the endpoint mapping in (20.18):

$$
\begin{aligned}
&Q : T_{\tilde{u}}\mathcal{U} \to \mathbb{R}, \\
&Q(v) = \int_0^{t_1} h''_\tau(\lambda_\tau)(v(\tau), v(\tau))\, d\tau + \int_0^{t_1} \lambda_0\left[\int_0^{\tau_1} g'_{\tau_2}v(\tau_2)\, d\tau_2, g'_{\tau_1}v(\tau_1)\right] d\tau_1.
\end{aligned}
$$

Then (20.18) reads

$$\lambda_{t_1} \operatorname{Hess}_{\tilde{u}} F_{t_1}(v, v) = Q(v), \qquad v \in \operatorname{Ker} D_{\tilde{u}}F_{t_1}.$$

By Theorem 20.3, if a control $\tilde{u}$ is locally geometrically optimal (i.e., the endpoint mapping F_{t_1} is not locally open at $\tilde{u}$), then there exists a Lagrange multiplier λ_0 such that the corresponding form Q satisfies the condition

$$\operatorname{ind}_- Q|_{\operatorname{Ker} D_{\tilde{u}}F_{t_1}} < m = \operatorname{corank} D_{\tilde{u}}F_{t_1}. \qquad (20.20)$$

The kernel of the differential $D_{\tilde u} F_{t_1}$ is defined by a finite number of scalar linear equations:

$$\operatorname{Ker} D_{\tilde u} F_{t_1} = \left\{ v \in T_{\tilde u}\mathcal{U} \mid q_0 \circ \int_0^{t_1} g_t' v(t)\, dt = 0 \right\},$$

i.e., it has a finite codimension in $T_{\tilde u}\mathcal{U}$. Thus inequality (20.20) implies that

$$\operatorname{ind}_- Q < +\infty$$

for the corresponding extremal λ_t. If we take the extremal $-\lambda_t$ projecting to the same extremal curve $q(t)$, then we obtain a form Q with a finite positive index. So local geometric optimality of $\tilde u$ implies finiteness of index of the form Q for some Lagrange multiplier λ_0.

Proposition 20.1. *If the quadratic form Q has a finite positive index, then there holds the following inequality along the corresponding extremal λ_t:*

$$h_t''(\lambda_t)(v, v) \leq 0, \qquad t \in [0, t_1], \quad v \in \mathbb{R}^m. \tag{20.21}$$

Inequality (20.21) is called *Legendre condition*.

In particular, if a trajectory $q(t)$ is locally geometrically optimal, then Legendre condition holds for some extremal λ_t, $\pi(\lambda_t) = q(t)$. However, necessity of Legendre condition for optimality follows directly from the maximality condition of PMP (exercise). But we will need in the sequel the stronger statement related to index of Q as in Proposition 20.1.

Notice once more that in the study of geometric optimality, all signs may be reversed: multiplying λ_t by -1, we obtain a quadratic form with $\operatorname{ind}_- Q < +\infty$ and the reversed Legendre condition $h_t''(\lambda_t)(v, v) \geq 0$. Of course, this is true also for subsequent conditions related to geometric optimality.

Now we prove Proposition 20.1.

Proof. Take a smooth vector function

$$v : \mathbb{R} \to \mathbb{R}^m, \qquad \operatorname{supp} v \subset [0, 1],$$

and introduce a family of variations of the form:

$$v_{\bar\tau,\varepsilon}(\tau) = v\left(\frac{\tau - \bar\tau}{\varepsilon}\right), \qquad \bar\tau \in [0, t_1), \quad \varepsilon > 0.$$

Notice that the vector function $v_{\bar\tau,\varepsilon}$ is concentrated at the segment $[\bar\tau, \bar\tau + \varepsilon]$. Compute asymptotics of the form Q on the family introduced:

$$\begin{aligned} Q(v_{\bar\tau,\varepsilon}) &= \varepsilon \int_0^1 h''_{\bar\tau+\varepsilon s}(\lambda_{\bar\tau+\varepsilon s})(v(s), v(s))\, ds \\ &\qquad + \varepsilon^2 \int_0^1 \lambda_0 \left[\int_0^1 g'_{\bar\tau+\varepsilon s_2} v(s_2)\, ds_2,\, g'_{\bar\tau+\varepsilon s_1} v(s_1) \right] ds_1 \\ &= \varepsilon \int_0^1 h''_{\bar\tau}(\lambda_{\bar\tau})(v(s), v(s))\, ds + O(\varepsilon^2), \end{aligned} \tag{20.22}$$

where $O(\varepsilon^2)$ is uniform w.r.t. v in the L_∞ norm.

Suppose, by contradiction, that

$$h''_{\bar\tau}(\lambda_{\bar\tau})(v, v) > 0$$

for some $\bar\tau \in [0, t_1)$, $v \in \mathbb{R}^m$. In principal axes, the quadratic form becomes a sum of squares:

$$h''_{\bar\tau}(\lambda_{\bar\tau})(v, v) = \sum_{i=1}^{m} \alpha^i_{\bar\tau}(v^i)^2$$

with at least one coefficient

$$\alpha^i_{\bar\tau} > 0.$$

Choose a vector function v of the form

$$v(s) = \begin{pmatrix} v^1(s) \\ \dots \\ v^i(s) \\ \dots \\ v^m(s) \end{pmatrix} = \begin{pmatrix} 0 \\ \dots \\ v^i(s) \\ \dots \\ 0 \end{pmatrix}$$

with the only nonzero component $v^i(s)$. For sufficiently small $\varepsilon > 0$, $Q(v_{\bar\tau,\varepsilon}) > 0$. But for any fixed $\bar\tau$ and ε, the space of vector functions $v_{\bar\tau,\varepsilon}$ is infinite-dimensional. Thus the quadratic form Q has an infinite positive index. By contradiction, the proposition follows. $\square$

20.4.2 Regular extremals

We proved that Legendre condition is necessary for finiteness of positive index of the quadratic form Q. The corresponding sufficient condition is given by the *strong Legendre condition*:

$$\begin{aligned} &h''_t(\lambda_t)(v, v) < -\alpha|v|^2, \qquad t \in [0, t_1], \quad v \in \mathbb{R}^m, \\ &\alpha > 0. \end{aligned} \tag{20.23}$$

An extremal that satisfies the strong Legendre condition is called *regular* (notice that this definition is valid only in the case of open space of control parameters U, where Legendre condition is related to maximality of h_u).

Proposition 20.2. *If λ_t, $t \in [0, t_1]$, is a regular extremal, then:*

(1) *For any $\tau \in [0, t_1)$ there exists $\varepsilon > 0$ such that the form Q is negative on the space $L^m_\infty[\tau, \tau + \varepsilon]$,*

(2) *The form Q has a finite positive index on the space $T_{\tilde u}\mathcal{U} = L^m_\infty[0, t_1]$.*

Proof. (1) We have:

$$\begin{aligned}
Q(v) &= Q_1(v) + Q_2(v), \\
Q_1(v) &= \int_0^{t_1} h''_\tau(\lambda_\tau)(v(\tau), v(\tau))\, d\tau, \\
Q_2(v) &= \int_0^{t_1} \lambda_0 \left[\int_0^{\tau_1} g'_{\tau_2} v(\tau_2)\, d\tau_2, g'_{\tau_1} v(\tau_1) \right] d\tau_1 \\
&= \int_0^{t_1} \sigma_{\lambda_0} \left(\int_0^{\tau_1} \frac{\partial}{\partial u} \overrightarrow{h_{u,\tau_2}}\, v(\tau_2), \frac{\partial}{\partial u} \overrightarrow{h_{u,\tau_1}}\, v(\tau_1) \right) d\tau_2\, d\tau_1.
\end{aligned}$$

By continuity of $h''_\tau(\lambda_\tau)$ w.r.t. τ, the strong Legendre condition implies that

$$Q_1\left(v|_{[\tau,\tau+\varepsilon]} \right) < -\frac{\alpha}{2}\, \varepsilon\, \|v\|^2_{L_2}$$

for small $\varepsilon > 0$. It follows by the same argument as in (20.22) that the term Q_1 dominates on short segments:

$$Q_2\left(v|_{[\tau,\tau+\varepsilon]} \right) = O(\varepsilon^2)\|v\|^2_{L_2}, \qquad \varepsilon \to 0,$$

thus

$$Q\left(v|_{[\tau,\tau+\varepsilon]} \right) < 0$$

for sufficiently small $\varepsilon > 0$ and all $v \in L^m_\infty[0, t_1]$, $v \neq 0$.

(2) We show that the form Q is negative on a finite codimension subspace in $L^m_\infty[0, t_1]$, this implies that $\operatorname{ind}_+ Q < \infty$.

By the argument used in the proof of item (1), any point $\tau \in [0, t_1]$ can be covered by a segment $[\tau - \varepsilon, \tau + \varepsilon]$ such that the form Q is negative on the space $L^m_\infty[\tau - \varepsilon, \tau + \varepsilon]$. Choose points $0 = \tau_0 < \tau_1 < \cdots < \tau_N = t_1$ such that Q is negative on the spaces $L^m_\infty[\tau_{i-1}, \tau_i]$, $i = 1, \ldots, N$. Define the following finite codimension subspace of $L^m_\infty[0, t_1]$:

$$L = \left\{ v \in L^m_\infty[0, t_1] \mid \lambda_0 \circ \int_{\tau_{i-1}}^{\tau_i} \frac{\partial}{\partial u} \overrightarrow{h_{u,\tau}}\, v(\tau)\, d\tau = 0, \quad i = 1, \ldots, N \right\}.$$

For any $v \in L$, $v \neq 0$,

$$Q(v) = \sum_{i=1}^{N} Q\left(v|_{[\tau_{i-1},\tau_i]} \right) < 0.$$

Thus L is the required finite codimension negative subspace of the quadratic form Q. Consequently, the form Q has a finite positive index. □

Propositions 20.1 and 20.2 relate sign-definiteness of the form $h''_\tau(\lambda_t)$ with sign-definiteness of the form Q, thus, in the corank one case, with local geometric optimality of the reference control $\tilde{u}$ (via Theorem 20.1). Legendre condition is necessary for finiteness of $\operatorname{ind}_+ Q$, thus for local geometric optimality of $\tilde{u}$. On the other hand, strong Legendre condition is sufficient for negativeness of Q on short segments, thus for local finite-dimensional optimality of $\tilde{u}$ on short segments. Notice that we can easily obtain a much stronger result from the theory of fields of extremals (Section 17.1). Indeed, under the strong Legendre condition the maximized Hamiltonian of PMP is smooth, and Corollary 17.1 gives local optimality on short segments (in $C([0,t_1],M)$ topology, thus in $L_\infty([0,t_1],U)$ topology and in topology of convergence on finite-dimensional submanifolds in $\mathcal{U}$).

20.4.3 Singular extremals

Now we consider the case where the second derivative of the Hamiltonian h_u vanishes identically along the extremal, in particular, the case of affine in control systems $\dot{q} = f_0(q) + \sum_{i=1}^m u_i f_i(q)$. So we assume that an extremal λ_t satisfies the identity

$$h_t''(\lambda_t) \equiv 0, \qquad t \in [0,t_1]. \tag{20.24}$$

Such an extremal is called *totally singular*. As in the case of regular extremals, this definition is valid only if the set of control parameters U is open.

For a totally singular extremal, expression (20.18) for the Hessian takes the form:

$$\lambda_{t_1} \operatorname{Hess}_{\tilde{u}} F_{t_1}(v_1,v_2) = \lambda_0 \int_0^{t_1} \left[\int_0^{\tau_1} g'_{\tau_2} v_1(\tau_2)\, d\tau_2, g'_{\tau_1} v_2(\tau_1) \right] d\tau_1.$$

In order to find the dominating term of the Hessian (concentrated on the diagonal $\tau_1 = \tau_2$), we integrate by parts. Denote

$$w_i(\tau) = \int_\tau^{t_1} v_i(s)\, ds,$$
$$\dot{g}'_\tau = \frac{d}{d\tau} g'_\tau.$$

Then

$$\begin{aligned}
&\lambda_{t_1} \operatorname{Hess}_{\tilde{u}} F_{t_1}(v_1,v_2) \\
&\quad = \lambda_0 \left(\int_0^{t_1} \left[-g'_{\tau_1} w_1(\tau_1) + g'_0 w_1(0) + \int_0^{\tau_1} \dot{g}'_{\tau_2} w_1(\tau_2)\, d\tau_2,\ g'_{\tau_1} v_2(\tau_1) \right] d\tau_1 \right) \\
&\quad = \lambda_0 \left(- \int_0^{t_1} [g'_\tau w_1(\tau), g'_\tau v_2(\tau)]\, d\tau + [g'_0 w_1(0), g'_0 w_2(0)] \right. \\
&\qquad + \left[g'_0 w_1(0), \int_0^{t_1} \dot{g}'_\tau w_2(\tau)\, d\tau \right] + \int_0^{t_1} [\dot{g}'_\tau w_1(\tau), g'_\tau w_2(\tau)]\, d\tau \\
&\qquad \left. + \int_0^{t_1} \left[\dot{g}'_{\tau_2} w_1(\tau_2), \int_{\tau_2}^{t_1} \dot{g}'_{\tau_1} w_2(\tau_1)\, d\tau_1 \right] d\tau_2 \right).
\end{aligned} \tag{20.25}$$

We integrate by parts also the admissibility condition $q_0 \circ \int_0^{t_1} g_t' v_i(t)\, dt = 0$:

$$q_0 \circ \left(\int_0^{t_1} \dot{g}_t' w_i(t)\, dt + g_0' w_i(0) \right) = 0. \tag{20.26}$$

In the sequel we take variations v_i subject to the restriction

$$w_i(0) = \int_0^{t_1} v_i(t)\, dt = 0, \qquad i = 1,\ 2.$$

We assume that functions $v(s)$ used in construction of the family $v_{\bar{\tau},\varepsilon}(\tau) = v\left(\frac{\tau - \bar{\tau}}{\varepsilon}\right)$ satisfy the equality

$$\int_0^1 v(s)\, ds = 0,$$

then the primitive

$$w(s) = \int_0^s v(s')\, ds'$$

is also concentrated at the segment $[0, 1]$. Then the last term in expression (20.25) of the Hessian vanishes, and equality (20.26) reduces to

$$q_0 \circ \int_0^{t_1} \dot{g}_t' w_i(t)\, dt = 0.$$

Asymptotics of the Hessian on the family $v_{\bar{\tau},\varepsilon}$ has the form:

$$\lambda_{t_1} \operatorname{Hess}_{\tilde{u}} F_{t_1}(v_{\bar{\tau},\varepsilon}, v_{\bar{\tau},\varepsilon}) = Q(v_{\bar{\tau},\varepsilon}) = \varepsilon^2 \lambda_0 \int_0^1 [g_{\bar{\tau}}' w(s), g_{\bar{\tau}}' v(s)]\, ds + O(\varepsilon^3).$$

The study of this dominating term provides necessary optimality conditions.

Proposition 20.3. *Let λ_t, $t \in [0, t_1]$, be a totally singular extremal. If the quadratic form $Q = \lambda_{t_1} \operatorname{Hess}_{\tilde{u}} F_{t_1}$ has a finite positive index, then*

$$\lambda_0 [g_t' v_1, g_t' v_2] = 0 \qquad \forall\ v_1,\ v_2 \in \mathbb{R}^m,\ t \in [0, t_1]. \tag{20.27}$$

Equality (20.27) is called *Goh condition*. It can be written also as follows:

$$\lambda_t \left[\frac{\partial f_u}{\partial u} v_1, \frac{\partial f_u}{\partial u} v_2 \right] = 0,$$

or in Hamiltonian form:

$$\left\{ \frac{\partial h_u}{\partial u_i}, \frac{\partial h_u}{\partial u_j} \right\} (\lambda_t) = \sigma_{\lambda_t} \left(\frac{\partial}{\partial u_i} \vec{h}_u, \frac{\partial}{\partial u_j} \vec{h}_u \right) = 0,$$
$$i,\ j = 1, \dots, m, \quad t \in [0, t_1].$$

As before, derivatives w.r.t. u are evaluated at $u = \tilde{u}(t)$.

Now we prove Proposition 20.3.

Proof. Take a smooth vector function $v : \mathbb{R} \to \mathbb{R}^m$ concentrated at the segment $[0, 2\pi]$ such that $\int_0^{2\pi} v(s)\, ds = 0$, and construct as before the variation of controls

$$v_{\bar{\tau},\varepsilon}(\tau) = v\left(\frac{\tau - \bar{\tau}}{\varepsilon}\right).$$

Then

$$Q(v_{\bar{\tau},\varepsilon}) = \varepsilon^2 \int_0^{2\pi} \lambda_0[g'_{\bar{\tau}} w(s), g'_{\bar{\tau}} v(s)]\, ds + O(\varepsilon^3)\|v\|^2_{L_2},$$

where $w(s) = \int_0^s v(s')\, ds'$. The leading term is the integral

$$\int_0^{2\pi} \lambda_0[g'_{\bar{\tau}} w(s), g'_{\bar{\tau}} v(s)]\, ds = \int_0^{2\pi} \omega(w(s), v(s))\, ds, \tag{20.28}$$
$$\omega(x, y) = \lambda_0[g'_{\bar{\tau}} x, g'_{\bar{\tau}} y], \qquad x,\ y \in \mathbb{R}^m,$$

notice that the bilinear skew-symmetric form ω enters Goh condition (20.27). In order to prove the proposition, we show that if $\omega \not\equiv 0$, then the leading term (20.28) of Hessian has a positive subspace of arbitrarily large dimension.

Let $\omega \not\equiv 0$ for some $\bar{\tau} \in [0, t_1]$, then $\operatorname{rank} \omega = 2l > 0$, and there exist coordinates in $\mathbb{R}^m$ in which the form ω reads

$$\omega(x, y) = \sum_{i=1}^{l} (x^i y^{i+l} - x^{i+l} y^i),$$
$$x = \begin{pmatrix} x^1 \\ \cdots \\ x_m \end{pmatrix}, \qquad y = \begin{pmatrix} y^1 \\ \cdots \\ y_m \end{pmatrix}.$$

Take a vector function v of the form

$$v(s) = \begin{pmatrix} v^1(s) \\ 0 \\ \cdots \\ 0 \\ v^{l+1}(s) \\ 0 \\ \cdots \\ 0 \end{pmatrix},$$
$$v^1(s) = \sum_{k>0} \xi_k \cos ks, \qquad v^{l+1}(s) = \sum_{k>0} \eta_k \sin ks.$$

Substituting $v(s)$ to (20.28), we obtain:

$$\int_0^{2\pi} \omega(w(s), v(s))\, ds = -2\pi \sum_{k>0} \frac{1}{k} \xi_k \eta_k.$$

This form obviously has a positive subspace of infinite dimension.

For an arbitrarily great N, we can find an N-dimensional positive space L_N for form (20.28). There exists $\varepsilon_N > 0$ such that $Q(v_{\bar{\tau},\varepsilon_N}) > 0$ for any $v \in L_N$. Thus $\operatorname{ind}_+ Q = \infty$. By contradiction, Goh condition follows. □

Exercise 20.5. Show that Goh condition holds not only for piecewise smooth, but also for measurable bounded extremal control $\tilde{u}$ at Lebesgue points.

Goh condition imposes a strong restriction on a totally singular optimal control $\tilde{u}$. For a totally singular extremal, the first two terms in (20.25) vanish by Goh condition, and the third term in (20.25) is zero under the condition $w(0) = 0$, which we assume fulfilled. Thus the expression for Hessian (20.25) reduces to the following two terms:

$$\lambda_{t_1} \operatorname{Hess}_{\tilde{u}} F_{t_1}(v,v) = Q(v)$$
$$= \lambda_0 \left(\int_0^{t_1} [\dot{g}'_\tau w(\tau), g'_\tau w(\tau)] \, d\tau + \int_0^{t_1} \left[\dot{g}'_{\tau_2} w(\tau_2), \int_{\tau_2}^{t_1} \dot{g}'_{\tau_1} w(\tau_1) \, d\tau_1 \right] d\tau_2 \right). \tag{20.29}$$

Suppose that the quadratic form Q has a finite positive index. Then by the same argument as in Proposition 20.1 we prove one more pointwise condition:

$$\lambda_0 [\dot{g}'_t v, g'_t v] \leq 0 \qquad \forall\, v \in \mathbb{R}^m, \quad t \in [0, t_1]. \tag{20.30}$$

This inequality is called *generalized Legendre condition*.

Notice that generalized Legendre condition can be rewritten in Hamiltonian terms:

$$\left\{ \left\{ h_{\tilde{u}(t)}, h'_t v \right\}, h'_t v \right\} (\lambda_t) + \left\{ h''_t(\dot{\tilde{u}}(t), v), h'_t v \right\} (\lambda_t) \leq 0, \qquad v \in \mathbb{R}^m, \quad t \in [0, t_1].$$

This easily follows from the equalities:

$$g'_t v = P_{t*}^{-1} \frac{\partial f_u}{\partial u} v = \operatorname{Ad} P_t \frac{\partial f_u}{\partial u} v,$$
$$\dot{g}'_t v = \frac{d}{dt} \overrightarrow{\exp} \int_0^t \operatorname{ad} f_{\tilde{u}(\tau)} \, d\tau \, \frac{\partial f_u}{\partial u} v$$
$$= P_{t*}^{-1} \left[f_{\tilde{u}(t)}, \frac{\partial f_u}{\partial u} v \right] + P_{t*}^{-1} \frac{\partial^2 f_u}{\partial u^2} (\dot{\tilde{u}}(t), v).$$

The strong version (20.31) of generalized Legendre condition plays in the totally singular case the role similar to that of the strong Legendre condition in the regular case.

Proposition 20.4. *Let an extremal λ_t be totally singular, satisfy Goh condition, the* strong generalized Legendre condition*:*

$$\left\{ \left\{ h_{\tilde{u}(t)}, h'_t v \right\}, h'_t v \right\} (\lambda_t) + \left\{ h''_t(\dot{\tilde{u}}(t), v), h'_t v \right\} (\lambda_t) \leq -\alpha |v|^2,$$
$$v \in \mathbb{R}^m, \quad t \in [0, t_1], \tag{20.31}$$

for some $\alpha > 0$, and the following nondegeneracy condition:

$$\text{the linear mapping } \left.\frac{\partial f_u(q_0)}{\partial u}\right|_{\tilde{u}(0)} : \mathbb{R}^m \to T_{q_0}M \text{ is injective.} \tag{20.32}$$

Then the quadratic form $Q|_{\operatorname{Ker} D_{\tilde{u}} F_t}$ is negative on short segments and has a finite positive index on $L_\infty^m[0, t_1]$.

Proof. This proposition is proved similarly to Proposition 20.2. In decomposition (20.25) the first two terms vanish by Goh condition, and the fourth term is negative and dominates on short segments. The third term is small on short segments since

$$q_0 \circ g_0' w_1(0) = \left.\frac{\partial f_u(q_0)}{\partial u}\right|_{\tilde{u}(0)} w_1(0),$$

and condition (20.32) allows to express $w_1(0)$ through the integral $\displaystyle\int_0^{t_1} w_1(\tau)\, d\tau$ on the kernel of $D_{\tilde{u}} F_{t_1}$, which is defined by equality (20.26). □

We call an extremal that satisfies all hypotheses of Proposition 20.4 a *nice singular extremal*.

20.4.4 Necessary conditions

Summarizing the results obtained in this section, we obtain the following necessary conditions for the quadratic form Q to have a finite positive index.

Theorem 20.4. *Let a piecewise smooth control $\tilde{u} = \tilde{u}(t)$, $t \in [0, t_1]$, be a critical point of the endpoint mapping F_{t_1}. Let a covector $\lambda_{t_1} \in T^*_{F_{t_1}(\tilde{u})}M$ be a Lagrange multiplier:*

$$\lambda_{t_1} D_{\tilde{u}} F_{t_1} = 0, \qquad \lambda_{t_1} \neq 0.$$

If the quadratic form Q has a finite positive index, then:

(I) *The trajectory λ_t of the Hamiltonian system of PMP*

$$\dot{\lambda}_t = \vec{h}_{\tilde{u}(t)}(\lambda_t),$$
$$h_u(\lambda) = \langle \lambda, f_u(q) \rangle,$$

satisfies the equality

$$h_t'(\lambda_t) = 0, \qquad t \in [0, t_1].$$

(II.1) *Legendre condition is satisfied:*

$$h_t''(\lambda_t)(v, v) \leq 0, \qquad v \in \mathbb{R}^m, \quad t \in [0, t_1].$$

(II.2) *If the extremal* λ_t *is totally singular:*

$$h_t''(\lambda_t)(v,v) \equiv 0, \qquad v \in \mathbb{R}^m, \quad t \in [0,t_1],$$

then there hold Goh condition:

$$\{h_t' v_1, h_t' v_2\}(\lambda_t) \equiv 0, \qquad v_1,\ v_2 \in \mathbb{R}^m, \quad t \in [0,t_1], \tag{20.33}$$

and generalized Legendre condition:

$$\left\{\left\{h_{\tilde{u}(t)}, h_t' v\right\}, h_t' v\right\}(\lambda_t) + \left\{h_t''(\dot{\tilde{u}}(t), v), h_t' v\right\}(\lambda_t) \leq 0, \qquad v \in \mathbb{R}^m, \quad t \in [0,t_1]. \tag{20.34}$$

Remark. If the Hamiltonian h_u is affine in u (for affine in control systems), then the second term in generalized Legendre condition (20.34) vanishes.

Recall that the corresponding sufficient conditions for finiteness of index of the second variation are given in Propositions 20.2 and 20.4.

Combining Theorems 20.4 and 20.3, we come to the following necessary optimality conditions.

Corollary 20.1. *If a piecewise smooth control* $\tilde{u} = \tilde{u}(t)$ *is locally geometrically optimal for control system* (20.14)*, then first-order conditions* (I) *and second-order conditions* (II.1)*,* (II.2) *of Theorem* 20.4 *hold along the corresponding extremal* λ_t*.*

20.5 Applications

In this section we apply the second order optimality conditions obtained to particular problems.

20.5.1 Abnormal sub-Riemannian geodesics

Consider the sub-Riemannian problem:

$$\begin{aligned}
&\dot{q} = \sum_{i=1}^m u_i f_i(q), \qquad q \in M, \quad u = (u_1, \dots, u_m) \in \mathbb{R}^m,\\
&q(0) = q_0, \qquad q(1) = q_1,\\
&J(u) = \frac{1}{2}\int_0^1 \sum_{i=1}^m u_i^2\, dt = \frac{1}{2}\int_0^1 |u|^2\, dt \to \min.
\end{aligned}$$

The study of optimality is equivalent to the study of boundary of attainable set for the extended system:

$$\begin{cases}
\dot{q} = \sum_{i=1}^m u_i f_i(q), & q \in M,\\
\dot{y} = \frac{1}{2}|u|^2, & y \in \mathbb{R}.
\end{cases}$$

The Hamiltonian is

$$h_u(\lambda, \nu) = \sum_{i=1}^{m} u_i \langle \lambda, f_i(q) \rangle + \frac{\nu}{2}|u|^2, \qquad \lambda \in T^*M, \quad \nu \in \mathbb{R}^* = \mathbb{R}.$$

The parameter ν is constant along any geodesic (extremal). If $\nu \neq 0$ (the normal case), then extremal control can be recovered via PMP. In the sequel we consider the abnormal case:

$$\nu = 0.$$

Then

$$h_u(\lambda) = h_u(\lambda, 0) = \sum_{i=1}^{m} u_i h_i(\lambda),$$
$$h_i(\lambda) = \langle \lambda, f_i(q) \rangle, \qquad i = 1, \dots, m.$$

The maximality condition of PMP does not determine controls in the abnormal case directly (abnormal extremals are totally singular). What that condition implies is that abnormal extremals λ_t satisfy, in addition to the Hamiltonian system

$$\dot{\lambda}_t = \sum_{i=1}^{m} u_i(t) \vec{h}_i(\lambda_t),$$

the following identities:

$$h_i(\lambda_t) \equiv 0, \qquad i = 1, \dots, m.$$

We apply second order conditions. As we already noticed, Legendre condition degenerates. Goh condition reads:

$$\{h_i, h_j\}(\lambda_t) \equiv 0, \qquad i, \ j = 1, \dots m.$$

If an abnormal extremal λ_t projects to an optimal trajectory $q(t)$, then at any point q of this trajectory there exists a covector

$$\lambda \in T_q^*M, \qquad \lambda \neq 0,$$

such that

$$\langle \lambda, f_i(q) \rangle = 0, \qquad i = 1, \dots, m,$$
$$\langle \lambda, [f_i, f_j](q) \rangle = 0, \qquad i, \ j = 1, \dots, m.$$

Consequently, if

$$\operatorname{span}(f_i(q), \ [f_i, f_j](q)) = T_q M, \tag{20.35}$$

then no locally optimal strictly abnormal trajectory passes through the point q. An extremal trajectory is called *strictly abnormal* if it is a projection of an abnormal extremal and it is not a projection of a normal extremal. Notice that in the case corank > 1 extremal trajectories can be abnormal but not strictly abnormal (i.e., can be abnormal and normal simultaneously), there can be two Lagrange multipliers $(\lambda, 0)$ and $(\lambda', \nu \neq 0)$. Small arcs of such trajectories are always local minimizers since the normal Hamiltonian $H = \frac{1}{2}\sum_{i=1}^{m} h_i^2$ is smooth (see Corollary 17.1).

Distributions $\operatorname{span}(f_i(q))$ that satisfy condition (20.35) are called *2-generating*. E.g., the left-invariant bracket-generating distributions appearing in the sub-Riemannian problem on a compact Lie group in Section 19.2 and Exercise 19.1 are 2-generating, thus there are no optimal strictly abnormal trajectories in those problems.

Example 20.1. Consider the following left-invariant sub-Riemannian problem on $\mathrm{GL}(n)$ with a natural cost:

$$\dot{Q} = QV, \qquad Q \in \mathrm{GL}(n), \quad V = V^*, \tag{20.36}$$

$$J(V) = \frac{1}{2}\int_0^1 \operatorname{tr} V^2 \, dt \to \min. \tag{20.37}$$

Exercise 20.6. Show that normal extremals in this problem are products of 2 one-parameter subgroups. (Hint: repeat the argument of Section 19.2.) Then it follows that any matrix can be represented as a product of two exponentials $e^V e^{(V-V^*)/2}$. Notice that not any matrix can be represented as a single exponential e^V.

There are many abnormal extremals in problem (20.36), (20.37), but they are never optimal. Indeed, the distribution defined by the right-hand side of the system is 2-generating. We have

$$[QV_1, QV_2] = Q[V_1, V_2],$$

and if matrices V_i are symmetric then their commutator $[V_1, V_2]$ is antisymmetric. Moreover, any antisymmetric matrix appears in this way. But any $n \times n$ matrix is a sum of symmetric and antisymmetric matrices. Thus the distribution $\{QV \mid V^* = V\}$ is 2-generating, and strictly abnormal extremal trajectories are not optimal.

20.5.2 Local controllability of bilinear system

Consider a *bilinear* control system of the form

$$\dot{x} = Ax + uBx + vb, \qquad u, \ v \in \mathbb{R}, \quad x \in \mathbb{R}^n. \tag{20.38}$$

We are interested, when the system is *locally controllable* at the origin, i.e.,

$$0 \in \operatorname{int} \mathcal{A}_0(t) \qquad \forall \, t > 0.$$

Negation of necessary conditions for geometric optimality gives sufficient conditions for local controllability. Now we apply second order conditions of Corollary 20.1 to our system. Suppose that

$$0 \in \partial \mathcal{A}_0(t) \text{ for some } t > 0.$$

Then the reference trajectory $x(t)$ is geometrically optimal, thus it satisfies PMP. The control-dependent Hamiltonian is

$$h_{u,v}(p,x) = pAx + upBx + vpb, \qquad \lambda = (p,x) \in T^*\mathbb{R}^n = \mathbb{R}^{n*} \times \mathbb{R}^n.$$

The vertical part of the Hamiltonian system along the reference trajectory reads:

$$\dot{p} = -pA, \qquad p \in \mathbb{R}^{n*}. \tag{20.39}$$

It follows from PMP that

$$p(\tau)b = p(0)e^{-A\tau}b \equiv 0, \qquad \tau \in [0,t],$$

i.e.,

$$p(0)A^i b = 0, \qquad i = 0, \dots, n-1, \tag{20.40}$$

for some covector $p(0) \neq 0$, thus

$$\operatorname{span}(b, Ab, \dots, A^{n-1}b) \neq \mathbb{R}^n.$$

We pass to second order conditions. Legendre condition degenerates since the system is affine in control, and Goh condition takes the form:

$$p(\tau)Bb \equiv 0, \qquad \tau \in [0,t].$$

Differentiating this identity by virtue of Hamiltonian system (20.39), we obtain, in addition to (20.40), new restrictions on $p(0)$:

$$p(0)A^i Bb = 0, \qquad i = 0, \dots, n-1.$$

Generalized Legendre condition degenerates.

Summing up, the inequality

$$\operatorname{span}(b, Ab, \dots, A^{n-1}b, Bb, ABb, \dots, A^{n-1}Bb) \neq \mathbb{R}^n$$

is necessary for geometric optimality of the trajectory $x(t) \equiv 0$. In other words, the equality

$$\operatorname{span}(b, Ab, \dots, A^{n-1}b, Bb, ABb, \dots, A^{n-1}Bb) = \mathbb{R}^n$$

is sufficient for local controllability of bilinear system (20.38) at the origin.

20.6 Single-input case

In this section we apply first- and second-order optimality conditions to the simplest (and the hardest to control) case with scalar input:

$$\dot{q} = f_0(q) + uf_1(q), \qquad u \in [\alpha, \beta] \subset \mathbb{R}, \quad q \in M. \tag{20.41}$$

Since the system is affine in control, Legendre condition automatically degenerates. Further, control is one-dimensional, thus Goh condition is trivial. Although, generalized Legendre condition works (we write it down later). We apply first Pontryagin Maximum Principle. Introduce the linear on fibers Hamiltonians

$$h_i(\lambda) = \langle \lambda, f_i(q) \rangle, \qquad i = 0,\ 1,$$

then the Hamiltonian of the system reads

$$h_u(\lambda) = h_0(\lambda) + u h_1(\lambda).$$

We look for extremals corresponding to a control

$$u(t) \in (\alpha, \beta). \tag{20.42}$$

The Hamiltonian system of PMP reads

$$\dot{\lambda}_t = \vec{h}_0(\lambda_t) + u(t)\vec{h}_1(\lambda_t), \tag{20.43}$$

and maximality condition reduces to the identity

$$h_1(\lambda_t) \equiv 0. \tag{20.44}$$

Extremals λ_t are Lipschitzian, so we can differentiate the preceding identity:

$$\dot{h}_1(\lambda_t) = \frac{d}{dt} h_1(\lambda_t) = \{h_0 + u(t)h_1, h_1\}(\lambda_t) = \{h_0, h_1\}(\lambda_t) \equiv 0. \tag{20.45}$$

Equalities (20.44), (20.45), which hold identically along any extremal λ_t that satisfies (20.42), do not allow us to determine the corresponding control $u(t)$. In order to obtain an equality involving $u(t)$, we proceed with differentiation:

$$\begin{aligned} \ddot{h}_1(\lambda_t) &= \{h_0 + u(t)h_1, \{h_0, h_1\}\}(\lambda_t) \\ &= \{h_0, \{h_0, h_1\}\}(\lambda_t) + u(t)\{h_1, \{h_0, h_1\}\}(\lambda_t) \equiv 0. \end{aligned}$$

Introduce the notation for Hamiltonians:

$$h_{i_1 i_2 \dots i_k} = \{h_{i_1}, \{h_{i_2}, \dots, \{h_{i_{k-1}}, h_{i_k}\} \dots\}\}, \qquad i_j \in \{0, 1\}.$$

then any extremal λ_t with (20.42) satisfies the identities

$$h_1(\lambda_t) = h_{01}(\lambda_t) \equiv 0, \tag{20.46}$$

$$h_{001}(\lambda_t) + u(t) h_{101}(\lambda_t) \equiv 0. \tag{20.47}$$

If $h_{101}(\lambda_t) \neq 0$, then extremal control $u = u(\lambda_t)$ is uniquely determined by λ_t:

$$u(\lambda_t) = -\frac{h_{001}(\lambda_t)}{h_{101}(\lambda_t)}. \tag{20.48}$$

Notice that the regularity condition $h_{101}(\lambda_t) \neq 0$ is closely related to generalized Legendre condition. Indeed, for the Hamiltonian $h_u = h_0 + uh_1$ generalized Legendre condition takes the form

$$\{\{h_0 + uh_1, h_1\}, h_1\}(\lambda_t) = -h_{101}(\lambda_t) \leq 0,$$

i.e.,

$$h_{101}(\lambda_t) \geq 0.$$

And if this inequality becomes strong, then the control is determined by relation (20.48).

Assume that $h_{101}(\lambda_t) \neq 0$ and plug the control $u(\lambda) = -h_{001}(\lambda)/h_{101}(\lambda)$ given by (20.48) to the Hamiltonian system (20.43):

$$\dot{\lambda} = \vec{h}_0(\lambda) + u(\lambda)\vec{h}_1(\lambda). \tag{20.49}$$

Any extremal with (20.42) and $h_{101}(\lambda_t) \neq 0$ is a trajectory of this system.

Lemma 20.3. *The manifold*

$$\{\lambda \in T^*M \mid h_1(\lambda) = h_{01}(\lambda) = 0, \ h_{101}(\lambda) \neq 0\} \tag{20.50}$$

is invariant for system (20.49).

Proof. Notice first of all that the regularity condition $h_{101}(\lambda) \neq 0$ guarantees that conditions (20.50) determine a smooth manifold since $d_\lambda h_1$ and $d_\lambda h_{01}$ are linearly independent. Introduce a Hamiltonian

$$\varphi(\lambda) = h_0(\lambda) + u(\lambda)h_1(\lambda).$$

The corresponding Hamiltonian vector field

$$\vec{\varphi}(\lambda) = \vec{h}_0(\lambda) + u(\lambda)\,\vec{h}_1(\lambda) + h_1(\lambda)\,\vec{u}(\lambda)$$

coincides with field (20.49) on the manifold $\{h_1 = h_{01} = 0\}$, so it is sufficient to show that $\vec{\varphi}$ is tangent to this manifold.

Compute derivatives by virtue of the field $\vec{\varphi}$:

$$\begin{aligned}
\dot{h}_1 &= \{h_0 + uh_1, h_1\} = h_{01} - (\vec{h}_1 u)h_1, \\
\dot{h}_{01} &= \{h_0 + uh_1, h_{01}\} = \underbrace{h_{001} + uh_{101}}_{\equiv 0} - (\vec{h}_{01}u)h_1 = -(\vec{h}_{01}u)h_1.
\end{aligned}$$

The linear system with variable coefficients for $h_1(t) = h_1(\lambda_t)$, $h_{01}(t) = h_{01}(\lambda_t)$

$$\begin{cases}
\dot{h}_1(t) = h_{01}(t) - (\vec{h}_1 u)(\lambda_t)\,h_1(t), \\
\dot{h}_{01}(t) = -(\vec{h}_{01}u)(\lambda_t)\,h_1(t)
\end{cases}$$

has a unique solution. Thus for the initial condition $h_1(0) = h_{01}(0) = 0$ we obtain the solution $h_1(t) = h_{01}(t) \equiv 0$. So manifold (20.50) is invariant for the field $\vec{\varphi}(\lambda)$, thus for field (20.49). $\square$

Now we can describe all extremals of system (20.41) satisfying the conditions (20.42) and $h_{101} \neq 0$. Any such extremal belongs to the manifold $\{h_1 = h_{01} = 0\}$, and through any point λ_0 of this manifold with the boundary restrictions on control satisfied:

$$u(\lambda_0) = -\frac{h_{001}(\lambda_0)}{h_{101}(\lambda_0)} \in (\alpha, \beta),$$

passes a unique such extremal — the trajectory λ_t of system (20.49).

In problems considered in Chapters 13 and 18 (Dubins car, rotation around 2 axes in SO(3)), all singular extremals appeared exactly in this way. Generically, $h_{101} \neq 0$, thus all extremals with (20.42) can be studied as above. But in important examples h_{101} can vanish. E.g., consider a mechanical system with a controlled force:

$$\ddot{y} = g(y) + ub, \qquad y,\ b \in \mathbb{R}^n, \quad u \in [\alpha, \beta] \subset \mathbb{R},$$

or, in the standard form:

$$\begin{cases} \dot{y}_1 = y_2, \\ \dot{y}_2 = g(y_1) + ub. \end{cases}$$

The vector fields in the right-hand side are

$$\begin{aligned} f_0 &= y_2 \frac{\partial}{\partial y_1} + g(y_1) \frac{\partial}{\partial y_2}, \\ f_1 &= b \frac{\partial}{\partial y_2}, \end{aligned}$$

thus

$$h_{101}(\lambda) = \langle \lambda, \underbrace{[f_1, [f_0, f_1]]}_{\equiv 0} \rangle \equiv 0.$$

More generally, h_{101} vanishes as well for systems of the form

$$\begin{cases} \dot{x} = f(x, y), \\ \ddot{y} = g(x, y) + ub, \end{cases} \qquad x \in M, \quad y,\ b \in \mathbb{R}^n, \quad u \in [\alpha, \beta] \subset \mathbb{R}. \tag{20.51}$$

An interesting example of such systems is *Dubins car with control of angular acceleration*:

$$\begin{cases} \dot{x}_1 = \cos\theta, \\ \dot{x}_2 = \sin\theta, \\ \dot{\theta} = y, \\ \dot{y} = u, \end{cases} \qquad (x_1, x_2) \in \mathbb{R}^2, \quad \theta \in S^1, \quad y \in \mathbb{R}, \quad |u| \leq 1.$$

Having such a motivation in mind, we consider now the case where

$$h_{101}(\lambda) \equiv 0. \tag{20.52}$$

Then equality (20.47) does not contain $u(t)$, and we continue differentiation in order to find an equation determining the control:

$$h_1^{(3)}(\lambda_t) = \dot h_{001}(\lambda_t) = h_{0001}(\lambda_t) + u(t)h_{1001}(\lambda_t) \equiv 0.$$

It turns out that the term near $u(t)$ vanishes identically under condition (20.52):

$$h_{1001} = \{h_1, \{h_0, \{h_0, h_1\}\}\} = \underbrace{\{\{h_1, h_0\}, \{h_0, h_1\}\}}_{=0} + \{h_0, \{h_1, \{h_0, h_1\}\}\}$$
$$= \{h_0, h_{101}\} = 0.$$

So we obtain, in addition to (20.46), (20.47), and (20.52), one more identity without $u(t)$ for extremals:

$$h_{0001}(\lambda_t) \equiv 0.$$

Thus we continue differentiation:

$$h_1^{(4)}(\lambda_t) = \dot h_{0001}(\lambda_t) = h_{00001}(\lambda_t) + u(t)h_{10001}(\lambda_t) \equiv 0. \qquad (20.53)$$

In Dubins car with angular acceleration control $h_{10001}(\lambda_t) \neq 0$, and generically (in the class of systems (20.51)) this is also the case. Under the condition $h_{10001}(\lambda_t) \neq 0$ we can express control as $u = u(\lambda)$ from equation (20.53) and find all extremals in the same way as in the case $h_{101}(\lambda_t) \neq 0$.

Exercise 20.7. Show that for Dubins car with angular acceleration control, singular trajectories are straight lines in the plane (x_1, x_2):

$$x_1 = x_1^0 + t\cos\theta_0, \quad x_2 = x_2^0 + t\sin\theta_0, \quad \theta = \theta_0, \quad y = 0.$$

Although, now geometry of the system is new. There appears a new pattern for optimal control, where control has an infinite number of switchings on compact time intervals.

For the standard Dubins car (with angular velocity control) singular trajectories can join bang trajectories as follows:

$$u(t) = \pm 1, \quad t < \bar t; \qquad u(t) = 0, \quad t > \bar t, \qquad (20.54)$$

or

$$u(t) = 0, \quad t < \bar t; \qquad u(t) = \pm 1, \quad t > \bar t. \qquad (20.55)$$

We show that now such controls cannot be optimal.

The following argument, which proves Proposition 20.5, shows how our methods can be applied to problems not covered directly by the formal theory.

Consider the time-optimal problem for our single-input system (20.41). We prove that there do not exist time-optimal trajectories containing a singular piece followed by a bang piece. Suppose, by contradiction, that such a trajectory

$q(t)$ exists. Consider restriction of this trajectory to the singular and bang pieces:

$$
\begin{aligned}
&q(t), \qquad t \in [0, t_1], \\
&u(t) \in (\alpha, \beta), \qquad t \in [0, \bar{t}], \\
&u(t) = \gamma \in \{\alpha, \beta\}, \qquad t \in [\bar{t}, t_1].
\end{aligned}
$$

Let λ_t be an extremal corresponding to the extremal trajectory $q(t)$. We suppose that such λ_t is unique up to a nonzero factor (generically, this is the case). Reparametrizing control (i.e., taking $u - u(\bar{t} - 0)$ as a new control), we obtain

$$u(\bar{t} - 0) = 0, \qquad \alpha < 0 < \beta,$$

without any change of the structure of Lie brackets. Notice that now we study a time-optimal trajectory, not geometrically optimal one as before. Although, the Hamiltonian of PMP $h_u = h_0 + u h_1$ for the time-optimal problem is the same as for the geometric problem, thus the above analysis of singular extremals applies. In fact, we prove below that a singular piece and a bang piece cannot follow one another not only for a time-minimal trajectory, but also for a time-maximal trajectory or for a geometrically optimal one.

We suppose that the fields f_0, f_1 satisfy the identity

$$[f_1, [f_0, f_1]] \equiv 0$$

and the extremal λ_t satisfies the inequality

$$h_{10001}(\lambda_{\bar{t}}) \neq 0.$$

Since $u(\bar{t} - 0) = 0$, then equality (20.53) implies that $h_{00001}(\lambda_{\bar{t}}) = 0$.

It follows from the maximality condition of PMP that

$$h_{u(t)}(\lambda_t) = h_0(\lambda_t) + u(t) h_1(\lambda_t) \geq h_0(\lambda_t),$$

i.e., along the whole extremal

$$u(t) h_1(\lambda_t) \geq 0, \qquad t \in [0, t_1].$$

But along the singular piece $h_1(\lambda_t) \equiv 0$, thus

$$u(t) h_1(\lambda_t) \equiv 0, \qquad t \in [0, \bar{t}].$$

The first nonvanishing derivative of $u_1(t) h_1(\lambda_t)$ at $t = \bar{t} + 0$ is positive. Keeping in mind that $u(t) \equiv \gamma$ at the singular piece $t \in [\bar{t}, t_1]$, we compute this derivative. Since $h_1(\lambda_{\bar{t}}) = h_{01}(\lambda_{\bar{t}}) = h_{001}(\lambda_{\bar{t}}) = h_{0001}(\lambda_{\bar{t}}) = h_{1001}(\lambda_{\bar{t}}) = 0$, then the first three derivatives vanish:

$$\left.\frac{d^k}{dt^k}\right|_{t = \bar{t} + 0} u(t) h_1(\lambda_t) = 0, \qquad k = 0, 1, 2, 3.$$

Thus the fourth derivative is nonnegative:

$$\left.\frac{d^4}{dt^4}\right|_{t=\bar{t}+0} u(t)h_1(\lambda_t) = \gamma(h_{00001}(\lambda_{\bar{t}}) + \gamma h_{10001}(\lambda_{\bar{t}})) \\ = \gamma^2 h_{10001}(\lambda_{\bar{t}}) \geq 0.$$

Since $\gamma^2 > 0$, then

$$h_{10001}(\lambda_{\bar{t}}) > 0. \tag{20.56}$$

Now we apply this inequality in order to obtain a contradiction via the theory of second variation.

Recall expression (20.29) for Hessian of the endpoint mapping:

$$\lambda_t \operatorname{Hess}_u F_t(v) \\ = \int_0^t \lambda_0 \, [\dot{g}'_\tau, g'_\tau] \, w^2(\tau) \, d\tau + \int_0^t \int_0^{\tau_1} \lambda_0 \, [\dot{g}'_{\tau_2}, \dot{g}'_{\tau_1}] \, w(\tau_2) w(\tau_1) \, d\tau_2 d\tau_1. \tag{20.57}$$

Here

$$w(\tau) = \int_\tau^t v(\theta) \, d\theta, \qquad w(0) = 0,$$
$$g'_\tau = P_{\tau *}^{-1} f_1,$$
$$\dot{g}'_\tau = P_{\tau *}^{-1} [f_0, f_1],$$
$$P_\tau = \overrightarrow{\exp} \int_0^\tau f_{u(\theta)} \, d\theta.$$

The first term in expression (20.57) for the Hessian vanishes:

$$\lambda_0 \, [\dot{g}'_\tau, g'_\tau] = -h_{101}(\lambda_\tau) \equiv 0.$$

Integrating the second term by parts twice, we obtain:

$$\lambda_t \operatorname{Hess}_u F_t(v) = \int_0^t \lambda_0 \, [\ddot{g}'_\tau, \dot{g}'_\tau] \, \eta^2(\tau) \, d\tau + \int_0^t \int_0^{\tau_1} \lambda_0 \, [\ddot{g}'_{\tau_2}, \ddot{g}'_{\tau_1}] \, \eta(\tau_2)\eta(\tau_1) \, d\tau_2 \, d\tau_1 \tag{20.58}$$

where

$$\ddot{g}'_\tau = P_{\tau *}^{-1} [f_0, [f_0, f_1]],$$
$$\eta(\tau) = \int_0^\tau w(\tau_1) \, d\tau_1, \qquad \eta(t) = 0.$$

The first term in (20.58) dominates on needle-like variations $v = v_{\bar{t},\varepsilon}$:

$$\lambda_{\bar{t}} \operatorname{Hess}_u F_{\bar{t}}(v_{\bar{t},\varepsilon}) = \varepsilon^4 \lambda_0 [\, \ddot{g}'_{\bar{t}}, \, \dot{g}'_{\bar{t}}] + O(\varepsilon^5),$$

we compute the leading term in the Hamiltonian form:

$$\begin{aligned}\lambda_0[\ddot{g}'_{\bar t}, \dot{g}'_{\bar t}] &= \lambda_{\bar t}[[f_0,[f_0,f_1]],[f_0,f_1]] = \{h_{001},h_{01}\}(\lambda_{\bar t}) = \{\{h_1,h_0\},h_{001}\}(\lambda_{\bar t}) \\ &= \{h_1,\{h_0,h_{001}\}\}(\lambda_{\bar t}) - \{h_0,\underbrace{\{h_1,h_{001}\}}_{=h_{1001}\equiv 0}\}(\lambda_{\bar t}) = h_{10001}(\lambda_{\bar t}).\end{aligned}$$

By virtue of inequality (20.56),

$$\lambda_{\bar t}\,\mathrm{Hess}_u\, F_{\bar t}(v_0) > 0,$$

where

$$v_0 = v_{\bar t,\varepsilon}$$

for small enough $\varepsilon > 0$. This means that

$$\left.\frac{d^2}{d\,s^2}\right|_{s=0} a \circ F_{\bar t}(u + s v_0) = \lambda_{\bar t}\,\mathrm{Hess}_u\, F_{\bar t}(v_0) > 0$$

for any function $a \in C^\infty(M)$, $a(q(\bar t)) = 0$, $d_{q(\bar t)} a = \lambda_{\bar t}$. Then

$$a \circ F_{\bar t}(u + s v_0) = \frac{s^2}{2}\lambda_{\bar t}\,\mathrm{Hess}_u\, F_{\bar t}(v_0) + O(s^3), \qquad s \to 0,$$

i.e., the curve $F_{\bar t}(u + \sqrt{s} v_0)$ is smooth at $s = +0$ and has the tangent vector

$$\begin{aligned}&\left.\frac{d}{d\,s}\right|_{s=+0} F_{\bar t}(u + \sqrt{s} v_0) = \xi_0, \\ &\langle \lambda_{\bar t}, \xi_0 \rangle > 0. \end{aligned} \tag{20.59}$$

That is, variation of the optimal control u in direction of v_0 generates a tangent vector ξ_0 to the attainable set $\mathcal{A}_{q_0}(\bar t)$ that belongs to the half-space $\langle \lambda_{\bar t}, \cdot\, \rangle > 0$ in $T_{q(\bar t)}M$.

Since the extremal trajectory $q(t)$ is a projection of a unique, up to a scalar factor, extremal λ_t, then the control u is a corank one critical point of the endpoint mapping:

$$\dim \mathrm{Im}\, D_u F_{\bar t} = \dim M - 1 = n - 1.$$

This means that there exist variations of control that generate a hyperplane of tangent vectors to $\mathcal{A}_{q_0}(\bar t)$:

$$\begin{aligned}&\exists\ v_1, \dots, v_{n-1} \in T_u\mathcal{U} \quad \text{such that} \\ &\left.\frac{d}{d\,s}\right|_{s=0} F_{\bar t}(u + s v_i) = \xi_i, \qquad i = 1, \dots, n-1, \\ &\mathrm{span}(\xi_1, \dots, \xi_{n-1}) = \mathrm{Im}\, D_u F_{\bar t}.\end{aligned}$$

Summing up, the variations $v_0, v_1, \dots, v_{n-1}$ of the control u at the singular piece generate a nonnegative half-space of the covector $\lambda_{\bar{t}}$:

$$u_s = u + \sqrt{s_0}v_0 + \sum_{i=1}^{n-1} s_i v_i, \qquad s = (s_0, s_1, \dots, s_{n-1}) \in \mathbb{R}_+ \times \mathbb{R}^{n-1},$$

$$\left.\frac{\partial}{\partial s_i}\right|_{s=0} F_{\bar{t}}(u_s) = \xi_i, \qquad i = 0, 1, \dots, n-1,$$

$$\mathbb{R}_+\xi_0 + \operatorname{span}(\xi_1, \dots, \xi_{n-1}) = \{\langle \lambda_{\bar{t}}, \cdot \rangle \geq 0\}.$$

Now we add a needle-like variation on the bang piece. Since the control $u(t)$, $t \in [\bar{t}, t_1]$, is nonsingular, then the switching function $h_1(\lambda_t) \not\equiv 0$, $t \in [\bar{t}, t_1]$. Choose any instant

$$\bar{t}_1 \in (\bar{t}, t_1) \text{ such that } h_1(\lambda_{\bar{t}_1}) \neq 0.$$

Add a needle-like variation concentrated at small segments near $\bar{t}_1$:

$$u_{s,\varepsilon}(t) = \begin{cases} u_s(t), & t \in [0, \bar{t}], \\ u(t) = \gamma, & t \in [\bar{t}, \bar{t}_1] \cup [\bar{t}_1 + \varepsilon, t_1], \\ 0, & t \in [\bar{t}_1, \bar{t}_1 + \varepsilon]. \end{cases}$$

The needle-like variation generates the tangent vector

$$\left.\frac{\partial}{\partial \varepsilon}\right|_{(\varepsilon,s)=(+0,0)} F_{t_1}(u_{s,\varepsilon}) = -\gamma \left[\left(P_{\bar{t}_1}^{t_1}\right)_* f_1 \right] (q(t_1)),$$

$$P_\tau^t = \overrightarrow{\exp} \int_\tau^t f_{u(\tau)}\, d\tau,$$

this derivative is computed as in the proof of PMP, see Lemma 12.2. We determine disposition of the vector

$$\eta_n = -\gamma \left[\left(P_{\bar{t}_1}^{t_1}\right)_* f_1 \right] (q(t_1))$$

w.r.t. the hyperplane $\operatorname{Im} D_u F_{t_1}$:

$$\langle \lambda_{t_1}, \eta_n \rangle = -\gamma \langle \lambda_{\bar{t}_1}, f_1 \rangle = -\gamma h_1(\lambda_{\bar{t}_1}).$$

Since $h_1(\lambda_{\bar{t}_1}) \neq 0$, then it follows from PMP that $\gamma h_1(\lambda_{\bar{t}_1}) = u(\bar{t}_1) h_1(\lambda_{\bar{t}_1}) > 0$, thus

$$\langle \lambda_{t_1}, \eta_n \rangle < 0.$$

Now we translate the tangent vectors ξ_i, $i = 0, \dots, n-1$, from $q(\bar{t})$ to $q(t_1)$:

$$\left.\frac{\partial}{\partial \varepsilon}\right|_{(\varepsilon,s)=(0,0)} F_{t_1}(u_{s,\varepsilon}) = \left.\frac{\partial}{\partial \varepsilon}\right|_{(\varepsilon,s)=(0,0)} P_{\bar{t}}^{t_1}\left(F_{\bar{t}}(u_s)\right)$$

$$= \left(P_{\bar{t}}^{t_1}\right)_* \xi_i = \eta_i, \quad i = 0, \dots, n-1.$$

Inequality (20.59) translates to

$$\langle \lambda_{t_1}, \eta_0 \rangle = \langle \lambda_{\bar{t}}, \xi_0 \rangle > 0$$

and, of course,

$$\langle \lambda_{t_1}, \eta_i \rangle = \langle \lambda_{\bar{t}}, \xi_i \rangle = 0, \qquad i = 1, \dots, n-1.$$

The inequality $\langle \lambda_{t_1}, \eta_n \rangle < 0$ means that the needle-like variation on the bang piece generates a tangent vector in the half-space $\langle \lambda_{t_1}, \cdot \rangle < 0$ complementary to the half-space $\langle \lambda_{t_1}, \cdot \rangle \geq 0$ generated by variations on the singular piece.

Summing up, the mapping

$$F : \mathbb{R}_+ \times \mathbb{R}^{n-1} \times \mathbb{R}_+ \to M, \qquad F(s, \varepsilon) = F_{t_1}(u_{s,\varepsilon}),$$

satisfies the condition

$$D_{(0,0)} F(\mathbb{R}_+ \times \mathbb{R}^{n-1} \times \mathbb{R}_+) = \mathbb{R}_+ \eta_0 + \operatorname{span}(\eta_1, \dots, \eta_{n-1}) + \mathbb{R}_+ \eta_n = T_{q(t_1)} M.$$

By Lemma 12.1 and remark after it, the mapping F is locally open at $(s, \varepsilon) = (0, 0)$. Thus the image of the mapping $F_{t_1}(u_{s,\varepsilon})$ contains a neighborhood of the terminal point $q(t_1)$. By continuity, $q(t_1)$ remains in the image of $F_{t_1 - \delta}(u_{s,\varepsilon})$ for sufficiently small $\delta > 0$. In other words, the point $q(t_1)$ is reachable from q_0 at $t_1 - \delta$ instants of time, i.e., the trajectory $q(t)$, $t \in [0, t_1]$, is not time-optimal, a contradiction.

We proved that a time-optimal trajectory $q(t)$ cannot have a singular piece followed by a bang piece. Similarly, a singular piece cannot follow a bang piece.

We obtain the following statement on the possible structure of optimal control.

Proposition 20.5. *Assume that vector fields in the right-hand side of system* (20.41) *satisfy the identity*

$$[f_1, [f_0, f_1]] = 0. \tag{20.60}$$

Let a time-optimal trajectory $q(t)$ of this system be a projection of a unique, up to a scalar factor, extremal λ_t, and let $h_{10001}(\lambda_t) \neq 0$. Then the trajectory $q(t)$ cannot contain a singular piece and a bang piece adjacent one to another.

Remark. In this proposition, time-optimal (i.e., time-minimal) control can be replaced by a time-maximal control or by a geometrically optimal one.

What happens near singular trajectories under hypothesis (20.60)? Assume that a singular trajectory is optimal (as straight lines for Dubins car with angular acceleration control). Notice that optimal controls exist, thus the cost function is everywhere defined. For boundary conditions sufficiently close to the singular trajectory, there are two possible patterns of optimal control:

(1) either it makes infinite number of switchings on a compact time segment adjacent to the singular part, so that the optimal trajectory "gets off" the singular trajectory via infinite number of switchings,

(2) or optimal control is bang-bang, but the number of switchings grows infinitely as the terminal point approaches the singular trajectory.

Pattern (1) of optimal control is called *Fuller's phenomenon*. It turns out that Fuller's phenomenon takes place in Dubins car with angular acceleration control. As our preceding arguments suggest, this phenomenon is not a pathology, but is ubiquitous for certain classes of systems (in particular, in applications). One can observe this phenomenon trying to stop a ping-pong ball jumping between the table and descending racket. The theory of Fuller's phenomenon is described in book [26]. It follows from this theory that possibility (1) is actually realized for Dubins car with angular acceleration control.

Chapter 21

Jacobi equation

In Chapter 20 we established that the sign of the quadratic form $\lambda_t \operatorname{Hess}_{\tilde{u}} F_t$ is related to optimality of the reference control $\tilde{u}$. Under natural assumptions, the second variation is negative on short segments. Now we wish to catch the instant of time where this quadratic form fails to be negative. We derive an ODE (Jacobi equation) that allows to find such instants (conjugate times). Moreover, we give necessary and sufficient optimality conditions in these terms.

Recall expression (20.18) for the quadratic form Q, $\lambda_t \operatorname{Hess}_{\tilde{u}} F_t = Q|_{\operatorname{Ker} D_{\tilde{u}} F_t}$, obtained in Section 20.3:

$$Q(v) = \int_0^t h''_\tau(v(\tau))\, d\tau + \int_0^t \lambda_0 \left[\int_0^{\tau_1} g'_{\tau_2} v(\tau_2)\, d\tau_2,\, g'_{\tau_1} v(\tau_1) \right] d\tau_1.$$

We extend the form Q from L_∞ to L_2 by continuity.

We will consider a family of problems on segments $[0,t]$, $t \in [0,t_1]$, so we introduce the corresponding sets of admissible controls:

$$\mathcal{U}_t = \{u \in L_2([0,t_1], U) \mid u(\tau) = 0 \text{ for } \tau > t\},$$

and spaces of variations of controls:

$$\mathcal{V}_t = T_{\tilde{u}} \mathcal{U}_t = \{v \in L_2^m[0,t_1] \mid v(\tau) = 0 \text{ for } \tau > t\} \cong L_2^m[0,t].$$

We denote the second variation on the corresponding segment as

$$Q_t = Q|_{\mathcal{V}_t}.$$

Notice that the family of spaces $\mathcal{V}_t$ is ordered by inclusion:

$$t' < t'' \quad \Rightarrow \quad \mathcal{V}_{t'} \subset \mathcal{V}_{t''},$$

and the family of forms Q_t respects this order:

$$Q_{t'} = Q_{t''}|_{\mathcal{V}_{t'}}.$$

In particular,

$$Q_{t''} < 0 \qquad \Rightarrow \qquad Q_{t'} < 0.$$

Denote the instant of time where the forms Q_t lose their negative sign:

$$t_* \stackrel{\text{def}}{=} \sup \left\{ t \in (0, t_1] \mid Q_t|_{K_t} < 0 \right\},$$

where

$$K_t = \left\{ v \in \mathcal{V}_t \mid q_0 \circ \int_0^t g'_\tau v(\tau)\, d\tau = 0 \right\}$$

is the closure of the space $\operatorname{Ker} D_{\tilde{u}} F_t$ in L_2. If $Q_t|_{K_t}$ is negative for all $t \in (0, t_1]$, then, by definition, $t_* = +\infty$.

21.1 Regular case: derivation of Jacobi equation

Proposition 21.1. *Let λ_t be a regular extremal with $t_* \in (0, t_1]$. Then the quadratic form $Q_{t_*}|_{K_{t_*}}$ is degenerate.*

Proof. By the strong Legendre condition, the norm

$$\|v\|_{h''} = \left(\int_0^{t_*} -h''_\tau(v(\tau))\, d\tau \right)^{1/2}$$

is equivalent to the standard L_2^m-norm. Then

$$\begin{aligned} Q_{t_*} &= \int_0^{t_*} h''_\tau(v(\tau))\, d\tau + \int_0^{t_*} \lambda_0 \left[\int_0^{\tau_1} g'_{\tau_2} v(\tau_2)\, d\tau_2,\, g'_{\tau_1} v(\tau_1) \right] d\tau_1 \\ &= -\|v\|^2_{h''} + \langle Rv, v \rangle, \end{aligned}$$

where R is a compact operator in $L_2^m[0, t_*]$.

First we prove that the quadratic form Q_{t_*} is nonpositive on the kernel K_{t_*}. Assume, by contradiction, that there exists $v \in \mathcal{V}_{t_*}$ such that

$$Q_{t_*}(v) > 0, \qquad v \in K_{t_*}.$$

The linear mapping $D_{\tilde{u}} F_{t_*}$ has a finite-dimensional image, thus

$$\mathcal{V}_{t_*} = K_{t_*} \oplus E, \qquad \dim E < \infty.$$

The family $D_{\tilde{u}} F_t$ is weakly continuous in t, hence $D_{\tilde{u}} F_{t_* - \varepsilon}|_E$ is invertible and

$$\mathcal{V}_{t_*} = K_{t_* - \varepsilon} \oplus E$$

for small $\varepsilon > 0$. Consider the corresponding decomposition

$$v = v_\varepsilon + x_\varepsilon, \qquad v_\varepsilon \in K_{t_* - \varepsilon}, \quad x_\varepsilon \in E.$$

Then $x_\varepsilon \to 0$ weakly as $\varepsilon \to 0$, so $x_\varepsilon \to 0$ strongly since E is finite-dimensional. Consequently, $v_\varepsilon \to v$ strongly as $\varepsilon \to 0$. Further, $Q_{t_* - \varepsilon}(v_\varepsilon) = Q_{t_*}(v_\varepsilon) \to$

$Q_{t_*}(v)$ as $\varepsilon \to 0$ since the quadratic forms Q_t are continuous. Summing up, $Q_{t_*-\varepsilon}(v_\varepsilon) > 0$ for small $\varepsilon > 0$, a contradiction with definition of t_*. We proved that

$$Q_{t_*}|_{K_{t_*}} \leq 0. \tag{21.1}$$

Now we show that

$$\exists\, v \in K_{t_*}, \quad v \neq 0, \qquad \text{such that} \qquad Q_{t_*}(v) = 0.$$

By the argument similar to the proof of Proposition 16.2 (in the study of conjugate points for the linear-quadratic problem), we show that the function

$$\mu(t) = \sup\ \{Q_t(v) \mid v \in K_t,\ \|v\|_{h''} = 1\} \tag{21.2}$$

satisfies the properties: $\mu(t)$ is monotone nondecreasing, the supremum in (21.2) is attained, and $\mu(t)$ is continuous from the right.

Inequality (21.1) means that $\mu(t_*) \leq 0$. If $\mu(t_*) < 0$, then $\mu(t_* + \varepsilon) < 0$ for small $\varepsilon > 0$, which contradicts definition of the instant t_*. Thus $\mu(t_*) = 0$, moreover, there exists

$$v \in K_{t_*}, \qquad \|v\|_{h''} = 1,$$

such that

$$Q_{t_*}(v) = 0.$$

Taking into account that the quadratic form Q_{t_*} is nonpositive, we conclude that the element $v \neq 0$ is in the kernel of $Q_{t_*}|_{K_{t_*}}$. □

Proposition 21.1 motivates the introduction of the following important notion. An instant $t_c \in (0, t_1]$ is called a *conjugate time* (for the initial instant $t = 0$) along a regular extremal λ_t if the quadratic form $Q_{t_c}|_{K_{t_c}}$ is degenerate. Notice that by Proposition 20.2, the forms $Q_t|_{K_t}$ are negative for small $t > 0$, thus short arcs of regular extremals have no conjugate points: for them $t_* > 0$. Proposition 21.1 means that the instant t_* where the quadratic forms $Q_t|_{K_t}$ lose their negative sign is the first conjugate time.

We start to derive a differential equation on conjugate time for a regular extremal pair $(\tilde{u}(t), \lambda_t)$. The symplectic space

$$\Sigma = T_{\lambda_0}(T^*M)$$

will be the state space of that ODE. Introduce the family of mappings

$$\begin{aligned} J_t &: \mathbb{R}^m \to \Sigma, \\ J_t\, v &= \left.\frac{\partial}{\partial u}\right|_{\tilde{u}(t)} \overrightarrow{h_{u,t}}\ v. \end{aligned}$$

In these terms, the bilinear form Q_t reads

$$Q_t(v_1, v_2) = \int_0^t h''_\tau(v_1(\tau), v_2(\tau))\, dt + \iint\limits_{0 \leq \tau_2 \leq \tau_1 \leq t} \sigma(J_{\tau_2} v_1(\tau_2), J_{\tau_1} v_2(\tau_1))\, d\tau_1 d\tau_2, \tag{21.3}$$

see (20.18), (20.19). We consider the form Q_t on the subspace

$$K_t = \operatorname{Ker} D_{\tilde{u}} F_t = \left\{ v_i \in \mathcal{V}_t \mid \int_0^t J_\tau v_i(\tau)\, d\tau \in \Pi_0 \right\}, \tag{21.4}$$

where

$$\Pi_0 = T_{\lambda_0}(T^*_{q_0} M) \subset \Sigma$$

is the vertical subspace.

A variation of control $v \in \mathcal{V}_t$ satisfies the inclusion

$$v \in \operatorname{Ker} \left(Q_t|_{K_t} \right)$$

iff the linear form $Q_t(v, \cdot)$ annihilates the subspace $K_t \subset \mathcal{V}_t$. Since the vertical subspace $\Pi_0 \subset \Sigma$ is Lagrangian, equality (21.4) can be rewritten as follows:

$$K_t = \left\{ v_i \in \mathcal{V}_t \mid \sigma \left(\int_0^t J_\tau v_i(\tau)\, d\tau, \nu \right) = 0 \quad \forall \nu \in \Pi_0 \right\}.$$

That is, the annihilator of the subspace $K_t \subset \mathcal{V}_t$ coincides with the following finite-dimensional space of linear forms on $\mathcal{V}_t$:

$$\left\{ \int_0^t \sigma(J_\tau \cdot, \nu)\, d\tau \mid \nu \in \Pi_0 \right\}. \tag{21.5}$$

Summing up, we obtain that $v \in \operatorname{Ker} \left(Q_t|_{K_t} \right)$ iff the form $Q_t(v, \cdot)$ on $\mathcal{V}_t$ belongs to subspace (21.5). That is, $v \in \operatorname{Ker} \left(Q_t|_{K_t} \right)$ iff there exists $\nu \in \Pi_0$ such that

$$Q_t(v, \cdot) = \int_0^t \sigma(J_\tau \cdot, \nu)\, d\tau. \tag{21.6}$$

We transform equality of forms (21.6):

$$\begin{aligned} \int_0^t \sigma(J_\tau \cdot, \nu)\, d\tau &= \int_0^t h''_\tau(v(\tau), \cdot)\, d\tau + \iint\limits_{0 \le \tau_2 \le \tau_1 \le t} \sigma(J_{\tau_2} v(\tau_2), J_{\tau_1} \cdot)\, d\tau_1 d\tau_2 \\ &= \int_0^t h''_\tau(v(\tau), \cdot)\, d\tau + \int_0^t \sigma \left(\int_0^\tau J_\theta v(\theta)\, d\theta, J_\tau \cdot \right) d\tau. \end{aligned}$$

This equality of forms means that the integrands coincide one with another:

$$\sigma(J_\tau \cdot, \nu) = h''_\tau(v(\tau), \cdot) + \sigma \left(\int_0^\tau J_\theta v(\theta)\, d\theta, J_\tau \cdot \right), \qquad \tau \in [0, t]. \tag{21.7}$$

In terms of the curve in the space Σ

$$\eta_\tau = \int_0^\tau J_\theta v(\theta)\, d\theta + \nu, \qquad \tau \in [0, t], \tag{21.8}$$

equality of forms (21.7) can be rewritten as follows:

$$h''_\tau(v(\tau),\,\cdot\,) + \sigma(\eta_\tau, J_\tau\,\cdot\,) = 0, \qquad \tau \in [0,t]. \tag{21.9}$$

The strong Legendre condition implies that the linear mapping

$$h''_\tau : \mathbb{R}^m \to \mathbb{R}^{m*}$$

is nondegenerate (we denote here and below the linear mapping into the dual space by the same symbol as the corresponding quadratic form), thus the inverse mapping is defined:

$$(h''_\tau)^{-1} : \mathbb{R}^{m*} \to \mathbb{R}^m.$$

Then equality (21.9) reads

$$v(\tau) + (h''_\tau)^{-1}\,\sigma(\eta_\tau, J_\tau\,\cdot\,) = 0, \qquad \tau \in [0,t]. \tag{21.10}$$

We come to the following statement.

Theorem 21.1. *Let λ_t, $t \in [0,t_1]$, be a regular extremal. An instant $t \in (0,t_1]$ is a conjugate time iff there exists a nonconstant solution η_τ to* Jacobi equation

$$\dot\eta_\tau = J_\tau\,(h''_\tau)^{-1}\,\sigma(J_\tau\,\cdot\,,\eta_\tau), \qquad \tau \in [0,t], \tag{21.11}$$

that satisfies the boundary conditions

$$\eta_0 \in \Pi_0, \qquad \eta_t \in \Pi_0. \tag{21.12}$$

Jacobi equation (21.11) *is a linear nonautonomous Hamiltonian system on Σ:*

$$\dot\eta_\tau = \vec{b}_\tau(\eta_\tau) \tag{21.13}$$

with the quadratic Hamiltonian function

$$b_\tau(\eta) = -\frac{1}{2}\,(h''_\tau)^{-1}\left(\sigma(J_\tau\,\cdot\,,\eta), \sigma(J_\tau\,\cdot\,,\eta)\right), \qquad \eta \in \Sigma,$$

where $(h''_\tau)^{-1}$ is a quadratic form on $\mathbb{R}^{m*}$.

Proof. We already proved that existence of $v \in \operatorname{Ker} Q_t|_{K_t}$ is equivalent to existence of a solution η_τ to Jacobi equation that satisfies the boundary conditions (21.12).

If $v \equiv 0$, then $\eta_\tau \equiv \text{const}$ by virtue of (21.8). Conversely, if $\eta_\tau \equiv \text{const}$, then $J_\tau v(\tau) = \dot\eta_\tau \equiv 0$. By (21.3), the second variation takes the form

$$Q_t(v) = \int_0^t h''_\tau(v(\tau))\,d\tau < -\alpha\|v\|^2_{L_2} \qquad \text{for some } \alpha > 0.$$

But $v \in \operatorname{Ker} Q_t$, so $Q_t(v) = 0$, consequently $v \equiv 0$. Thus nonzero v correspond to nonconstant η_τ and vice versa.

It remains to prove that b_τ is the Hamiltonian function for Jacobi equation (21.11). Denote

$$A_\tau(\eta) = (h''_\tau)^{-1}\,\sigma(J_\tau\,\cdot\,,\eta) \in \mathbb{R}^m, \qquad \eta \in \Sigma,$$

then Jacobi equation reads

$$\dot\eta_\tau = J_\tau A_\tau(\eta_\tau),$$

so we have to prove that

$$J_\tau A_\tau(\eta) = \vec{b}_\tau(\eta), \qquad \eta \in \Sigma. \tag{21.14}$$

Since

$$b_\tau(\eta) = -\frac{1}{2}\left\langle \sigma(J_\tau\,\cdot\,,\eta), (h''_\tau)^{-1}\,\sigma(J_\tau\,\cdot\,,\eta)\right\rangle = -\frac{1}{2}\left\langle \sigma(J_\tau\,\cdot\,,\eta), A_\tau(\eta)\right\rangle,$$

then

$$\langle d_\eta b_\tau, \xi\rangle = -\langle \sigma(J_\tau\,\cdot\,,\xi), A_\tau(\eta)\rangle = \sigma(\xi, J_\tau A_\tau(\eta)).$$

Thus equality (21.14) follows and the proof is complete. □

21.2 Singular case: derivation of Jacobi equation

In this section we obtain Jacobi equation for a nice singular extremal pair $(\tilde{u}(t), \lambda_t)$.

In contrast to the regular case, the second variation in the singular case can be nondegenerate at the instant t_* where it loses its negative sign. In order to develop the theory of conjugate points for the singular case, we introduce a change of variables in the form Q_t. We denote, as before, the integrals

$$w_i(\tau) = \int_\tau^t v_i(s)\,ds, \qquad i = 1,\ 2,$$

and denote the bilinear form that enters generalized Legendre condition:

$$l_t(w_1, w_2) = \sigma(\dot J_t w_1, J_t w_2), \qquad w_i \in \mathbb{R}^m.$$

For a nice singular extremal, expression (20.25) for the second variation reads

$$Q_t(v_1, v_2) = \int_0^t l_\tau(w_1(\tau), w_2(\tau))\,d\tau + \int_0^t \sigma\left(\dot J_\tau w_1(\tau), \int_\tau^t \dot J_\theta w_2(\theta)\,d\theta\right)\,d\tau$$
$$+\,\sigma\left(J_0 w_1(0), \int_0^t \dot J_\tau w_2(\tau)\,d\tau\right).$$

Admissibility condition (20.26) for variations of control $v_i(\cdot)$ can be written as follows:

$$\int_0^t \dot J_\tau w(\tau)\,d\tau + J_0 w(0) \in \Pi_0. \tag{21.15}$$

The mapping
$$v(\cdot) \mapsto (w(\cdot), w(0)) \in L_2^m \times \mathbb{R}^m$$
has a dense image in $L_2^m \times \mathbb{R}^m$, and the Hessian Q_t and admissibility condition (21.15) are extended to $L_2^m \times \mathbb{R}^m$ by continuity.

Denote
$$\gamma = J_0 w(0) \in \Gamma_0$$
and consider the extended form
$$Q_t(w_1, w_2) = \int_0^t l_\tau(w_1(\tau), w_2(\tau))\, d\tau + \int_0^t \sigma\left(\dot{J}_\tau w_1(\tau), \int_\tau^t \dot{J}_\theta w_2(\theta)\, d\theta\right) d\tau + \sigma\left(\gamma_1, \int_0^t \dot{J}_\tau w_2(\tau)\, d\tau\right)$$
on the space
$$\int_0^t \dot{J}_\tau w(\tau)\, d\tau + \gamma \in \Pi_0. \tag{21.16}$$

Then in the same way as in the regular case, it follows that the restriction of the quadratic form $Q_t(w)$ to the space (21.16) is degenerate at the instant $t = t_*$. An instant t that satisfies such a property is called a *conjugate time*.

Similarly to the regular case, we derive now a Hamiltonian Jacobi equation on conjugate times for nice singular extremals, although the Hamiltonian function and boundary conditions differ from the ones for the regular case.

Let $t \in (0, t_1]$ be a conjugate time, i.e., let the form $Q_t(w_1, w_2)$ have a nontrivial kernel on the space (21.16). That is, there exists a pair
$$(w, \gamma) \in L_2^m[0,t] \times \Gamma_0, \qquad \int_0^t \dot{J}_\tau w(\tau)\, d\tau + \gamma \in \Pi_0,$$
such that the linear form on the space $L_2^m[0,t] \times \Gamma_0$
$$Q_t(\,\cdot\,, w) = \int_0^t l_\tau(\,\cdot_{L_2}, w(\tau))\, d\tau + \int_0^t \sigma\left(\dot{J}_\tau \,\cdot_{L_2}, \int_\tau^t \dot{J}_\theta w(\theta)\, d\theta\right) d\tau + \sigma\left(\,\cdot_{\Gamma_0}, \int_0^t \dot{J}_\tau w(\tau)\, d\tau\right) \tag{21.17}$$
annihilates the admissible space (21.16). In turn, the annihilator of the admissible space (21.16) is the space of linear forms
$$\int_0^t \sigma\left(\dot{J}_\tau \,\cdot_{L_2}, \nu\right) d\tau + \sigma\left(\,\cdot_{\Gamma_0}, \nu\right), \qquad \nu \in \Pi_0.$$
Thus, similarly to the regular case, there exists $\nu \in \Pi_0$ such that
$$Q_t(\,\cdot\,, w) = \int_0^t \sigma\left(\dot{J}_\tau \,\cdot_{L_2}, \nu\right) d\tau + \sigma\left(\,\cdot_{\Gamma_0}, \nu\right).$$

By virtue of (21.17), the previous equality of forms splits:

$$l_\tau\left(\cdot_{\mathbb{R}^m}, w(\tau)\right) + \sigma\left(\dot{J}_\tau \cdot_{\mathbb{R}^m}, \int_\tau^t \dot{J}_\theta w(\theta)\, d\theta\right) = \sigma\left(\dot{J}_\tau \cdot_{\mathbb{R}^m}, \nu\right), \qquad \tau \in [0,t],$$

$$\sigma\left(\cdot_{\Gamma_0}, \int_0^t \dot{J}_\tau w(\tau)\, d\tau\right) = \sigma\left(\cdot_{\Gamma_0}, \nu\right).$$

That is,

$$l_\tau w(\tau) = -\sigma\left(\dot{J}_\tau \cdot_{\mathbb{R}^m}, \int_\tau^t \dot{J}_\theta w(\theta)\, d\theta - \nu\right), \tag{21.18}$$

$$\sigma\left(\cdot_{\Gamma_0}, \int_0^t \dot{J}_\tau w(\tau)\, d\tau - \nu\right) = 0. \tag{21.19}$$

In terms of the curve in the space $\Sigma = T_{\lambda_0}(T^*M)$

$$\eta_\tau = \int_\tau^t \dot{J}_\theta w(\theta)\, d\theta - \nu, \qquad \tau \in [0,t], \tag{21.20}$$

equalities (21.18), (21.19) take the form

$$l_\tau w(\tau) = -\sigma\left(\dot{J}_\tau \cdot_{\mathbb{R}^m}, \eta_\tau\right), \qquad \tau \in [0,t], \tag{21.21}$$

$$\sigma\left(\Gamma_0, \eta_0\right) = 0.$$

The last equality means that η_0 belongs to the skew-orthogonal complement $\Gamma_0^\angle$. On the other hand, $\eta_0 \in \Gamma_0 + \Pi_0$, compare definition (21.20) with (21.16). That is,

$$\eta_0 \in (\Pi_0 + \Gamma_0) \cap \Gamma_0^\angle = \Pi_0^{\Gamma_0}.$$

Recall that $\Pi_0^{\Gamma_0}$ is a Lagrangian subspace in the symplectic space Σ containing the isotropic subspace Γ_0, see definition (11.28). Notice that Goh condition

$$\sigma(J_t v_1, J_t v_2) \equiv 0, \qquad v_1,\, v_2 \in \mathbb{R}^m, \quad t \in [0, t_1]$$

means that the subspaces

$$\Gamma_t = \operatorname{span}\{J_t v \mid v \in \mathbb{R}^m\} \subset \Sigma$$

are isotropic. We obtain boundary conditions for the curve η_τ:

$$\eta_0 \in \Pi_0^{\Gamma_0}, \qquad \eta_t \in \Pi_0. \tag{21.22}$$

Moreover, equality (21.21) yields an ODE for η_τ:

$$\dot{\eta}_\tau = -\dot{J}_\tau w(\tau) = \dot{J}_\tau\, l_\tau^{-1}(\sigma(\dot{J}_\tau \cdot, \eta_\tau)), \qquad \tau \in [0,t]. \tag{21.23}$$

Similarly to the regular case, it follows that this equation is Hamiltonian with the Hamiltonian function

$$\hat{b}_\tau(\eta) = -\frac{1}{2}\, l_\tau^{-1}(\sigma(\dot{J}_\tau \cdot, \eta), \sigma(\dot{J}_\tau \cdot, \eta)), \qquad \eta \in \Sigma.$$

The linear nonautonomous equation (21.23) is *Jacobi equation* for the totally singular case.

Now the next statement follows in the same way as in the regular case.

Theorem 21.2. *Let λ_t be a nice singular extremal. An instant $t \in (0, t_1]$ is a conjugate time iff there exists a nonconstant solution η_τ to Jacobi equation*

$$\dot\eta_\tau = \dot J_\tau \, l_\tau^{-1} \left(\sigma(\dot J_\tau \,\cdot\,, \eta_\tau) \right), \qquad \tau \in [0, t], \tag{21.24}$$

with the boundary conditions

$$\eta_0 \in \Pi_0^{\Gamma_0}, \qquad \eta_t \in \Pi_0. \tag{21.25}$$

Jacobi equation (21.24) *is Hamiltonian:*

$$\dot\eta_\tau = \vec{\hat b}_\tau(\eta_\tau) \tag{21.26}$$

with the nonautonomous quadratic Hamiltonian function

$$\hat b_\tau(\eta) = -\frac{1}{2} \, l_\tau^{-1} \left(\sigma(\dot J_\tau \,\cdot\,, \eta), \sigma(\dot J_\tau \,\cdot\,, \eta) \right), \qquad \eta \in \Sigma.$$

The following statement provides a first integral of equation (21.23), it can be useful in the study of Jacobi equation in the singular case.

Lemma 21.1. *For any constant vector $v \in \mathbb{R}^m$, the function $\sigma(\eta, J_\tau v)$ is an integral of Jacobi equation* (21.23).

Proof. We have to show that

$$\sigma(\dot\eta_\tau, J_\tau v) + \sigma(\eta_\tau, \dot J_\tau v) \equiv 0 \tag{21.27}$$

for a solution η_τ to (21.23). The first term can be computed via Jacobi equation:

$$\begin{aligned} \sigma(\dot\eta_\tau, J_\tau v) &= -\langle d_{\eta_\tau} \hat b_\tau, J_\tau v \rangle \\ &= l_\tau^{-1} \left(\sigma(\dot J_\tau \,\cdot\,, J_\tau v), \sigma(\dot J_\tau \,\cdot\,, \eta_\tau) \right) \end{aligned}$$

where l_τ^{-1} is a bilinear form

$$= \left\langle \sigma(\dot J_\tau \,\cdot\,, \eta_\tau), l_\tau^{-1} \sigma(\dot J_\tau \,\cdot\,, J_\tau v) \right\rangle$$

where l_τ^{-1} is a linear mapping to the dual space

$$= \left\langle \sigma(\dot J_\tau \,\cdot\,, \eta_\tau), v \right\rangle = -\sigma(\eta_\tau, \dot J_\tau v),$$

and equality (21.27) follows. □

In particular, this lemma means that

$$\eta_0 \in \Gamma_0^{\angle} \quad \Leftrightarrow \quad \eta_\tau \in \Gamma_\tau^{\angle},$$

i.e., the flow of Jacobi equation preserves the family of spaces $\Gamma_\tau^{\angle}$. Since this equation is Hamiltonian, its flow preserves also the family Γ_τ. Consequently, boundary conditions (21.22) can equivalently be written in the form

$$\eta_0 \in \Pi_0, \qquad \eta_t \in \Pi_0^{\Gamma_t}.$$

21.3 Necessary optimality conditions

Proposition 21.2. *Let $(\tilde{u}, \lambda_t)$ be a corank one extremal pair. Suppose that λ_t is regular or nice singular. Let $t_* \in (0, t_1]$. Then:*

(1) *Either for any nonconstant solution η_t, $t \in [0, t_*]$, to Jacobi equation* (21.13) *or* (21.26) *that satisfies the boundary conditions* (21.12) *or* (21.25) *the continuation*

$$\bar{\eta}_t = \begin{cases} \eta_t, & t \in [0, t_*], \\ \eta_{t_*}, & t \in [t_*, t_1], \end{cases} \tag{21.28}$$

satisfies Jacobi equation on $[0, t_1]$,

(2) *Or the control $\tilde{u}$ is not locally geometrically optimal on $[0, t_1]$.*

Proof. Assume that condition (1) does not hold, we prove that condition (2) is then satisfied. Take any nonzero $v \in \operatorname{Ker}\left(Q_{t_*}|_{K_{t_*}}\right)$ and let η_t, $t \in [0, t_*]$, be the corresponding nonconstant solution to Jacobi equation with the boundary conditions. Consider the continuation of v by zero:

$$\bar{v}(t) = \begin{cases} v(t), & t \in [0, t_*], \\ 0, & t \in [t_*, t_1]. \end{cases}$$

and the corresponding continuation by constant $\bar{\eta}_t$ as in (21.28). Since $\bar{\eta}_t$ does not satisfy Jacobi equation on $[0, t_1]$, then $\bar{v} \notin \operatorname{Ker}(Q_{t_1}|_{K_{t_1}})$. Notice that $Q_{t_1}(\bar{v}) = Q_{t_*}(v) = 0$. On the other hand, there exists $w \in K_{t_1}$ such that $Q_{t_1}(\bar{v}, w) \neq 0$. Then the quadratic form Q_{t_1} takes values of both signs in the plane $\operatorname{span}(\bar{v}, w)$.

In the singular case, since the extended form Q_t is sign-indefinite, then the initial form is sign-indefinite as well.

Summing up, the form Q_{t_1} is sign-indefinite on K_{t_1}. By Theorem 20.1, the control $\tilde{u}(t)$ is not optimal on $[0, t_1]$. □

Notice that case (1) of Proposition 21.2 imposes a strong restriction on an extremal λ_t. If this case realizes, then the set of conjugate points coincides with the segment $[t_*, t_1]$.

Assume that the reference control $\tilde{u}(t)$ is analytic, then solutions η_t to Jacobi equation are analytic as well. If η_t is constant on some segment, then it is constant on the whole domain. Thus in the analytic case alternative (1) of Proposition 21.2 is impossible, and the first conjugate time t_* provides a necessary optimality condition: a trajectory cannot be locally geometrically optimal after t_*.

Absence of conjugate points implies finite-dimensional local optimality in the corank one case, see Theorem 20.1. In the following two sections, we prove a much stronger result for the regular case: absence of conjugate points is sufficient for strong optimality.

21.4 Regular case: transformation of Jacobi equation

Let λ_t be a regular extremal, and assume that the maximized Hamiltonian $H(\lambda)$ is smooth in a neighborhood of λ_t. The maximality condition of PMP yields the equation

$$\frac{\partial h_u}{\partial u}(\lambda) = 0,$$

which can be resolved in the neighborhood of λ_t:

$$\frac{\partial h_u}{\partial u}(\lambda) = 0 \quad \Leftrightarrow \quad u = u(\lambda).$$

The mapping $\lambda \mapsto u(\lambda)$ is smooth near λ_t and satisfies the equality

$$u(\lambda_t) = \tilde{u}(t).$$

The maximized Hamiltonian of PMP is expressed in the neighborhood of λ_t as

$$H(\lambda) = h_{u(\lambda)}(\lambda),$$

see Proposition 12.1. Consider the flow on T^*M:

$$e^{t\vec{H}} \circ \overleftarrow{\exp} \int_0^t -\vec{h}_{\tilde{u}(\tau)}\, d\tau = e^{t\vec{H}} \circ P_t^*.$$

By the variations formula in the Hamiltonian form, see (2.27) and (11.24), this flow is Hamiltonian:

$$e^{t\vec{H}} \circ P_t^* = \overrightarrow{\exp} \int_0^t \vec{\varphi}_\tau\, d\tau \tag{21.29}$$

with the Hamiltonian function

$$\varphi_t(\lambda) = (H - h_{\tilde{u}(t)})(P_t^{*\,-1}(\lambda)).$$

Notice that

$$\lambda_0 \circ e^{t\vec{H}} \circ P_t^* = \lambda_t \circ P_t^* = \lambda_0,$$

i.e., λ_0 is an equilibrium point of the field $\vec{\varphi}_t$. In other words, λ_0 is a critical point of the Hamiltonian function:

$$\varphi_t(\lambda) \geq 0 = \varphi_t(\lambda_0) \quad \Rightarrow \quad d_{\lambda_0}\varphi_t = 0.$$

It is natural to expect that the corresponding Hessian is related to optimality of the extremal λ_t.

The following statement relates two Hamiltonian systems: Jacobi equation on Σ and the maximized Hamiltonian system on T^*M. We will use this relation in the proof of sufficient optimality conditions in Section 21.5.

Proposition 21.3. *The Hamiltonian b_t of Jacobi equation coincides with one half of Hessian of the Hamiltonian φ_t at λ_0:*

$$b_t = \frac{1}{2}\operatorname{Hess}_{\lambda_0}\varphi_t.$$

Proof. Recall that Hamiltonian of Jacobi equation for the regular case is

$$b_t(\eta) = -\frac{1}{2}\left\langle \sigma(J_t\,\cdot\,,\eta), (h_t'')^{-1}\sigma(J_t\,\cdot\,,\eta)\right\rangle.$$

Transform the linear form:

$$\sigma(J_t\,\cdot\,,\eta) = \sigma\left(\frac{\partial}{\partial u}\,\overrightarrow{h_{u,t}}\ \cdot\,,\eta\right)$$

where $h_{u,t}(\lambda) = h_u(P_t^{*-1}(\lambda))$

$$= -\left\langle d_{\lambda_0}\frac{\partial}{\partial u}h_{u,t}\,\cdot\,,\eta\right\rangle = -\left\langle \left(d_{\lambda_t}\frac{\partial h_u}{\partial u}\,\cdot\right)\left(P_t^{*-1}\right)_{*\lambda_0},\eta\right\rangle$$

where $\left(P_t^{*-1}\right)_*$ is differential of the diffeomorphism $\left(P_t^{*-1}\right) : T^*M \to T^*M$

$$= -\left\langle d_{\lambda_t}\frac{\partial h_u}{\partial u}\,\cdot\,,\xi\right\rangle,$$

$$\xi = \left(P_t^{*-1}\right)_{*\lambda_0}\eta \in T_{\lambda_t}(T^*M).$$

Then the Hamiltonian b_t can be rewritten as

$$b_t(\eta) = -\frac{1}{2}\left\langle\left\langle d_{\lambda_t}\frac{\partial h_u}{\partial u}\,\cdot\,,\xi\right\rangle, (h_t'')^{-1}\left\langle d_{\lambda_t}\frac{\partial h_u}{\partial u}\,\cdot\,,\xi\right\rangle\right\rangle.$$

Now we compute Hessian of the Hamiltonian

$$\varphi_t(\lambda) = (h_{u(\lambda)} - h_{\tilde{u}(t)})(P_t^{*-1}(\lambda)).$$

We have

$$\operatorname{Hess}_{\lambda_0}\varphi_t(\eta) = \operatorname{Hess}_{\lambda_t}(h_{u(\lambda)} - h_{\tilde{u}(t)})(\xi).$$

Further,

$$d_\lambda(h_{u(\lambda)} - h_{\tilde u(t)}) = \underbrace{\left.\frac{\partial h_u}{\partial u}\right|_{u(\lambda)}}_{\equiv 0} d_\lambda u + (d_\lambda h_u)|_{u(\lambda)} - d_\lambda h_{\tilde u(t)},$$

$$D^2_{\lambda_t}(h_{u(\lambda)} - h_{\tilde u(t)}) = \left(d_{\lambda_t} \left.\frac{\partial h_u}{\partial u}\right|_{u(\lambda_t)} \right) d_{\lambda_t} u.$$

The differential $d_{\lambda_t} u$ can be found by differentiation of the identity

$$\left.\frac{\partial h_u}{\partial u}\right|_{u(\lambda)} \equiv 0$$

at $\lambda = \lambda_t$. Indeed, we have

$$\frac{\partial^2 h_u}{\partial u^2} d_\lambda u + d_\lambda \frac{\partial h_u}{\partial u} = 0,$$

thus

$$d_{\lambda_t} u = -(h''_t)^{-1} d_{\lambda_t} \frac{\partial h_u}{\partial u}.$$

Consequently,

$$D^2_{\lambda_t}(h_{u(\lambda)} - h_{\tilde u(t)}) = -d_{\lambda_t} \frac{\partial h_u}{\partial u} (h''_t)^{-1} d_{\lambda_t} \frac{\partial h_u}{\partial u},$$

i.e.,

$$\operatorname{Hess}_{\lambda_0} \varphi_t(\eta) = \operatorname{Hess}_{\lambda_t}(h_{u(\lambda)} - h_{\tilde u(t)})(\xi) = 2b_t(\eta),$$

and the statement follows. □

Since the Hamiltonian φ_t attains minimum at λ_0, the quadratic form b_t is nonnegative:

$$b_t \geq 0.$$

Denote by C_t the space of constant vertical solutions to Jacobi equation at the segment $[0, t]$:

$$C_t = \left\{ \eta \in \Pi_0 \mid \vec{b}_\tau(\eta) \equiv 0, \ \tau \in [0, t] \right\}. \qquad (21.30)$$

Now we can give the following simple characterization of this space:

$$C_t = \Pi_0 \bigcap \left(\cap_{\tau \in [0,t]} \operatorname{Ker} b_\tau \right).$$

Indeed, equilibrium points of a Hamiltonian vector field are critical points of the Hamiltonian, and critical points of a nonnegative quadratic form are elements of its kernel.

21.5 Sufficient optimality conditions

In this section we prove sufficient conditions for optimality in the problem with integral cost:

$$\begin{aligned}
&\dot{q} = f_u(q), \qquad q \in M, \quad u \in U = \operatorname{int} U \subset \mathbb{R}^m,\\
&q(0) = q_0, \qquad q(t_1) = q_1,\\
&\int_0^{t_1} \varphi(q(t), u(t))\, dt \to \min,
\end{aligned}$$

with fixed or free terminal time. Notice that now we study an optimal problem, not a geometric one as before. Although, the theory of Jacobi equation can be applied here since Jacobi equation depends only on a Hamiltonian $h_u(\lambda)$ and an extremal pair $(\tilde{u}(t), \lambda_t)$.

For the normal Hamiltonian of PMP

$$h_u(\lambda) = \langle \lambda, f_u(q) \rangle - \varphi(q, u), \qquad \lambda \in T^*M,$$

and a regular extremal pair $(\tilde{u}(t), \lambda_t)$ of the optimal control problem, consider Jacobi equation

$$\dot{\eta} = \vec{b}_t(\eta), \qquad \eta \in \Sigma = T_{\lambda_0}(T^*M).$$

In Section 21.3 we showed that absence of conjugate points at the interval $(0, t_1)$ is necessary for geometric optimality (at least in the corank one analytic case).

Exercise 21.1. Show that absence of conjugate points on $(0, t_1)$ is necessary also for optimality (in the analytic case) reducing the optimal control problem to a geometric one.

Now we can show that absence of conjugate points is also sufficient for optimality (in the regular case).

A trajectory $q(t)$, $t \in [0, t_1]$, is called *strongly optimal* for an optimal control problem if it realizes a local minimum of the cost functional w.r.t. all trajectories of the system close to $q(t)$ in the uniform topology $C([0, t_1], M)$ and having the same endpoints as $q(t)$.

Theorem 21.3. *Let λ_t, $t \in [0, t_1]$, be a regular normal extremal in the problem with integral cost and fixed time, and let the maximized Hamiltonian $H(\lambda)$ be smooth in a neighborhood of λ_t. If the segment $(0, t_1]$ does not contain conjugate points, then the extremal trajectory $q(t) = \pi(\lambda_t)$, $t \in [0, t_1]$, is strongly optimal.*

Proof. We apply the theory of fields of extremals (see Section 17.1) and embed λ_t into a family of extremals well projected to M.

The maximized Hamiltonian

$$H(\lambda) = \max_{u \in U} h_u(\lambda), \qquad \lambda \in T^*M,$$

is defined and smooth. Then by Theorem 17.1, it is enough to construct a function $a \in C^\infty(M)$ such that the family of manifolds

$$\begin{aligned} &\mathcal{L}_t = e^{t\vec{H}}(\mathcal{L}_0) \subset T^*M, \qquad t \in [0, t_1], \\ &\mathcal{L}_0 = \{\lambda = d_q a\} \subset T^*M, \\ &\lambda_0 \in \mathcal{L}_0, \end{aligned}$$

has a good projection to M:

$$\pi : \mathcal{L}_t \to M \text{ is a diffeomorphism near } \lambda_t, \quad t \in [0, t_1].$$

In other words, we require that the tangent spaces $T_{\lambda_t}\mathcal{L}_t = e_*^{t\vec{H}}(T_{\lambda_0}\mathcal{L}_0)$ have zero intersection with the vertical subspaces $\Pi_t = T_{\lambda_t}(T^*_{q(t)}M)$:

$$e_*^{t\vec{H}}(T_{\lambda_0}\mathcal{L}_0) \cap \Pi_t = \{0\}, \qquad t \in [0, t_1].$$

Below we show that such a manifold $\mathcal{L}_0$ exists by passing to its tangent space L_0 — a Lagrangian subspace in Σ (see definition in Subsection 11.5.3). For any Lagrangian subspace $L_0 \subset \Sigma$ transversal to Π_0, one can find a function $a \in C^\infty(M)$ such that the graph of its differential $\mathcal{L}_0 = \{\lambda = d_q a\} \subset T^*M$ satisfies the conditions:

(1) $\lambda_0 \in \mathcal{L}_0$,

(2) $T_{\lambda_0}\mathcal{L}_0 = L_0$.

Indeed, in canonical coordinates (p, q) on T^*M, take a function of the form

$$a(q) = \langle p_0, q\rangle + \frac{1}{2}q^T S q, \qquad \lambda_0 = (p_0, 0),$$

with a symmetric $n \times n$ matrix S. Then

$$\begin{aligned} &\mathcal{L}_0 = \{\lambda = (p, q) \mid p = p_0 + Sq\}, \\ &T_{\lambda_0}\mathcal{L}_0 = \{(dp, dq) \mid dp = S dq\} \end{aligned}$$

and it remains to choose the linear mapping S with the graph L_0. Notice that the symmetry of the matrix S corresponds to the Lagrangian property of the subspace L_0. Below we use a similar construction for parametrization of Lagrangian subspaces by quadratic forms.

To complete the proof, we have to find a Lagrangian subspace $L_0 \subset \Sigma$ such that

$$\left(e_*^{t\vec{H}} L_0\right) \cap \Pi_t = \{0\}, \qquad t \in [0, t_1].$$

By (21.29), the flow of the maximized Hamiltonian decomposes:

$$e^{t\vec{H}} = \Phi_t \circ P_t^{*-1}, \qquad \Phi_t = \overrightarrow{\exp} \int_0^t \vec{\varphi}_\tau \, d\tau.$$

Notice that the flow P_t^{*-1} on T^*M is induced by the flow P_t on M, thus it preserves the family of vertical subspaces:

$$\left(P_t^{*-1}\right)_* \Pi_0 = \Pi_t.$$

So it remains to show that there exists a Lagrangian subspace $L_0 \subset \Sigma$ for which

$$(\Phi_{t*} L_0) \cap \Pi_0 = \{0\}, \qquad t \in [0, t_1]. \tag{21.31}$$

Proposition 21.3 relates the Hamiltonian b_t of Jacobi equation to the Hamiltonian φ_t:

$$\frac{1}{2} \operatorname{Hess}_{\lambda_0} \varphi_t = b_t.$$

Thus the field $\vec{b}_t$ is the linearization of the field $\vec{\varphi}_t$ at the equilibrium point λ_0: the Hamiltonian b_t and the Hamiltonian field $\vec{b}_t$ are respectively the main terms in Taylor expansion of φ_t and $\vec{\varphi}_t$ at λ_0. Linearization of a flow is the flow of the linearization, thus

$$\left(\overrightarrow{\exp} \int_0^t \vec{\varphi}_\tau \, d\tau \right)_{*\lambda_0} = \overrightarrow{\exp} \int_0^t \vec{b}_\tau \, d\tau.$$

Introduce notation for the flow of Jacobi equation:

$$B_t = \overrightarrow{\exp} \int_0^t \vec{b}_\tau \, d\tau,$$

then

$$\Phi_{t*\lambda_0} = B_t,$$

and equality (21.31) reads

$$(B_t L_0) \cap \Pi_0 = \{0\}, \qquad t \in [0, t_1]. \tag{21.32}$$

It remains to prove existence of a Lagrangian subspace L_0 that satisfies this equality.

Recall that the segment $(0, t_1]$ does not contain conjugate points:

$$(B_t \Pi_0) \cap \Pi_0 = C_t, \qquad t \in (0, t_1],$$

where C_t is the space of constant vertical solutions to Jacobi equation on $[0, t]$, see (21.30).

In order to make the main ideas of the proof more clear, we consider first the simple case where

$$C_t = \{0\}, \qquad t \in (0, t_1], \tag{21.33}$$

i.e.,

$$(B_t \Pi_0) \cap \Pi_0 = \{0\}, \qquad t \in (0, t_1].$$

Fix any $\varepsilon \in (0, t_1)$. By continuity of the flow B_t, there exists a neighborhood of the vertical subspace Π_0 such that for any Lagrangian subspace L_0 from this neighborhood

$$(B_t L_0) \cap \Pi_0 = \{0\}, \qquad t \in [\varepsilon, t_1].$$

In order to complete the proof, it remains to find such a Lagrangian subspace L_0 satisfying the condition

$$(B_t L_0) \cap \Pi_0 = \{0\}, \qquad t \in [0, \varepsilon].$$

We introduce a parametrization of the set of Lagrangian subspaces $L_0 \subset \Sigma$ sufficiently close to Π_0. Take any Lagrangian subspace $H \subset \Sigma$ which is horizontal, i.e., transversal to the vertical subspace Π_0. Then the space Σ splits:

$$\Sigma = \Pi_0 \oplus H.$$

Introduce Darboux coordinates (p, q) on Σ such that

$$\Pi_0 = \{(p, 0)\}, \qquad H = \{(0, q)\}.$$

Such coordinates can be chosen in many ways. Indeed, the symplectic form σ defines a nondegenerate pairing of the mutually transversal Lagrangian subspaces Π_0 and H:

$$\begin{aligned} &H = \Pi_0^*, \\ &\langle f, e \rangle = \sigma(e, f), \qquad e \in \Pi_0, \quad f \in H. \end{aligned}$$

Taking any basis $e_1, \dots, e_n$ in Π_0 and the corresponding basis $f_1, \dots, f_n$ in H dual w.r.t. this pairing, we obtain a Darboux basis in Σ. In Darboux coordinates the symplectic form reads

$$\sigma((p_1, q_1), (p_2, q_2)) = \langle p_1, q_2 \rangle - \langle p_2, q_1 \rangle.$$

Any n-dimensional subspace $L \subset \Sigma$ transversal to H is a graph of a linear mapping

$$S : \Pi_0 \to H,$$

i.e.,

$$L = \{(p, Sp) \mid p \in \Pi_0\}.$$

A subspace L is Lagrangian iff the corresponding mapping S has a symmetric matrix in a symplectic basis (exercise):

$$S = S^*.$$

Introduce the quadratic form on Π_0 with the matrix S:

$$S(p, p) = \langle p, Sp \rangle.$$

So the set of Lagrangian subspaces $L \subset \Sigma$ transversal to the horizontal space H is parametrized by quadratic forms S on Π_0. We call such parametrization of Lagrangian subspaces $L \subset \Sigma$, $L \cap H = \{0\}$, a (Π_0, H)-parametrization.

Consider the family of quadratic forms S_t that parametrize a family of Lagrangian subspaces of the form

$$L_t = B_t L_0,$$

i.e.,

$$L_t = \{(p, S_t p) \mid p \in \Pi_0\}.$$

Lemma 21.2.

$$\dot{S}_t(p,p) = 2b_t(p, S_t p).$$

Proof. Take any trajectory $(p,q) = (p_t, q_t)$ of the Hamiltonian field $\vec{b}_t$. We have

$$q = S_t p,$$

thus

$$\dot{q} = \dot{S}_t p + S_t \dot{p},$$

i.e.,

$$\vec{b}_t(p,q) = \left(\dot{p}, \dot{S}_t p + S_t \dot{p}\right).$$

Since the Hamiltonian b_t is quadratic, we have

$$\sigma\left((p,q), \vec{b}_t(p,q)\right) = 2b_t(p,q).$$

But the left-hand side is easily computed:

$$\begin{aligned}
\sigma\left((p,q), \vec{b}_t(p,q)\right) &= \sigma\left((p,q), (\dot{p}, \dot{q})\right) \\
&= \sigma\left((p, S_t p), (\dot{p}, \dot{S}_t p + S_t \dot{p})\right) = \left\langle p, \dot{S}_t p + S_t \dot{p}\right\rangle - \langle \dot{p}, S_t p\rangle \\
&= \left\langle p, \dot{S}_t p\right\rangle
\end{aligned}$$

by symmetry of S_t. $\square$

Since the Hamiltonian φ_t attains minimum at λ_0, then $b_t \geq 0$, thus

$$\dot{S}_t \geq 0.$$

The partial order on the space of quadratic forms induced by positive forms explains how one should choose the initial subspace L_0. Taking any Lagrangian subspace $L_0 \subset \Sigma$ with the corresponding quadratic form

$$S_0 > 0$$

sufficiently close to the zero form, we obtain

$$S_t > 0, \qquad t \in [0, \varepsilon].$$

That is,

$$L_t \cap \Pi_0 = \{0\}$$

on $[0, \varepsilon]$, thus on the whole segment $[0, t_1]$.

We proved equality (21.32) in the simple case (21.33). Now we consider the general case. The intersection $(B_t \Pi_0) \cap \Pi_0 = C_t$ is nonempty now, but we can get rid of it by passing to Jacobi equation on the quotient $C_t^\angle / C_t$.

The family of constant vertical solutions C_t is monotone nonincreasing:

$$C_{t'} \supset C_{t''} \text{ for } t' < t''.$$

We have $C_0 = \Pi_0$ and set, by definition, $C_{t_1+0} = \{0\}$. The family C_t is continuous from the left, denote its discontinuity points:

$$0 \le s_1 < s_2 < \cdots < s_k \le t_1$$

(notice that in the simple case (21.33), we have $k = 1$, $s_1 = 0$). The family C_t is constant on the segments $(s_i, s_{i+1}]$.

Construct subspaces $E_i \subset \Pi_0$, $i = 1, \ldots, k$, such that

$$C_t = E_{i+1} \oplus E_{i+2} \oplus \cdots \oplus E_k, \qquad t \in (s_i, s_{i+1}].$$

Notice that for $t = 0$, we obtain a splitting of the vertical subspace:

$$\Pi_0 = C_0 = E_1 \oplus \cdots \oplus E_k.$$

For any horizontal Lagrangian subspace $H \subset \Sigma$, one can construct the corresponding splitting of H:

$$H = F_1 \oplus \cdots \oplus F_k, \qquad \sigma(E_i, F_j) = 0, \quad i \ne j. \tag{21.34}$$

Fix any initial horizontal subspace $H_0 \subset \Sigma$, $H_0 \cap \Pi_0 = \{0\}$. The following statement completes the proof of Theorem 21.3 in the general case.

Lemma 21.3. *For any $i = 1, \ldots, k$, there exist a number $\varepsilon_i > 0$ and a Lagrangian subspace $H_i \subset \Sigma$, $H_i \cap \Pi_0 = \{0\}$, such that any Lagrangian subspace $L_0 \subset \Sigma$, $L_0 \cap H_0 = \{0\}$, with a (Π_0, H_0)-parametrization $S_0(p,p) = \varepsilon\langle p, p\rangle$, $0 < \varepsilon < \varepsilon_i$, satisfies the conditions:*

(1) $L_t \cap \Pi_0 = \{0\}, \qquad t \in [0, s_i]$,

(2) $L_t \cap H_i = \{0\}$, $t \in [0, s_i]$, *and the Lagrangian subspace L_t has a (Π_0, H_i)-parametrization $S_t > 0$.*

Proof. We prove this lemma by induction on i.

Let $i = 1$. For $s_1 = 0$, the statement is trivial, so we assume that $s_1 > 0$. Take any $\varepsilon_1 > 0$ and any Lagrangian subspace $L_0 \subset \Sigma$ with a quadratic form $\varepsilon\langle p, p\rangle$, $0 < \varepsilon < \varepsilon_1$, in the (Π_0, H_0)-parametrization.

Notice that $C_t = \Pi_0$, i.e., $B_t|_{\Pi_0} = \mathrm{Id}$, for $t \in (0, s_1]$. We have

$$L_t \cap \Pi_0 = B_t L_0 \cap B_t \Pi_0 = B_t(L_0 \cap \Pi_0) = \{0\}, \qquad t \in [0, s_1].$$

By continuity of the flow B_t, there exists a horizontal Lagrangian subspace H_1 with a (Π_0, H_0)-parametrization $-\delta\langle p, p\rangle$, $\delta > 0$, such that $L_t \cap H_1 = \{0\}$, $t \in [0, s_1]$. One can easily check that the subspace L_0 in (Π_0, H_1)-parametrization is given by the quadratic form $S_0(p, p) = \varepsilon'\langle p, p\rangle > 0$, $\varepsilon' = \varepsilon/(1 + \varepsilon/\delta) < \varepsilon$. We already proved that $\dot{S}_t \geq 0$, thus

$$S_t > 0, \qquad t \in [0, s_1],$$

in the (Π_0, H_1)-parametrization.

The induction basis $(i = 1)$ is proved.

Now we prove the induction step. Fix $i \geq 1$, assume that the statement of Lemma 21.3 is proved for i, and prove it for $i + 1$.

Let $t \in (s_i, s_{i+1}]$, then $C_t = E_{i+1} \oplus \cdots \oplus E_k$. Introduce a splitting of the horizontal subspace H_i as in (21.34):

$$H_i = F_1 \oplus \cdots \oplus F_k.$$

Denote

$$\begin{aligned} E_1' &= E_1 \oplus \cdots \oplus E_i, & E_2' &= C_t = E_{i+1} \oplus \cdots \oplus E_k, \\ F_1' &= F_1 \oplus \cdots \oplus F_i, & F_2' &= F_{i+1} \oplus \cdots \oplus F_k, \\ L_0^1 &= L_0 \cap (E_1' \oplus F_1'), & L_0^2 &= L_0 \cap (E_2' \oplus F_2'). \end{aligned}$$

Since $B_t E_2' = E_2'$, then the skew-orthogonal complement $(E_2')^\angle = E_1' \oplus E_2' \oplus F_1'$ is also invariant for the flow of Jacobi equation: $B_t(E_2')^\angle = (E_2')^\angle$.

In order to prove that $L_t \cap \Pi_0 = \{0\}$, compute this intersection. We have $\Pi_0 \subset (E_2')^\angle$, thus

$$B_t L_0 \cap \Pi_0 = B_t L_0 \cap B_t (E_2')^\angle \cap \Pi_0 = B_t(L_0 \cap (E_2')^\angle) \cap \Pi_0 = B_t L_0^1 \cap \Pi_0. \tag{21.35}$$

So we have to prove that $B_t L_0^1 \cap \Pi_0 = \{0\}$, $t \in (s_i, s_{i+1}]$.

Since the subspaces E_2' and $(E_2')^\angle$ are invariant w.r.t. the flow B_t, the quotient flow is well-defined:

$$\widetilde{B}_t : \widetilde{\Sigma} \to \widetilde{\Sigma}, \qquad \widetilde{\Sigma} = (E_2')^\angle / E_2'.$$

In the quotient, the flow $\widetilde{B}_t$ has no constant vertical solutions:

$$\begin{aligned} &\widetilde{B}_t \widetilde{\Pi}_0 \cap \widetilde{\Pi}_0 = \{0\}, \qquad t \in (s_i, s_{i+1}], \\ &\widetilde{\Pi}_0 = \Pi_0 / E_2'. \end{aligned}$$

By the argument already used in the proof of the simple case (21.33), it follows that

$$\widetilde{B}_t \widetilde{L}_0^1 \cap \widetilde{\Pi}_0 = \{0\}, \qquad t \in (s_i, s_{i+1}],$$
$$\widetilde{L}_0^1 = L_0^1 / E_2',$$

for L_0 sufficiently close to Π_0, i.e., for ε sufficiently small. That is,

$$B_t L_0^1 \cap \Pi_0 \subset E_2', \qquad t \in (s_i, s_{i+1}].$$

Now it easily follows that this intersection is empty:

$$B_t L_0^1 \cap \Pi_0 \subset B_t L_0^1 \cap E_2' = B_t L_0^1 \cap B_t E_2' = B_t(L_0^1 \cap E_2') = \{0\}, \qquad t \in (s_i, s_{i+1}].$$

In view of chain (21.35),

$$L_t \cap \Pi_0 = \{0\}, \qquad t \in (s_i, s_{i+1}],$$

that is, we proved condition (1) in the statement of Lemma 21.3 for $i+1$.

Now we pass to condition (2). In the same way as in the proof of the induction basis, it follows that there exists a horizontal Lagrangian subspace $H_{i+1} \subset \Sigma$ such that the curve of Lagrangian subspaces L_t, $t \in [0, s_{i+1}]$, is transversal to H_{i+1}. In the (Π_0, H_{i+1})-parametrization, the initial subspace L_0 is given by a positive quadratic form $S_0(p,p) = \varepsilon' \langle p, p \rangle$, $0 < \varepsilon' < \varepsilon$. Since $\dot{S}_t \geq 0$, then

$$S_t > 0, \qquad t \in [0, s_{i+1}].$$

Condition (2) is proved for $i+1$.

The induction step is proved, and the statement of this lemma follows. □

By this lemma,

$$L_t \cap \Pi_0 = \{0\}, \qquad t \in [0, t_1],$$

for all initial subspaces L_0 given by quadratic forms $S_0 = \varepsilon \langle p, p \rangle$, $0 < \varepsilon < \varepsilon_k$, for some $\varepsilon_k > 0$, in a (Π_0, H_0)-parametrization. This means that we constructed a family of extremals containing λ_t and having a good projection to M. By Theorem 17.1, the extremal λ_t, $t \in [0, t_1]$, is strongly optimal. Theorem 21.3 is proved. □

For the problem with integral cost and free terminal time t_1, a similar argument and Theorem 17.2 yield the following sufficient optimality condition.

Theorem 21.4. *Let λ_t, $t \in [0, t_1]$, be a regular normal extremal in the problem with integral cost and free time, and let $H(\lambda)$ be smooth in a neighborhood of λ_t. If there are no conjugate points at the segment $(0, t_1]$, then the extremal trajectory $q(t) = \pi(\lambda_t)$, $t \in [0, t_1]$, is strongly optimal.*

Chapter 22

Reduction

In this chapter we consider a method for reducing a system affine in control to a nonlinear system on a manifold of a less dimension.

22.1 Reduction

Consider an affine in control system

$$\dot{q} = f(q) + \sum_{i=1}^{m} u_i g_i(q), \qquad u_i \in \mathbb{R}, \quad q \in M, \tag{22.1}$$

with pairwise commuting vector fields near controls:

$$[g_i, g_j] \equiv 0, \qquad i,\ j = 1, \dots, m.$$

The flow of the system can be decomposed by the variations formula:

$$\overrightarrow{\exp} \int_0^t \left(f + \sum_{i=1}^{m} u_i(\tau) g_i \right) d\tau = \overrightarrow{\exp} \int_0^t e^{\sum_{i=1}^{m} w_i(\tau) \operatorname{ad} g_i} f \, d\tau \circ e^{\sum_{i=1}^{m} w_i(t) g_i}, \tag{22.2}$$

$$w_i(t) = \int_0^t u_i(\tau) \, d\tau.$$

Here we treat $\sum_{i=1}^{m} u_i(\tau) g_i$ as a nonperturbed flow and take into account that the fields g_i mutually commute. Introduce the partial system corresponding to the second term in composition (22.2):

$$\dot{q} = e^{\sum_{i=1}^{m} w_i \operatorname{ad} g_i} f(q), \qquad w_i \in \mathbb{R}, \quad q \in M, \tag{22.3}$$

where w_i are new controls. Attainable sets $\mathcal{A}_1(t)$ of the initial system (22.1) and $\mathcal{A}_2(t)$ of the partial system (22.3) for time t from a point $q_0 \in M$ are closely related one to another:

$$\mathcal{A}_1(t) \subset \mathcal{A}_2(t) \circ \left\{ e^{\sum_{i=1}^{m} w_i g_i} \mid w_i \in \mathbb{R} \right\} \subset \operatorname{cl}(\mathcal{A}_1(t)). \tag{22.4}$$

Indeed, the first inclusion follows directly from decomposition (22.2). To prove the second inclusion in (22.4), notice that the mapping

$$w(\cdot) \mapsto q_0 \circ \overrightarrow{\exp} \int_0^t e^{\sum_{i=1}^m w_i(\tau) \operatorname{ad} g_i} f \, d\tau$$

is continuous in L_1 topology, this follows from the asymptotic expansion of the chronological exponential. Thus the mapping

$$(w(\cdot), v) \mapsto q_0 \circ \overrightarrow{\exp} \int_0^t e^{\sum_{i=1}^m w_i(\tau) \operatorname{ad} g_i} f \, d\tau \circ e^{\sum_{i=1}^m v_i g_i}$$

is continuous in topology of $L_1 \times \mathbb{R}^m$. Finally, the mapping

$$u(\cdot) \mapsto (w(\cdot), v) = \left(\int_0^{\cdot} u(\tau) \, d\tau, \int_0^t u(\tau) \, d\tau \right)$$

has a dense image in $L_1 \times \mathbb{R}^m$. Then decomposition (22.2) implies the second inclusion in (22.4).

The partial system (22.3) is invariant w.r.t. the fields g_i:

$$\left(e^{\sum_{i=1}^m v_i g_i} \right)_* e^{\sum_{i=1}^m w_i \operatorname{ad} g_i} f = e^{\sum_{i=1}^m (w_i - v_i) \operatorname{ad} g_i} f. \tag{22.5}$$

Thus chain (22.4) and equality (22.5) mean that the initial system (22.1) can be considered as a composition of the partial system (22.3) with the flow of the fields g_i: any time t attainable set of the initial system is (up to closure) the time t attainable set of the partial system plus a jump along g_i, moreover, the jump along g_i is possible at any instant.

Let $(u(t), \lambda_t)$ be an extremal pair of the initial affine in control system. The extremal λ_t is necessarily totally singular, moreover the maximality condition of PMP is equivalent to the identity

$$\langle \lambda_t, g_i \rangle \equiv 0.$$

It is easy to see that

$$\mu_t = \left(e^{\sum_{i=1}^m w_i(t) g_i} \right)^* \lambda_t$$

is an extremal of system (22.3) corresponding to the control

$$w(t) = \int_0^t u(\tau) \, d\tau,$$

moreover,

$$\langle \mu_t, g_i \rangle \equiv 0. \tag{22.6}$$

(Here, we use the term extremal as a synonym of a critical point of the endpoint mapping, i.e., we require that the extremal control be critical, not necessarily minimizing, for the control-dependent Hamiltonian of PMP.) Conversely, if μ_t is an extremal of (22.3) with a Lipschitzian control $w(t)$, and if identity (22.6) holds, then

$$\lambda_t = \left(e^{-\sum_{i=1}^m w_i(t) g_i}\right)^* \mu_t$$

is an extremal of the initial system (22.1) with the control

$$u(t) = \dot{w}(t).$$

Moreover, the strong generalized Legendre condition for an extremal λ_t of the initial system coincides with the strong Legendre condition for the corresponding extremal μ_t of the partial system. In other words, the passage from system (22.1) to system (22.3) transforms nice singular extremals λ_t into regular extremals μ_t.

Exercise 22.1. Check that the initial and reduced systems have the same conjugate times.

Since system (22.3) is invariant w.r.t. the fields g_i, this system can be considered on the quotient manifold of M modulo action of the fields g_i if the quotient manifold is well-defined. Consider the following equivalence relation on M:

$$q' \sim q \quad \Leftrightarrow \quad q' \in \mathcal{O}_q(g_1, \dots, g_m).$$

Suppose that all orbits $\mathcal{O}_q(g_1, \dots, g_m)$ have the same dimension and, moreover, the following nonrecurrence condition is satisfied: for each point $q \in M$ there exist a neighborhood $O_q \ni q$ and a manifold $N_q \subset M$, $q \in N_q$, transversal to $\mathcal{O}_q(g_1, \dots, g_m)$, such that any orbit $\mathcal{O}_{q'}(g_1, \dots, g_m)$, $q' \in O_q$, intersects N_q at a unique point. In particular, these conditions hold if $M = \mathbb{R}^n$ and g_i are constant vector fields, or if $m = 1$ and the field g_1 is nonsingular and nonrecurrent. If these conditions are satisfied, then the space of orbits $M/\!\sim$ is a smooth manifold. Then system (22.3) is well-defined on the quotient manifold $M/\!\sim$:

$$\dot{q} = e^{\sum_{i=1}^m w_i \operatorname{ad} g_i} f(q), \qquad w_i \in \mathbb{R}, \quad q \in M/\!\sim. \tag{22.7}$$

The passage from the initial system (22.1) affine in controls to the reduced system (22.7) nonlinear in controls decreases dimension of the state space and transforms singular extremals into regular ones.

Let $\pi : M \to M/\!\sim$ be the projection. For the attainable set $\mathcal{A}_3(t)$ of the reduced system (22.7) from the point $\pi(q_0)$, inclusions (22.4) take the form

$$\mathcal{A}_1(t) \subset \pi^{-1}(\mathcal{A}_3(t)) \subset \operatorname{cl}(\mathcal{A}_1(t)). \tag{22.8}$$

It follows from the analysis of extremals above that $q(t)$ is an extremal curve of the initial system (22.1) iff its projection $\pi(q(t))$ is an extremal curve of the reduced system (22.7). The first inclusion in (22.8) means that if $\pi(q(\tau))$, $\tau \in [0, t]$, is geometrically optimal, then $q(\tau)$, $\tau \in [0, t]$, is also geometrically optimal.

One can also define a procedure of inverse reduction. Given a control system

$$\dot{q} = f(q, w), \qquad q \in M, \quad w \in \mathbb{R}^n, \tag{22.9}$$

we restrict it to Lipschitzian controls $w(\cdot)$ and add an integrator:

$$\begin{cases} \dot{q} = f(q, w), \\ \dot{w} = u, \end{cases} \qquad (q, w) \in M \times \mathbb{R}^n, \quad u \in \mathbb{R}^n. \tag{22.10}$$

Exercise 22.2. Prove that system (22.9) is the reduction of of system (22.10).

22.2 Rigid body control

Consider the time-optimal problem for the system that describes rotations of a rigid body, see Section 19.4:

$$\dot{q} = q(a + ub), \qquad q \in \mathrm{SO}(3), \quad u \in \mathbb{R}, \tag{22.11}$$

where

$$a, \ b \in \mathrm{so}(3), \quad \langle a, b \rangle = 0, \quad |b| = 1, \quad a \neq 0.$$

Notice that in Section 19.4 we assumed $|a| = 1$, not $|b| = 1$ as now, but one case is obtained from another by dividing the right-hand side of the system by a constant.

We construct the reduced system for system (22.11).

The state space SO(3) factorizes modulo orbits qe^{sb}, $s \in \mathbb{R}$, of the field qb. The corresponding equivalence relation is:

$$q \sim qe^{sb}, \qquad s \in \mathbb{R},$$

and the structure of the factor space is described in the following statement.

Proposition 22.1.

$$\mathrm{SO}(3)/\!\sim \ \simeq \ S^2,$$

the canonical projection is

$$q \mapsto q\beta, \qquad q \in \mathrm{SO}(3), \quad \beta \in S^2. \tag{22.12}$$

Here $\beta \in S^2 \subset \mathbb{R}^3$ is the unit vector corresponding to the matrix $b \in \mathrm{so}(3)$:

$$\beta = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{pmatrix}.$$

Proof. The group SO(3) acts transitively on the sphere S^2. The subgroup of SO(3) leaving a point $\beta \in S^2$ fixed consists of rotations around the line β, i.e., it is

$$e^{\mathbb{R}b} = \{e^{sb} \mid s \in \mathbb{R}\}.$$

Thus the quotient $\mathrm{SO}(3)/\{e^{\mathbb{R}b}\} = \mathrm{SO}(3)/\!\sim$ is diffeomorphic to S^2, projection $\mathrm{SO}(3) \to S^2$ is given by (22.12), and level sets of this mapping coincide with orbits of the field qb. □

The partial system (22.3) in this example takes the form

$$\dot{q} = q e^{w \operatorname{ad} b} a, \qquad q \in \mathrm{SO}(3), \quad w \in \mathbb{R},$$

and the reduced system (22.7) is

$$\frac{d}{dt}(q\beta) = q e^{w \operatorname{ad} b} a\beta, \qquad q\beta \in S^2. \tag{22.13}$$

The right-hand side of this symmetric control system defines a circle of radius $|a|$ in the tangent plane $(q\beta)^{\perp} = T_{q\beta}S^2$. In other words, system (22.13) determines a Riemannian metric on S^2. Since the vector fields in the right-hand side of system (22.13) are constant by absolute value, then the time-optimal problem is equivalent to the Riemannian problem (time minimization is equivalent to length minimization if velocity is constant by absolute value).

Extremal curves (geodesics) of a Riemannian metric on S^2 are arcs of great circles, they are optimal up to semicircles. And the antipodal point is conjugate to the initial point. Conjugate points for the initial and reduced systems coincide, thus for both systems extremal curves are optimal up to the antipodal point.

22.3 Angular velocity control

Consider the system that describes angular velocity control of a rotating rigid body, see (6.19):

$$\dot{\mu} = \mu \times \beta\mu + ul, \qquad u \in \mathbb{R}, \quad \mu \in \mathbb{R}^3. \tag{22.14}$$

Here μ is the vector of angular velocity of the rigid body in a coordinate system connected with the body, and $l \in \mathbb{R}^3$ is a unit vector in general position along which the torque is applied. Notice that in Section 6.4 we allowed only torques $u = \pm 1$, while now the torque is unbounded. In Section 8.4 we proved that the system with bounded control is completely controllable (even in the six-dimensional space). Now we show that with unbounded control we have complete controllability in $\mathbb{R}^3$ for an arbitrarily small time.

We apply the reduction procedure to the initial system (22.14). The partial system reads now

$$\begin{aligned}\dot{\mu} &= e^{w \operatorname{ad} l}(\mu \times \beta\mu) \\ &= (\mu + wl) \times \beta(\mu + wl), \qquad w \in \mathbb{R}, \quad \mu \in \mathbb{R}^3.\end{aligned}$$

The quotient of $\mathbb{R}^3$ modulo orbits of the constant field l can be realized as the plane $\mathbb{R}^2$ passing through the origin and orthogonal to l. Then projection $\mathbb{R}^3 \to \mathbb{R}^2$ is the orthogonal projection along l, and the reduced system reads

$$\dot{x} = (x + wl) \times \beta(x + wl) - \langle x \times \beta(x + wl), l\rangle\, l, \qquad x \in \mathbb{R}^2, \quad w \in \mathbb{R}. \tag{22.15}$$

Introduce Cartesian coordinates in $\mathbb{R}^3$ corresponding to the orthonormal frame with basis vectors collinear to the vectors $l, l\times\beta l, l\times(l\times\beta l)$. In these coordinates $x = (x_1, x_2)$ and the reduced system (22.15) takes the form:

$$\dot{x}_1 = b_{13}x_2^2 + ((b_{11} - b_{33})x_2 - b_{23}x_1)w - b_{13}w^2, \tag{22.16}$$

$$\dot{x}_2 = -b_{13}x_1x_2 + ((b_{22} - b_{11})x_1 + b_{23}x_2)w, \tag{22.17}$$

where $b = (b_{ij})$ is the matrix of the operator β in the orthonormal frame. Direct computation shows that $b_{13} < 0$ and $b_{22} - b_{11} \neq 0$. In polar coordinates (r, φ) in the plane (x_1, x_2), the reduced system (22.16), (22.17) reads

$$\dot{r} = rF(\cos\varphi, \sin\varphi)w - b_{13}\cos\varphi\, w^2,$$

$$\dot{\varphi} = -b_{13}r\sin\varphi - (1/r)\sin\varphi\, w^2 + G(\cos\varphi, \sin\varphi)w,$$

where F and G are homogeneous polynomials of degree 2 with $G(\pm 1, 0) = b_{22} - b_{11}$.

Choosing appropriate controls, one can construct trajectories of the system in $\mathbb{R}^2$ of the following two types:

(1) "spirals", i.e., trajectories starting and terminating at the positive semi-axes x_1, not passing through the origin ($r \neq 0$), and rotating counterclockwise ($\dot{\varphi} > 0$),

(2) "horizontal" trajectories almost parallel to the axis x_1 ($\dot{x}_1 \gg \dot{x}_2$).

Moreover, we can move fast along these trajectories. Indeed, system (22.16), (22.17) has an obvious self-similarity — it is invariant with respect to the changes of variables $x_1 \mapsto \alpha x_1$, $x_2 \mapsto \alpha x_2$, $w \mapsto \alpha w$, $t \mapsto \alpha^{-1}t$ ($\alpha > 0$). Consequently, one can find "spirals" arbitrarily far from the origin and with an arbitrarily small time of complete revolution. Further, it is easy to see from equations (22.16), (22.17) that taking large in absolute value controls w one obtains arbitrarily fast motions along the "horizontal" trajectories in the positive direction of the axis x_1.

Combining motions of types (1) and (2), we can steer any point $x \in \mathbb{R}^2$ to any point $\hat{x} \in \mathbb{R}^2$ for any time $\varepsilon > 0$. Details of this argument are left to the reader as an exercise, see also [3].

That is, time t attainable sets $\mathcal{A}_x^3(t)$ of the reduced system (22.15) from a point x satisfy the property:

$$\mathcal{A}_x^3(\varepsilon) = \mathbb{R}^2 \qquad \forall\, x \in \mathbb{R}^2, \quad \varepsilon > 0.$$

By virtue of chain (22.8), attainable sets $\mathcal{A}_\mu^1(t)$ of the initial system (22.14) satisfy the equality

$$\mathrm{cl}(\mathcal{A}_\mu^1(\varepsilon)) = \mathbb{R}^3 \qquad \forall\, \mu \in \mathbb{R}^3, \quad \varepsilon > 0.$$

Since the vector l is in general position, the 3-dimensional system (22.14) has a full rank (see Proposition 6.1), thus it is completely controllable for arbitrarily small time:

$$\mathcal{A}_\mu^1(\varepsilon) = \mathbb{R}^3 \qquad \forall\, \mu \in \mathbb{R}^3, \quad \varepsilon > 0.$$

Chapter 23

Curvature

23.1 Curvature of 2-dimensional systems

Consider a control system of the form

$$\dot{q} = f_u(q), \qquad q \in M, \quad u \in U, \tag{23.1}$$

where

$$\dim M = 2, \qquad U = \mathbb{R} \text{ or } S^1.$$

We suppose that the right-hand side $f_u(q)$ is smooth in (u, q). A well-known example of such a system is given by a two-dimensional Riemannian problem: locally, such a problem determines a control system

$$\dot{q} = \cos u \, f_1(q) + \sin u \, f_2(q), \qquad q \in M, \quad u \in S^1,$$

where $f_1,\ f_2$ is a local orthonormal frame of the Riemannian structure. For control systems (23.1), we obtain a feedback-invariant form of Jacobi equation and construct the main feedback invariant — curvature (in the Riemannian case this invariant coincides with Gaussian curvature). We prove comparison theorem for conjugate points similar to those in Riemannian geometry.

We assume that the curve of admissible velocities of control system (23.1) satisfies the following regularity conditions:

$$\begin{gathered} f_u(q) \ \wedge \ \frac{\partial f_u(q)}{\partial u} \neq 0, \\ \frac{\partial f_u(q)}{\partial u} \ \wedge \ \frac{\partial^2 f_u(q)}{\partial u^2} \neq 0, \qquad q \in M, \quad u \in U. \end{gathered} \tag{23.2}$$

Condition (23.2) means that the curve $\{f_u(q) \mid u \in U\} \subset T_qM$ is strongly convex, it implies strong Legendre condition for extremals of system (23.1).

Introduce the linear on fibers control-dependent Hamiltonian

$$h_u(\lambda) = \langle \lambda, f_u(q) \rangle$$

and the maximized Hamiltonian

$$H(\lambda) = \max_{u \in U} h_u(\lambda). \tag{23.3}$$

We suppose that $H(\lambda)$ is defined in a domain in T^*M under consideration. Moreover, we assume that for any λ in this domain maximum in (23.3) is attained for a unique $u \in U$, this means that any line of support touches the curve of admissible velocities at a unique point. Then the convexity condition (23.2) implies that $H(\lambda)$ is smooth in this domain and strongly convex on fibers of T^*M. Moreover, H is homogeneous of order one on fibers, thus we restrict to the level surface

$$\mathcal{H} = H^{-1}(1) \subset T^*M.$$

Denote the intersection with a fiber

$$\mathcal{H}_q = \mathcal{H} \cap T_q^*M.$$

23.1.1 Moving frame

We construct a feedback-invariant moving frame on the 3-dimensional manifold $\mathcal{H}$ in order to write Jacobi equation in this frame. Notice that the maximized Hamiltonian H is feedback-invariant since it depends on the whole admissible velocity curve $f_U(q)$, not on its parametrization by u. Thus the level surface $\mathcal{H}$ and the fiber $\mathcal{H}_q$ are also feedback-invariant.

We start from a vertical field tangent to the curve $\mathcal{H}_q$. Introduce polar coordinates in a fiber:

$$p = (r\cos\varphi, r\sin\varphi) \in T_q^*M,$$

then $\mathcal{H}_q$ is parametrized by angle φ:

$$\mathcal{H}_q = \{p = p(\varphi)\}.$$

Since the curve $\mathcal{H}_q$ does not pass through the origin: $p(\varphi) \neq 0$, it follows that

$$p(\varphi) \ \wedge \ \frac{dp}{d\varphi}(\varphi) \neq 0. \tag{23.4}$$

Decompose the second derivative in the frame p, $\dfrac{dp}{d\varphi}$:

$$\frac{d^2p}{d\varphi^2}(\varphi) = a(\varphi)p(\varphi) + b(\varphi)\frac{dp}{d\varphi}(\varphi).$$

The curve $\mathcal{H}_q$ is strongly convex, thus

$$a(\varphi) < 0.$$

A change of parameter $\theta = \theta(\varphi)$ gives

$$\frac{d^2 p}{d\,\theta^2} = a(\varphi)\left(\frac{d\,\varphi}{d\,\theta}\right)^2 p(\theta) + \tilde{b}(\theta)\frac{d\,p}{d\,\theta}(\theta),$$

thus there exists a unique (up to translations and orientation) parameter θ on the curve $\mathcal{H}_q$ such that

$$\frac{d^2 p}{d\,\theta^2} = -p(\theta) + b(\theta)\frac{d\,p}{d\,\theta}(\theta).$$

We fix such a parameter θ and define the corresponding vertical vector field on $\mathcal{H}$:

$$v = \frac{\partial}{\partial\,\theta}.$$

In invariant terms, v is a unique (up to multiplication by ± 1) vertical field on $\mathcal{H}$ such that

$$L_v^2 s = -s + b\,L_v s, \tag{23.5}$$

where $s = p\,dq$ is the tautological form on T^*M.

We define the moving frame on $\mathcal{H}$ as follows:

$$V_1 = v, \quad V_2 = [v, \vec{H}], \quad V_3 = \vec{H}.$$

Notice that these vector fields are linearly independent since v is vertical and the other two fields have linearly independent horizontal parts:

$$\pi_* \vec{H} = f,$$
$$\pi_*[v, \vec{H}] = \frac{\partial\, f_u}{\partial\, u}\frac{d\,u}{d\,\theta}, \qquad \frac{d\,u}{d\,\theta} \neq 0.$$

Here we denote by $u(\theta)$ the maximizing control on $\mathcal{H}_q$:

$$\langle p(\theta), f_{u(\theta)}\rangle \geq \langle p(\theta), f_u\rangle, \qquad u \in U.$$

Differentiating the identity

$$\left\langle p(\theta), \left.\frac{\partial\, f_u}{\partial\, u}\right|_{u(\theta)}\right\rangle \equiv 0$$

w.r.t. θ, we obtain $\dfrac{d\,u}{d\,\theta} \neq 0$.

In order to write Jacobi equation along an extremal λ_t, we require Lie brackets of the Hamiltonian field $\vec{H}$ with the vector fields of the frame:

$$[\vec{H}, V_1] = -V_2,$$
$$[\vec{H}, V_2] = \,?,$$
$$[\vec{H}, V_3] = 0.$$

The missing second bracket is given by the following proposition.

Theorem 23.1.

$$[\vec{H}, [\vec{H}, v]] = -\kappa v. \tag{23.6}$$

The function $\kappa = \kappa(\lambda)$, $\lambda \in \mathcal{H}$, is called the *curvature* of the two-dimensional control system (23.1). The Hamiltonian field $\vec{H}$ is feedback-invariant, and the field v is feedback-invariant up to multiplication by ± 1. Thus the curvature κ is a feedback invariant of system (23.1).

Proof. The parameter θ provides an identification

$$\mathcal{H} \cong \{\theta\} \times M, \tag{23.7}$$

thus tangent spaces to $\mathcal{H}$ decompose into direct sum of the horizontal and vertical subspaces. By duality, any differential form on $\mathcal{H}$ has a horizontal and vertical part. Notice that trivialization (23.7) is not feedback invariant since the choice of the section $\theta = 0$ is arbitrary, thus the form $d\theta$ and the property of a subspace to be horizontal are not feedback-invariant.

For brevity, we denote in this proof

$$s = s|_{\mathcal{H}},$$

a horizontal form on $\mathcal{H}$. Denote the Lie derivative:

$$L_{\frac{\partial}{\partial \theta}} = {}'$$

and consider the following coframe on $\mathcal{H}$:

$$d\theta, \quad s, \quad s'. \tag{23.8}$$

It is easy to see that these forms are linearly independent: $d\theta$ is vertical, while the horizontal forms s, s' are linearly independent by (23.4). Now we construct a frame on $\mathcal{H}$ dual to coframe (23.8).

Decompose $\vec{H}$ into the horizontal and vertical parts:

$$\vec{H} = \underbrace{Y}_{\text{horizontal}} + \underbrace{\alpha \frac{\partial}{\partial \theta}}_{\text{vertical}}, \qquad \alpha = \alpha(\theta, q). \tag{23.9}$$

We prove that the fields

$$\frac{\partial}{\partial \theta}, \quad Y, \quad Y' = \left[\frac{\partial}{\partial \theta}, Y\right]$$

give a frame dual to coframe (23.8). We have to show only that the pair of horizontal fields Y, Y' is dual to the pair of horizontal forms s, s'. First,

$$\langle s_\lambda, Y \rangle = \langle s_\lambda, \vec{H} \rangle = \langle \lambda, f_u \rangle = H(\lambda) = 1.$$

Further,

$$\langle s_\lambda, Y' \rangle = \langle s_\lambda, \vec{H}' \rangle = \left\langle \lambda, \frac{\partial f_u}{\partial \theta} \right\rangle = \underbrace{\left\langle \lambda, \frac{\partial f_u}{\partial u} \right\rangle}_{=0} \frac{d\,u}{d\,\theta} = 0.$$

Consequently,

$$0 = \langle s, Y \rangle' = \langle s', Y \rangle + \langle s, Y' \rangle,$$

i.e.,

$$\langle s', Y \rangle = 0.$$

Finally,

$$0 = \langle s', Y \rangle' = \langle s'', Y \rangle + \langle s', Y' \rangle.$$

Equality (23.5) can be written as $s'' = -s + bs'$, thus

$$\langle s', Y' \rangle = -\langle s'', Y \rangle = \langle s - bs', Y \rangle = 1.$$

So we proved that the frame

$$\frac{\partial}{\partial\,\theta},\ Y,\ Y' \in \operatorname{Vec} \mathcal{H}$$

is dual to the coframe

$$d\theta,\ s,\ s' \in \Lambda^1(\mathcal{H}).$$

We complete the proof of this theorem computing the bracket $[\vec{H}, [\vec{H}, v]]$ using these frames.

First consider the standard symplectic form:

$$\sigma|_{\mathcal{H}} = d\,(\,s|_{\mathcal{H}}) = ds = d\theta \wedge s' + d_q s,$$

where $d_q s$ is the differential of the form s w.r.t. horizontal coordinates. The horizontal 2-form $d_q s$ decomposes:

$$d_q s = c\, s \wedge s', \qquad c = c(\theta, q),$$

thus

$$\sigma|_{\mathcal{H}} = d\theta \wedge s' + cs \wedge s'.$$

Since

$$i_{\vec{H}}\ \sigma|_{\mathcal{H}} = 0,$$

then

$$\begin{aligned} \sigma|_{\mathcal{H}}\,(\vec{H}, \cdot\,) &= \sigma|_{\mathcal{H}}\,(Y + \alpha \frac{\partial}{\partial\,\theta}, \cdot\,) \\ &= \alpha s' - \langle s', Y \rangle\, d\theta + c\,\langle s, Y \rangle s' - c\,\langle s', Y \rangle s \\ &= \alpha s' + cs' = 0, \end{aligned}$$

i.e., $\alpha = -c$, thus

$$\vec{H} = Y - c\frac{\partial}{\partial\,\theta}.$$

Now we can compute the required Lie bracket.

$$\vec{H}' = \left[\frac{\partial}{\partial\,\theta}, \vec{H}\right] = Y' - c'\frac{\partial}{\partial\,\theta},$$

consequently,

$$\begin{aligned}\left[\vec{H}, \left[\vec{H}, \frac{\partial}{\partial\,\theta}\right]\right] &= \left[\vec{H}, -\vec{H}'\right] = \left[Y - c\frac{\partial}{\partial\,\theta}, -Y' + c'\frac{\partial}{\partial\,\theta}\right] \\ &= \underbrace{\left(\vec{H}c' - \vec{H}'c\right)\frac{\partial}{\partial\,\theta}}_{\text{vertical part}} + \underbrace{[Y', Y] + cY'' - c'Y'}_{\text{horizontal part}}.\end{aligned}$$

In order to complete the proof, we have to show that the horizontal part of the bracket $[\vec{H}, [\vec{H}, \frac{\partial}{\partial\,\theta}]]$ vanishes.

The equality $s'' = -s + bs'$ implies, by duality of the frames Y, Y' and s, s', that

$$Y'' = -Y - bY'.$$

Further,

$$\begin{aligned} ds &= d\theta \wedge s' + cs \wedge s', \\ d(s') &= (ds)' = d\theta \wedge s'' + c's \wedge s' + cs \wedge s'' \\ &= -d\theta \wedge s - b\,d\theta \wedge s' + (c' + cb)s \wedge s', \end{aligned}$$

and we can compute the bracket $[Y', Y]$ using duality of the frames and Proposition 18.1:

$$[Y', Y] = cY + (c' + cb)Y'.$$

Summing up, the horizontal part of the field $[\vec{H}, [\vec{H}, v]]$ is

$$[Y', Y] + cY'' - c'Y' = cY + (c' + cb)Y' - cY - cbY' - c'Y' = 0.$$

We proved that

$$\left[\vec{H}, \left[\vec{H}, \frac{\partial}{\partial\,\theta}\right]\right] = -\kappa\frac{\partial}{\partial\,\theta},$$

where the curvature has the form

$$\kappa = -\vec{H}c' + \vec{H}'c.$$

□

Remark. Recall that the vertical vector field v that satisfies (23.5) is unique (up to a factor ± 1). On the other hand, the vertical field v that satisfies (23.6) is unique, up to a factor constant along trajectories of $\vec{H}$ (so this factor does not affect κ). Consequently, any vertical vector field v for which an equality of the form (23.6) holds can be used for computation of curvature κ.

So now we know all brackets of the Hamiltonian vector field $X = \vec{H}$ with the vector fields of the frame V_1, V_2, V_3:

$$[\vec{H}, V_1] = -V_2, \tag{23.10}$$

$$[\vec{H}, V_2] = \kappa V_1, \tag{23.11}$$

$$[\vec{H}, V_3] = 0. \tag{23.12}$$

23.1.2 Jacobi equation in moving frame

We apply the moving frame constructed to derive an ODE on conjugate time of our two-dimensional system — Jacobi equation in the moving frame.

As in Chapter 21, consider Jacobi equation along a regular extremal λ_t, $t \in [0, t_1]$, of the two-dimensional system (23.1):

$$\dot{\eta} = \vec{b}_t(\eta), \qquad \eta \in \Sigma = T_{\lambda_0}(T^*M),$$

and its flow

$$B_t = \overrightarrow{\exp} \int_0^t \vec{b}_\tau \, d\tau.$$

Recall that $\Pi_0 = T_{\lambda_0}(T^*_{q_0}M)$ is the vertical subspace in Σ and $C_t \subset \Pi_0$ is the subspace of constant vertical solutions to Jacobi equation at $[0, t]$, see (21.30). The intersection $B_t\Pi_0 \cap \Pi_0$ always contains the subspace C_t. An instant $t \in (0, t_1]$ is a conjugate time for the extremal λ_t iff that intersection is greater than C_t:

$$B_t\Pi_0 \cap \Pi_0 \neq C_t.$$

In order to complete the frame V_1, V_2, V_3 to a basis in $T_{\lambda_0}(T^*M)$, consider a vector field transversal to $\mathcal{H}$ — the vertical Euler field $E \in \mathrm{Vec}(T^*M)$ with the flow

$$\lambda \circ e^{tE} = e^t \cdot \lambda, \qquad \lambda \in T^*M, \quad t \in \mathbb{R}.$$

In coordinates (p, q) on T^*M, this field reads

$$E = p\frac{\partial}{\partial p}.$$

The vector fields V_1, V_2, V_3, E form a basis in $T_\lambda(T^*M)$, $\lambda \in \mathcal{H}$. The fields $V_1 = \frac{\partial}{\partial \theta}$ and E are vertical:

$$\Pi_0 = \mathrm{span}(V_1(\lambda_0), E(\lambda_0)).$$

To compute the constant vertical subspace C_t, evaluate the action of the flow B_t on these fields. In the proof of Theorem 21.3, we decomposed the flow of Jacobi equation:

$$B_t(\eta) = (P_t^*)_* \, e_*^{t\vec{H}}(\eta).$$

Thus

$$B_t E(\lambda_0) = (P_t^*)_* \, e_*^{t\vec{H}} E(\lambda_0).$$

The Hamiltonian H is homogeneous of order one on fibers, consequently the flow of $\vec{H}$ is homogeneous as well:

$$(k\lambda) \circ e^{t\vec{H}} = k\left(\lambda \circ e^{t\vec{H}}\right), \qquad k > 0,$$

and the fields $\vec{H}$ and E commute. That is, the Hamiltonian vector field $\vec{H}$ preserves the vertical Euler field E. Further, the flow P_t^* is linear on fibers, thus it also preserves the field E. Summing up, the vector $E(\lambda_0)$ is invariant under the action of the flow of Jacobi equation, i.e.,

$$\mathbb{R}E(\lambda_0) \subset C_t.$$

It is easy to see that this inclusion is in fact an equality. Indeed, in view of bracket (23.10),

$$e_*^{t\vec{H}} V_1(\lambda_0) = \lambda_t \circ e^{-t \operatorname{ad} \vec{H}} V_1 = \lambda_t \circ (V_1 + tV_2 + o(t)) \notin T_{\lambda_t}(T^*_{q(t)}M),$$

thus

$$B_t V_1(\lambda_0) \notin \Pi_0$$

for small $t > 0$. This means that

$$C_t = \mathbb{R}E(\lambda_0), \qquad t \in (0, t_1].$$

Thus an instant t is a conjugate time iff

$$B_t \Pi_0 \cap \Pi_0 \neq \mathbb{R}E(\lambda_0),$$

i.e.,

$$e_*^{t\vec{H}} V_1(\lambda_0) \in \mathbb{R}V_1(\lambda_t),$$

or, equivalently,

$$\lambda_0 \circ e^{t \operatorname{ad} \vec{H}} V_1 \in \mathbb{R}(\lambda_0 \circ V_1). \tag{23.13}$$

Now we describe the action of the flow of a vector field on a moving frame.

Lemma 23.1. *Let N be a smooth manifold,* $\dim N = m$, *and let vector fields $V_1, \dots, V_m \in \operatorname{Vec} N$ form a moving frame on N. Take a vector field $X \in \operatorname{Vec} N$. Let the operator* $\operatorname{ad} X$ *have a matrix $A = (a_{ij})$ in this frame:*

$$(\operatorname{ad} X)\, V_j = \sum_{i=1}^{m} a_{ij} V_i, \qquad a_{ij} \in C^\infty(N).$$

Then the matrix $\Gamma(t) = (\gamma_{ij}(t))$ of the operator $e^{t \operatorname{ad} \vec{H}}$ in the moving frame:

$$e^{t \operatorname{ad} X}\, V_j = \sum_{i=1}^{m} \gamma_{ij}(t) V_i, \qquad \gamma_{ij}(t) \in C^\infty(N), \tag{23.14}$$

is the solution to the following Cauchy problem:

$$\dot{\Gamma}(t) = \Gamma(t)A(t), \tag{23.15}$$

$$\Gamma(0) = \mathrm{Id}, \tag{23.16}$$

where $A(t) = (e^{tX} a_{ij})$.

Proof. Initial condition (23.16) is obvious. In order to derive the matrix equation (23.15), we differentiate identity (23.14) w.r.t. t:

$$\sum_{i=1}^{m} \dot{\gamma}_{ij}(t) V_i = e^{t\,\mathrm{ad}\,X}[X, V_j] = e^{t\,\mathrm{ad}\,X}\left(\sum_{k=1}^{m} a_{kj} V_k\right) = \sum_{k=1}^{m} \left(e^{tX} a_{kj}\right) e^{t\,\mathrm{ad}\,X} V_k$$
$$= \sum_{k,i=1}^{m} \left(e^{tX} a_{kj}\right) \gamma_{ik} V_i,$$

and the ODE follows. □

In view of inclusion (23.13), an instant t is a conjugate time iff the coefficients in the decomposition

$$\lambda_0 \circ e^{t\,\mathrm{ad}\,\vec{H}} V_j = \sum_{i=1}^{3} \gamma_{ij}(t)(\lambda_0 \circ V_i)$$

satisfy the equalities:

$$\gamma_{21}(t) = \gamma_{31}(t) = 0.$$

By the previous lemma, the matrix $\Gamma(t) = (\gamma_{ij}(t))$ is the solution to Cauchy problem (23.15), (23.16) with the matrix

$$A(t) = \begin{pmatrix} 0 & \kappa_t & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \kappa_t = \kappa(\lambda_t),$$

see Lie bracket relations (23.10)–(23.12).

Summing up, an instant $t \in (0, t_1]$ is a conjugate time iff the solutions to the Cauchy problems

$$\begin{cases} \dot{\gamma}_{21} = -\gamma_{22}, \\ \dot{\gamma}_{22} = \kappa_t \gamma_{21}, \end{cases} \qquad \gamma_{21}(0) = 0, \quad \gamma_{22}(0) = 1$$

and

$$\begin{cases} \dot{\gamma}_{31} = -\gamma_{32}, \\ \dot{\gamma}_{32} = \kappa_t \gamma_{31}, \end{cases} \qquad \gamma_{31}(0) = 0, \quad \gamma_{32}(0) = 0$$

satisfy the equalities

$$\gamma_{21}(t) = \gamma_{31}(t) = 0.$$

But Cauchy problem for γ_{31}, γ_{32} has only trivial solution. Thus for a conjugate time t, we obtain the linear nonautonomous system for $(x_1, x_2) = (\gamma_{21}, \gamma_{22})$:

$$\begin{cases} \dot{x}_1 = -x_2, \\ \dot{x}_2 = \kappa_t x_1, \end{cases} \qquad x_1(0) = x_1(t) = 0. \tag{23.17}$$

We call system (23.17), or, equivalently, the second order ODE

$$\ddot{x} + \kappa_t x = 0, \qquad x(0) = x(t) = 0, \tag{23.18}$$

Jacobi equation for system (23.1) *in the moving frame.* We proved the following statement.

Theorem 23.2. *An instant $t \in (0, t_1]$ is a conjugate time for the two-dimensional system* (23.1) *iff there exists a nontrivial solution to boundary problem* (23.18).

Sturm's comparison theorem for second order ODEs (see e.g. [5]) implies the following comparison theorem for conjugate points.

Theorem 23.3. (1) *If $\kappa < C^2$ for some $C > 0$ along an extremal λ_t, then there are no conjugate points at the time segment $[0, \frac{\pi}{C}]$. In particular, if $\kappa \leq 0$ along λ_t, then there are no conjugate points.*

(2) *If $\kappa \geq C^2$ along λ_t, then there is a conjugate point at the segment $[0, \frac{\pi}{C}]$.*

Example 23.1. Consider the control system corresponding to a Riemannian problem on a 2-dimensional manifold M:

$$\dot{q} = \cos u \, f_1(q) + \sin u \, f_2(q), \qquad q \in M, \quad u \in S^1,$$

where f_1, f_2 is an orthonormal frame of the Riemannian structure $\langle \cdot, \cdot \rangle$:

$$\langle f_i, f_j \rangle = \delta_{ij}, \qquad i, \ j = 1, \ 2.$$

In this case, κ is the *Gaussian curvature* of the Riemannian manifold M, and it is evaluated as follows:

$$\kappa = -c_1^2 - c_2^2 + f_1 c_2 - f_2 c_1,$$

where c_i are structure constants of the frame f_1, f_2: $[f_1, f_2] = c_1 f_1 + c_2 f_2$. We prove this formula for κ in Chapter 24.

For the Riemannian problem, the curvature $\kappa = \kappa(q)$ depends only on the base point $q \in M$, not on the coordinate θ in the fiber. Generally, this is not the case: the curvature is a function of $(q, \theta) \in \mathcal{H}$.

Optimality conditions in terms of conjugate points obtained in Chapter 21 can easily be applied to the two-dimensional system (23.1) under consideration.

Assume first that $t_c \in (0, t_1)$ is a conjugate time for an extremal λ_t, $t \in [0, t_1]$, of system (23.1). We verify hypotheses of Proposition 21.2. Condition (23.2) implies that the extremal is regular. The corresponding control $\tilde{u}$ has corank one since the Lagrange multiplier λ_t is uniquely determined by PMP (up to a scalar factor). Further, Jacobi equation cannot have solutions of form (21.28): if this were the case, Jacobi equation in the moving frame $\ddot{x} + \kappa_t x = 0$ would have a nontrivial solution with the terminal conditions $x(t_c) = \dot{x}(t_c) = 0$, which is impossible. Summing up, the extremal λ_t satisfies hypotheses of Proposition 21.2, and alternative (1) of this proposition is not realized. Thus the corresponding extremal trajectory is not locally geometrically optimal.

If the segment $[0, t_1]$ does not contain conjugate points, then by Theorem 21.4 the corresponding extremal trajectory is time-optimal compared with all other admissible trajectories sufficiently close in M.

23.2 Curvature of 3-dimensional affine systems

In this section we consider affine in control 3-dimensional systems:

$$\begin{aligned} &\dot{q} = f_0(q) + u f_1(q), \qquad u \in \mathbb{R}, \quad q \in M, \\ &\dim M = 3. \end{aligned} \tag{23.19}$$

We reduce such a system to a 2-dimensional one as in Chapter 22 and compute the curvature of the 2-dimensional system obtained — a feedback invariant of system (23.19).

We assume that the following regularity conditions hold on M:

$$f_0 \wedge f_1 \wedge [f_0, f_1] \neq 0, \tag{23.20}$$

$$f_1 \wedge [f_0, f_1] \wedge [f_1, [f_0, f_1]] \neq 0. \tag{23.21}$$

Any extremal λ_t of the affine in control system (23.19) is totally singular, it satisfies the equality

$$h_1(\lambda_t) = \langle \lambda_t, f_1 \rangle \equiv 0, \tag{23.22}$$

and the corresponding extremal control cannot be found immediately from this equality. Differentiation of (23.22) w.r.t. t yields

$$h_{01}(\lambda_t) = \langle \lambda_t, [f_0, f_1] \rangle \equiv 0,$$

and one more differentiation leads to an equality containing control:

$$h_{001}(\lambda_t) + u(t) h_{101}(\lambda_t) = \langle \lambda_t, [f_0, [f_0, f_1]] \rangle + u(t) \langle \lambda_t, [f_1, [f_0, f_1]] \rangle \equiv 0.$$

Then the singular control is uniquely determined:

$$u = \tilde{u}(q) = -\frac{h_{001}(\lambda)}{h_{101}(\lambda)}, \qquad h_1(\lambda) = h_{01}(\lambda) = 0.$$

We apply a feedback transformation to system (23.19):

$$u \mapsto u - \tilde{u}(q).$$

This transformation affects the field f_0, but preserves regularity conditions (23.20), (23.21). After this transformation the singular control is

$$u = 0.$$

in other words,

$$\lambda f_1 = \lambda[f_0, f_1] = 0 \quad \Rightarrow \quad \lambda[f_0, [f_0, f_1]] = 0.$$

So we assume below that

$$[f_0, [f_0, f_1]] \in \operatorname{span}(f_1, [f_0, f_1]). \tag{23.23}$$

Assume that the field f_1 is nonrecurrent. In a tubular neighborhood of a trajectory of the field f_1, consider the reduction of the three-dimensional system (23.19):

$$\frac{d\tilde{q}}{dt} = e^{w \operatorname{ad} f_1} f_0(\tilde{q}), \qquad w \in \mathbb{R}, \quad \tilde{q} \in \widetilde{M} = M/e^{\mathbb{R}f_1}. \tag{23.24}$$

This system has the same conjugate points as the initial one (23.19). If system (23.24) has no conjugate points, then the corresponding singular trajectory of system (23.19) is locally strongly geometrically optimal, i.e., comes, locally, to the boundary of attainable set.

Describe the cotangent bundle of the quotient $\widetilde{M}$. A tangent space to $\widetilde{M}$ consists of tangent vectors to M modulo f_1:

$$\begin{aligned} &T_{\tilde{q}}\widetilde{M} \cong T_q M/\mathbb{R}f_1(q), \\ &q \in M, \qquad \tilde{q} = q \circ e^{\mathbb{R}f_1} \in \widetilde{M}, \end{aligned} \tag{23.25}$$

identification (23.25) is given by the mapping

$$\begin{aligned} &v \mapsto \tilde{v}, \qquad v \in T_q M, \quad \tilde{v} \in T_{\tilde{q}}\widetilde{M}, \\ &v = \left.\frac{d}{dt}\right|_{t=0} q(t), \qquad \tilde{v} = \left.\frac{d}{dt}\right|_{t=0} \tilde{q}(t). \end{aligned}$$

Thus a cotangent space to $\widetilde{M}$ consists of covectors on M orthogonal to f_1:

$$\begin{aligned} &T^*_{\tilde{q}}\widetilde{M} \cong T^*_q M \cap \{h_1 = 0\}, \\ &\lambda \mapsto \tilde{\lambda}, \qquad \lambda \in T^*_q M \cap \{h_1 = 0\}, \quad \tilde{\lambda} \in T^*_{\tilde{q}}\widetilde{M}, \\ &\langle \tilde{\lambda}, \tilde{v} \rangle = \langle \lambda, v \rangle, \qquad v \in T_q M, \quad \tilde{v} \in T_{\tilde{q}}\widetilde{M}. \end{aligned}$$

Taking into account that the field f_1 is the projection of the Hamiltonian field $\vec{h}_1$, it is easy to see that

$$T^*\widetilde{M} \cong \{h_1 = 0\}/e^{\mathbb{R}\vec{h}_1},$$

where the mapping $\lambda \mapsto \tilde{\lambda}$ is defined above (exercise: show that $\tilde{\lambda}_1 = \tilde{\lambda}_2 \Leftrightarrow \lambda_2 \in \lambda_1 \circ e^{\mathbb{R}\vec{h}_1}$). Summing up, cotangent bundle to the quotient $\widetilde{M}$ is obtained from T^*M via Hamiltonian reduction by $\vec{h}_1$: restriction to the level surface of h_1 with subsequent factorization by the flow of $\vec{h}_1$.

Further, regularity condition (23.21) implies that the field $\vec{h}_1$ is transversal to the level surface $\{h_1 = h_{01} = 0\}$, so this level surface gives another realization of the cotangent bundle to the quotient:

$$T^*\widetilde{M} \cong \{h_1 = h_{01} = 0\}.$$

In this realization, $\vec{h}_0$ is the Hamiltonian field corresponding to the maximized Hamiltonian — generator of extremals ($\vec{H}$ in Section 23.1). The level surface of the maximized Hamiltonian ($\mathcal{H}$ in Section 23.1) realizes as the submanifold

$$\{h_1 = h_{01} = 0,\ h_0 = 1\} \subset T^*M.$$

Via the canonical projection $\pi : T^*M \to M$, this submanifold can be identified with M, so the level surface $\mathcal{H}$ of Section 23.1 realizes now as M. We use this realization to compute curvature of the three-dimensional system (23.19) as the curvature κ of its two-dimensional reduction (23.24).

The Hamiltonian field $\vec{H}$ of Section 23.1 is now f_0, and f_1 is a vertical field. It remains to normalize f_1, i.e., to find a vertical field af_1, $a \in C^\infty(M)$, such that

$$[f_0, [f_0, af_1]] = -\kappa a f_1, \tag{23.26}$$

see (23.6). The triple

$$f_0, \quad f_1, \quad f_2 = [f_0, f_1]$$

forms a moving frame on M, consider the structure constants of this frame:

$$[f_i, f_j] = \sum_{k=0}^{2} c_{ji}^k f_k, \qquad i,\ j = 0,\ 1,\ 2.$$

Notice that inclusion (23.23) obtained after preliminary feedback transformation reads now as $c_{02}^0 = 0$. That is why

$$[f_0, [f_0, f_1]] = -c_{02}^1 f_1 - c_{02}^2 [f_0, f_1].$$

Now we can find the normalizing factor a for f_1 such that (23.26) be satisfied. We have

$$\begin{aligned}[f_0, [f_0, af_1]] &= [f_0, (f_0 a) + a[f_0, f_1]] = (f_0^2 a) f_1 + 2(f_0 a)[f_0, f_1] + a[f_0, [f_0, f_1]] \\ &= (f_0^2 a - c_{02}^1 a) f_1 + (2 f_0 a - c_{01}^2)[f_0, f_1].\end{aligned}$$

Then the required function a is found from the first order PDE

$$2 f_0 a - c_{02}^2 a = 0,$$

and the curvature is computed:

$$\kappa = -\frac{f_0^2 a - c_{02}^1 a}{a}.$$

Summing up, *curvature of the affine in control 3-dimensional system* (23.19) is expressed through the structure constants as

$$\kappa = c_{02}^1 - \frac{1}{4}(c_{02}^2)^2 - \frac{1}{2} f_0 c_{02}^2,$$

a function on the state space M.

Bounds on curvature κ along a (necessarily singular) extremal of a 3-dimensional affine in control system allow one to obtain bounds on conjugate time, thus on segments where the extremal is locally optimal. Indeed, by construction, κ is the curvature of the reduced 2-dimensional system. As we know from Chapter 22, reduction transforms singular extremals into regular ones, and the initial and reduced systems have the same conjugate times. Thus Theorem 23.3 can be applied, via reduction, to the study of optimality of singular extremals of 3-dimensional affine in control systems.

Chapter 24

Rolling bodies

We apply the Orbit Theorem and Pontryagin Maximum Principle to an intrinsic geometric model of a pair of rolling rigid bodies. We solve the controllability problem: in particular, we show that the system is globally controllable iff the bodies are not isometric. We also state an optimal control problem and study its extremals.

24.1 Geometric model

Consider two solid bodies in the 3-dimensional space that roll one on another without slipping or twisting. Rather than embedding the problem into $\mathbb{R}^3$, we construct an intrinsic geometric model of the system.

Let M and $\widehat{M}$ be two-dimensional connected manifolds — surfaces of the rolling bodies. In order to measure lengths of paths in M and $\widehat{M}$, we suppose that each of these manifolds is *Riemannian*, i.e., endowed with a Riemannian structure — an inner product in the tangent space smoothly depending on the point in the manifold:

$$\langle v_1, v_2 \rangle_M, \qquad v_i \in T_x M,$$
$$\langle \widehat{v}_1, \widehat{v}_2 \rangle_{\widehat{M}}, \qquad \widehat{v}_i \in T_{\widehat{x}} \widehat{M}.$$

Moreover, we suppose that M and $\widehat{M}$ are oriented (which is natural since surfaces of solid bodies in $\mathbb{R}^3$ are oriented by the exterior normal vector).

At contact points of the bodies $x \in M$ and $\widehat{x} \in \widehat{M}$, their tangent spaces are identified by an isometry (i.e., a linear mapping preserving the Riemannian structures)

$$q : T_x M \to T_{\widehat{x}} \widehat{M},$$

We deal only with orientation-preserving isometries and omit the words "orientation-preserving" in order to simplify terminology. An isometry q is a state of the system, and the state space is the connected 5-dimensional manifold

$$Q = \{ \, q : T_x M \to T_{\widehat{x}} \widehat{M} \mid x \in M, \ \widehat{x} \in \widehat{M}, \ q \text{ an isometry} \, \}.$$

Denote the projections from Q to M and $\widehat{M}$:

$$\pi(q) = x, \quad \widehat{\pi}(q) = \widehat{x}, \quad q : T_x M \to T_{\widehat{x}} \widehat{M},$$
$$q \in Q, \quad x \in M, \quad \widehat{x} \in \widehat{M}.$$

Local coordinates on Q can be introduced as follows. Choose arbitrary local orthonormal frames e_1, e_2 on M and $\widehat{e}_1$, $\widehat{e}_2$ on $\widehat{M}$:

$$\langle e_i, e_j \rangle_M = \delta_{ij}, \qquad \langle \widehat{e}_i, \widehat{e}_j \rangle_{\widehat{M}} = \delta_{ij}, \qquad i, \ j = 1, \ 2.$$

For any contact configuration of the bodies $q \in Q$, denote by θ the angle of rotation from the frame $\widehat{e}_1$, $\widehat{e}_2$ to the frame qe_1, qe_2 at the contact point:

$$qe_1 = \cos\theta \ \widehat{e}_1 + \sin\theta \ \widehat{e}_2,$$
$$qe_2 = -\sin\theta \ \widehat{e}_1 + \cos\theta \ \widehat{e}_2.$$

Then locally points $q \in Q$ are parametrized by triples $(x, \widehat{x}, \theta)$, $x = \pi(q) \in M$, $\widehat{x} = \widehat{\pi}(q) \in \widehat{M}$, $\theta \in S^1$. Choosing local coordinates (x_1, x_2) on M and $(\widehat{x}_1, \widehat{x}_2)$ on $\widehat{M}$, we obtain local coordinates $(x_1, x_2, \widehat{x}_1, \widehat{x}_2, \theta)$ on Q.

Let $q(t) \in Q$ be a curve corresponding to a motion of the rolling bodies, then $x(t) = \pi(q(t))$ and $\widehat{x}(t) = \widehat{\pi}(q(t))$ are trajectories of the contact points in M and $\widehat{M}$ respectively. The condition of absence of slipping means that

$$q(t)\dot{x}(t) = \dot{\widehat{x}}(t), \tag{24.1}$$

and the condition of absence of twisting is geometrically formulated as follows:

$$q(t)\,(\text{vector field parallel along } x(t)) = (\text{vector field parallel along } \widehat{x}(t)). \tag{24.2}$$

Our model ignores the state constraints that correspond to admissibility of contact of the bodies embedded in $\mathbb{R}^3$. Notice although that if the surfaces M and $\widehat{M}$ have respectively positive and nonnegative Gaussian curvatures at a point, then their contact is locally admissible.

The admissibility conditions (24.1) and (24.2) imply that a curve $x(t) \in M$ determines completely the whole motion $q(t) \in Q$. That is, velocities of admissible motions determine a rank 2 distribution Δ on the 5-dimensional space Q. We show this formally and compute the distribution Δ explicitly below. Before this, we recall some basic facts of Riemannian geometry.

24.2 Two-dimensional Riemannian geometry

Let M be a 2-dimensional Riemannian manifold. We describe Riemannian geodesics, Levi-Civita connection and parallel translation on $T^*M \cong TM$.

Let $\langle \cdot , \cdot \rangle$ be the Riemannian structure and e_1, e_2 a local orthonormal frame on M.

24.2.1 Riemannian geodesics

For any fixed points $x_0,\ x_1 \in M$, we seek for the shortest curve in M connecting x_0 and x_1:

$$\begin{aligned}
&\dot{x} = u_1 e_1(x) + u_2 e_2(x), \qquad x \in M, \quad (u_1, u_2) \in \mathbb{R}^2,\\
&x(0) = x_0, \qquad x(t_1) = x_1,\\
&l = \int_0^{t_1} \langle \dot{x}, \dot{x} \rangle^{1/2}\, dt = \int_0^{t_1} (u_1^2 + u_2^2)^{1/2}\, dt \to \min .
\end{aligned}$$

In the same way as in Section 19.1, it easily follows from PMP that arc-length parametrized extremal trajectories in this problem (Riemannian geodesics) are projections of trajectories of the normal Hamiltonian field:

$$\begin{aligned}
&x(t) = \pi \circ e^{t\vec{H}}(\lambda), \qquad \lambda \in \mathcal{H} = \{H = 1/2\} \subset T^*M,\\
&H = \frac{1}{2}(h_1^2 + h_2^2),\\
&h_i(\lambda) = \langle \lambda, e_i \rangle, \qquad i = 1,\ 2.
\end{aligned}$$

The level surface $\mathcal{H}$ is a spherical bundle over M with a fiber

$$\mathcal{H}_q = \{h_1^2 + h_2^2 = 1\} \cap T_q^*M \cong S^1$$

parametrized by angle φ:

$$h_1 = \cos\varphi, \qquad h_2 = \sin\varphi.$$

Cotangent bundle of a Riemannian manifold can be identified with the tangent bundle via the Riemannian structure:

$$\begin{aligned}
&TM \cong T^*M,\\
&v \mapsto \lambda = \langle v, \cdot \rangle.
\end{aligned}$$

Then $\mathcal{H} \subset T^*M$ is identified with the spherical bundle

$$\mathcal{S} = \{v \in TM \mid \|v\| = 1\} \subset TM$$

of unit tangent vectors to M. After this identification, $e^{t\vec{H}}$ can be considered as a geodesic flow on $\mathcal{S}$.

24.2.2 Levi-Civita connection

A *connection* on the spherical bundle $\mathcal{S} \to M$ is an arbitrary horizontal distribution D:

$$\begin{aligned}
&D = \{D_v \subset T_v\mathcal{S} \mid v \in \mathcal{S}\},\\
&D_v \oplus T_v(\mathcal{S}_x) = T_v\mathcal{S}, \qquad \mathcal{S}_x = \mathcal{S} \cap T_xM.
\end{aligned}$$

Any connection D on M defines a *parallel translation* of unit tangent vectors along curves in M. Let $x(t)$, $t \in [0, t_1]$, be a curve in M, and let $v_0 \in T_{x(0)}M$ be a unit tangent vector. The curve $x(t)$ has a unique horizontal lift on $\mathcal{S}$ starting at v_0:

$$\begin{aligned} &v(t) \in \mathcal{S}, \qquad \pi \circ v(t) = x(t), \\ &\dot v(t) \in D_{v(t)}, \\ &v(0) = v_0. \end{aligned}$$

Indeed, if the curve $x(t)$ satisfies the nonautonomous ODE

$$\dot x = u_1(t)\, e_1(x) + u_2(t)\, e_2(x),$$

then its horizontal lift $v(t)$ is a solution to the lifted ODE

$$\dot v = u_1(t)\, \xi_1(v) + u_2(t)\, \xi_2(v), \tag{24.3}$$

where ξ_i are horizontal lifts of the basis fields e_i:

$$D_v = \operatorname{span}(\xi_1(v), \xi_2(v)), \qquad \pi_* \xi_i = e_i.$$

Notice that solutions of ODE (24.3) are continued to the whole time segment $[0, t_1]$ since the fibers $\mathcal{S}_x$ are compact. The vector $v(t_1)$ is the parallel translation of the vector v_0 along the curve $x(t)$.

A vector field $v(t)$ along a curve $x(t)$ is called *parallel* if it is preserved by parallel translations along $x(t)$.

Levi-Civita connection is the unique connection on the spherical bundle $\mathcal{S} \to M$ such that:

(1) velocity of a Riemannian geodesic is parallel along the geodesic (i.e., the geodesic field $\vec H$ is horizontal),

(2) parallel translation preserves angle, i.e., horizontal lifts of vector fields on the base M commute with the vector field $\frac{\partial}{\partial \varphi}$ that determines the element of length (or, equivalently, the element of angle) in the fiber $\mathcal{S}_x$.

Now we compute the Levi-Civita connection as a horizontal distribution on $\mathcal{H} \cong \mathcal{S}$. In Chapter 23 we constructed a feedback-invariant frame on the manifold $\mathcal{H}$:

$$T_\lambda \mathcal{H} = \operatorname{span}\left(\vec H, \frac{\partial}{\partial \varphi}, \vec H'\right), \qquad \vec H' = \left[\frac{\partial}{\partial \varphi}, \vec H\right].$$

For the two-dimensional Riemannian problem, decomposition (23.9) into the horizontal and vertical parts reads

$$\begin{aligned} \vec H &= Y - c\,\frac{\partial}{\partial \varphi}, \qquad & Y &= h_1 e_1 + h_2 e_2, \qquad & c &= -h_1 c_1 - h_2 c_2, \\ \vec H' &= Y' - c'\,\frac{\partial}{\partial \varphi}, \qquad & Y' &= -h_2 e_1 + h_1 e_2, \qquad & c' &= h_2 c_1 - h_1 c_2, \end{aligned}$$

where c_i are structure constants of the orthonormal frame on M:

$$[e_1, e_2] = c_1 e_1 + c_2 e_2, \qquad c_i \in C^\infty(M).$$

Notice that using the above decompositions, we can easily compute Gaussian curvature k of the Riemannian manifold M via the formula of Theorem 23.1:

$$\left[\vec{H}, \left[\vec{H}, \frac{\partial}{\partial \varphi}\right]\right] = -k \frac{\partial}{\partial \varphi}.$$

Since

$$[\vec{H}, \vec{H}'] = (c_1^2 + c_2^2 - e_1 c_2 + e_2 c_1) \frac{\partial}{\partial \varphi},$$

then

$$k = -c_1^2 - c_2^2 + e_1 c_2 - e_2 c_1. \tag{24.4}$$

A horizontal rank 2 distribution D on $\mathcal{H}$ containing the Hamiltonian field $\vec{H}$ (i.e., satisfying property (1) of Levi-Civita connections) has the form

$$D = \operatorname{span}\left(\vec{H},\ \vec{H}' + \beta \frac{\partial}{\partial \varphi}\right) \tag{24.5}$$

for some function β

$$\begin{aligned} &= \operatorname{span}\left(h_1 e_1 + h_2 e_2 - c \frac{\partial}{\partial \varphi}, -h_2 e_1 + h_1 e_2 + (-c' + \beta) \frac{\partial}{\partial \varphi}\right) \\ &= \operatorname{span}(V_1, V_2), \end{aligned}$$

where V_i are horizontal lifts of the basis vector fields on M:

$$V_1 = e_1 + (-h_1 c + h_2 c' - h_2 \beta) \frac{\partial}{\partial \varphi}, \qquad V_2 = e_2 + (-h_2 c - h_1 c' + h_1 \beta) \frac{\partial}{\partial \varphi}.$$

Property (2) of Levi-Civita connections reads

$$\left[\frac{\partial}{\partial \varphi}, V_1\right] = \left[\frac{\partial}{\partial \varphi}, V_2\right] = 0,$$

and straightforward computations yield that the function β in (24.5) vanishes identically. Summing up, the Levi-Civita connection on M has the form

$$D = \operatorname{span}\left(\vec{H}, \vec{H}'\right).$$

The 1-form of the connection D:

$$\omega \in \Lambda^1(\mathcal{H}), \qquad D = \operatorname{Ker} \omega,$$

reads

$$\omega = c_1 \omega_1 + c_2 \omega_2 - d\varphi,$$

where (ω_1, ω_2) is the dual coframe to (e_1, e_2):

$$\omega_i \in \Lambda^1(M), \qquad \langle \omega_i, e_j \rangle = \delta_{ij}, \quad i,\ j = 1,\ 2.$$

24.3 Admissible velocities

We return to the rolling bodies problem and write down admissibility conditions (24.1), (24.2) for a curve $q(t) \in Q$ as restrictions on velocity $\dot{q}(t)$. Decompose velocities of the contact curves in M and $\widehat{M}$ in the orthonormal frames:

$$\dot{x} = a_1 \, e_1(x) + a_2 \, e_2(x), \tag{24.6}$$

$$\dot{\widehat{x}} = \widehat{a}_1 \, \widehat{e}_1(\widehat{x}) + \widehat{a}_2 \, \widehat{e}_2(\widehat{x}). \tag{24.7}$$

Then the nonslipping condition (24.1) reads:

$$\widehat{a}_1 = a_1 \cos\theta - a_2 \sin\theta, \qquad \widehat{a}_2 = a_1 \sin\theta + a_2 \cos\theta. \tag{24.8}$$

Now we consider the nontwisting condition (24.2). Denote the structure constants of the frames:

$$[e_1, e_2] = c_1 e_1 + c_2 e_2, \qquad c_i \in C^\infty(M),$$

$$[\widehat{e}_1, \widehat{e}_2] = \widehat{c}_1 \widehat{e}_1 + \widehat{c}_2 \widehat{e}_2, \qquad \widehat{c}_i \in C^\infty(\widehat{M}).$$

Let $\tilde{q} : T_x^* M \to T_{\widehat{x}}^* \widehat{M}$ be the mapping induced by the isometry q via identification of tangent and cotangent spaces:

$$\tilde{q}\,\omega_1 = \cos\theta \, \widehat{\omega}_1 + \sin\theta \, \widehat{\omega}_2,$$

$$\tilde{q}\,\omega_2 = -\sin\theta \, \widehat{\omega}_1 + \cos\theta \, \widehat{\omega}_2.$$

In the cotangent bundle, the nontwisting condition means that if

$$\lambda(t) = (x(t), \varphi(t)) \in \mathcal{H}$$

is a parallel covector field along a curve $x(t) \in M$, then

$$\widehat{\lambda}(t) = \tilde{q}(t)\lambda(t) = (\widehat{x}(t), \widehat{\varphi}(t)) \in \widehat{\mathcal{H}}$$

is a parallel covector field along the curve $\widehat{x}(t) \in \widehat{M}$.

Since the isometry $q(t)$ rotates the tangent spaces at the angle $\theta(t)$, then the mapping $\tilde{q}(t)$ rotates the cotangent spaces at the same angle: $\widehat{\varphi}(t) = \varphi(t) + \theta(t)$, thus

$$\dot{\theta}(t) = \dot{\widehat{\varphi}}(t) - \dot{\varphi}(t). \tag{24.9}$$

A covector field $\lambda(t)$ is parallel along the curve in the base $x(t)$ iff $\dot{\lambda} \in \operatorname{Ker}\omega$, i.e.,

$$\dot{\varphi} = \langle c_1 \omega_1 + c_2 \omega_2, \dot{x} \rangle = c_1 a_1 + c_2 a_2.$$

Similarly, $\widehat{\lambda}(t)$ is parallel along $\widehat{x}(t)$ iff

$$\begin{aligned} \dot{\widehat{\varphi}} &= \left\langle \widehat{c}_1 \, \widehat{\omega}_1 + \widehat{c}_2 \, \widehat{\omega}_2, \dot{\widehat{x}} \right\rangle = \widehat{c}_1 \, \widehat{a}_1 + \widehat{c}_2 \, \widehat{a}_2 \\ &= a_1(\widehat{c}_1 \cos\theta + \widehat{c}_2 \sin\theta) + a_2(-\widehat{c}_1 \sin\theta + \widehat{c}_2 \cos\theta). \end{aligned}$$

In view of (24.9), the nontwisting condition reads

$$\begin{aligned}\dot\theta &= \widehat{c}_1\widehat{a}_1 + \widehat{c}_2\widehat{a}_2 - (c_1a_1 + c_2a_2)\\ &= a_1(-c_1 + \widehat{c}_1\cos\theta + \widehat{c}_2\sin\theta) + a_2(-c_2 - \widehat{c}_1\sin\theta + \widehat{c}_2\cos\theta).\end{aligned} \tag{24.10}$$

Summing up, admissibility conditions (24.1) and (24.2) for rolling bodies determine constraints (24.8) and (24.10) along contact curves (24.6), (24.7), i.e., a rank two distribution Δ on Q spanned locally by the vector fields

$$X_1 = e_1 + \cos\theta\,\widehat{e}_1 + \sin\theta\,\widehat{e}_2 + (-c_1 + \widehat{c}_1\cos\theta + \widehat{c}_2\sin\theta)\frac{\partial}{\partial\theta}, \tag{24.11}$$

$$X_2 = e_2 - \sin\theta\,\widehat{e}_1 + \cos\theta\,\widehat{e}_2 + (-c_2 - \widehat{c}_1\sin\theta + \widehat{c}_2\cos\theta)\frac{\partial}{\partial\theta}. \tag{24.12}$$

Admissible motions of the rolling bodies are trajectories of the control system

$$\dot q = u_1X_1(q) + u_2X_2(q), \quad q \in Q, \quad u_1, u_2 \in \mathbb{R}. \tag{24.13}$$

24.4 Controllability

Denote the Gaussian curvatures of the Riemannian manifolds M and $\widehat{M}$ by k and $\widehat{k}$ respectively. We lift these curvatures from M and $\widehat{M}$ to Q:

$$k(q) = k(\pi(q)), \quad \widehat{k}(q) = \widehat{k}(\widehat{\pi}(q)), \quad q \in Q.$$

Theorem 24.1. (1) *The reachable set O of system* (24.13) *from a point $q \in Q$ is an immersed smooth connected submanifold of Q with dimension equal to* 2 *or* 5. *Specifically:*

$$\begin{aligned}(k - \widehat{k})|_O \equiv 0 \quad &\Rightarrow \quad \dim O = 2,\\ (k - \widehat{k})|_O \not\equiv 0 \quad &\Rightarrow \quad \dim O = 5.\end{aligned}$$

(2) *There exists an injective correspondence between isometries $i : M \to \widehat{M}$ and* 2*-dimensional reachable sets O of system* (24.13). *In particular, if the manifolds M and $\widehat{M}$ are isometric, then system* (24.13) *is not globally controllable.*

(3) *Suppose that both manifolds M and $\widehat{M}$ are complete and simply connected. Then the correspondence between isometries $i : M \to \widehat{M}$ and* 2*-dimensional reachable sets O of system* (24.13) *is bijective. In particular, system* (24.13) *is globally controllable iff the manifolds M and $\widehat{M}$ are not isometric.*

Proof. (1) By the Orbit theorem, the reachable set of the symmetric system (24.13), i.e., the orbit of the distribution Δ through any point $q \in Q$, is an immersed smooth connected submanifold of Q. Now we show that any orbit O of Δ has dimension either 2 or 5.

Fix an orbit O and assume first that at some point $q \in O$ the manifolds M and $\widehat{M}$ have different Gaussian curvatures: $k(q) \neq \widehat{k}(q)$. In order to construct a frame on Q, compute iterated Lie brackets of the fields X_1, X_2:

$$
\begin{aligned}
X_3 &= [X_1, X_2] = c_1 X_1 + c_2 X_2 + (\widehat{k} - k)\frac{\partial}{\partial\,\theta}, && (24.14)\\
X_4 &= [X_1, X_3] \\
&= (X_1 c_1) X_1 + (X_1 c_2) X_2 + c_2 X_3 + (X_1(\widehat{k} - k))\frac{\partial}{\partial\,\theta} + (\widehat{k} - k)\left[X_1, \frac{\partial}{\partial\,\theta}\right], && (24.15)\\
X_5 &= [X_2, X_3] \\
&= (X_2 c_1) X_1 + (X_2 c_2) X_2 - c_1 X_3 + (X_2(\widehat{k} - k))\frac{\partial}{\partial\,\theta} + (\widehat{k} - k)\left[X_2, \frac{\partial}{\partial\,\theta}\right], && (24.16)
\end{aligned}
$$

$$
\begin{aligned}
\left[X_1, \frac{\partial}{\partial\,\theta}\right] &= \sin\theta\,\widehat{e}_1 - \cos\theta\,\widehat{e}_2 + (\cdots)\frac{\partial}{\partial\,\theta}, && (24.17)\\
\left[X_2, \frac{\partial}{\partial\,\theta}\right] &= \cos\theta\,\widehat{e}_1 + \cos\theta\,\widehat{e}_2 + (\cdots)\frac{\partial}{\partial\,\theta}. && (24.18)
\end{aligned}
$$

In the computation of bracket (24.14) we used expression (24.4) of Gaussian curvature through structure constants. It is easy to see that

$$
\begin{aligned}
\mathrm{Lie}(X_1, X_2)(q) &= \mathrm{span}\,(X_1, X_2, X_3, X_4, X_5)\,(q) = \mathrm{span}\left(e_1, e_2, \widehat{e}_1, \widehat{e}_2, \frac{\partial}{\partial\,\theta}\right)(q)\\
&= T_q Q.
\end{aligned}
$$

System (24.13) has the full rank at the point $q \in O$ where $k \neq \widehat{k}$, thus $\dim O = 5$.

On the other hand, if $k(q) = \widehat{k}(q)$ at all points $q \in O$, then equality (24.14) implies that the distribution Δ is integrable, thus $\dim O = 2$.

(2) Let $i : M \to \widehat{M}$ be an isometry. Its graph

$$
\Gamma = \left\{ q \in Q \mid q = i_{*x} : T_x M \to T_{\widehat{x}} \widehat{M}, \quad x \in M, \ \widehat{x} = i(x) \in \widehat{M} \right\}
$$

is a smooth 2-dimensional submanifold of Q. We prove that Γ is an orbit of Δ. Locally, choose an orthonormal frame e_1, e_2 in M and take the corresponding orthonormal frame $\widehat{e}_1 = i_* e_1$, $\widehat{e}_2 = i_* e_2$ in $\widehat{M}$. Then $\theta|_\Gamma = 0$. Since $\widehat{c}_1 = c_1$, $\widehat{c}_2 = c_2$, and $k(q) = \widehat{k}(q)$, restrictions of the fields X_1, X_2 read

$$
X_1|_\Gamma = e_1 + \widehat{e}_1, \quad X_2|_\Gamma = e_2 + \widehat{e}_2.
$$

Then it follows that the fields X_1, X_2 are tangent to Γ. Lie bracket (24.14) yields

$$
[X_1, X_2]|_\Gamma = c_1 X_1 + c_2 X_2,
$$

thus Γ is an orbit of Δ. Distinct isometries $i_1 \neq i_2$ have distinct graphs $\Gamma_1 \neq \Gamma_2$, i.e., the correspondence between isometries and 2-dimensional orbits is injective.

(3) Now assume that the manifolds M and $\widehat{M}$ are complete and simply connected. Let O be a 2-dimensional orbit of Δ. We construct an isometry $i : M \to \widehat{M}$ with the graph O.

Notice first of all that for any Lipschitzian curve $x(t) \in M$, $t \in [0, t_1]$, and any point $q_0 \in Q$, there exists a trajectory $q(t)$ of system (24.13) such that $\pi(q(t)) = x(t)$ and $q(0) = q_0$. Indeed, a Lipschitzian curve $x(t)$, $t \in [0, t_1]$, is a trajectory of a nonautonomous ODE $\dot{x} = u_1(t)e_1(x) + u_2(t)e_2(x)$ for some $u_i \in L_\infty[0, t_1]$. Consider the lift of this equation to Q:

$$\dot{q} = u_1(t)X_1(q) + u_2(t)X_2(q), \qquad q(0) = q_0. \tag{24.19}$$

We have to show that the solution to this Cauchy problem is defined on the whole segment $[0, t_1]$. Denote by R the Riemannian length of the curve $x(t)$ and by $B(x_0, 2R) \subset M$ the closed Riemannian ball of radius $2R$ centered at x_0. The curve $x(t)$ is contained in $B(x_0, 2R)$ and does not intersect with its boundary. Notice that the ball $B(x_0, 2R)$ is a closed and bounded subset of the complete space M, thus it is compact. The projection $\widehat{x}(t) \in \widehat{M}$ of the maximal solution $q(t)$ to Cauchy problem (24.19) has Riemannian length not greater than R, thus it is contained in the compact $B(\widehat{x}_0, 2R) \subset \widehat{M}$, $\widehat{x}_0 = \widehat{\pi}(q_0)$, and does not intersect with its boundary. Summing up, the maximal solution $q(t)$ to (24.19) is contained in the compact $K = B(x_0, 2R) \times B(\widehat{x}_0, 2R) \times S^1$ and does not come to its boundary. Thus the maximal solution $q(t)$ is defined at the whole segment $[0, t_1]$.

Now it easily follows that $\pi(O) = M$ for the two-dimensional orbit O. Indeed, let $q_0 \in O$, then $x_0 = \pi(q_0) \in \pi(O)$. Take any point $x_1 \in M$ and connect it with x_0 by a Lipschitzian curve $x(t)$, $t \in [0, t_1]$. Let $q(t)$ be the lift of $x(t)$ to the orbit O with the initial condition $q(0) = q_0$. Then $q(t_1) \in O$ and $x_1 = \pi(q(t_1)) \in \pi(O)$. Thus $\pi(O) = M$. Similarly, $\widehat{\pi}(O) = \widehat{M}$.

The projections

$$\pi : O \to M \quad \text{and} \quad \widehat{\pi} : O \to \widehat{M} \tag{24.20}$$

are local diffeomorphisms since

$$\begin{aligned} \pi_*(X_1) &= e_1, & \widehat{\pi}_*(X_1) &= \cos\theta\,\widehat{e}_1 + \sin\theta\,\widehat{e}_2, \\ \pi_*(X_2) &= e_2, & \widehat{\pi}_*(X_2) &= -\sin\theta\,\widehat{e}_1 + \cos\theta\,\widehat{e}_2. \end{aligned}$$

Moreover, it follows that projections (24.20) are global diffeomorphisms. Indeed, let $q \in O$. Any Lipschitzian curve $x(\cdot)$ on M starting from $\pi(q)$ has a unique lift to O starting from q and this lift continuously depends on $x(\cdot)$. Suppose that $q' \in O$, $q' \neq q$, $\pi(q') = \pi(q)$, and $q(\cdot)$ is a path on O connecting q with q'. Contracting the loop $\pi(q(\cdot))$ and taking the lift of the contraction, we come to a contradiction with the local invertibility of $\pi|_O$. Hence $\pi|_O$ is globally invertible, thus it is a global diffeomorphism. The same is true for $\widehat{\pi}|_O$.

Thus we can define a diffeomorphism

$$i = \widehat{\pi} \circ (\pi|_O)^{-1} : M \to \widehat{M}.$$

Since

$$i_* e_1 = \cos\theta\, \widehat{e}_1 + \sin\theta\, \widehat{e}_2,$$
$$i_* e_2 = -\sin\theta\, \widehat{e}_1 + \cos\theta\, \widehat{e}_2,$$

the mapping i is an isometry.

If the manifolds M and $\widehat{M}$ are not isometric, then all reachable sets of system (24.13) are 5-dimensional, thus open subsets of Q. But Q is connected, thus it is a single reachable set. □

24.5 Length minimization problem

24.5.1 Problem statement

Suppose that $k(x) \neq \widehat{k}(\widehat{x})$ for any $x \in M$, $\widehat{x} \in \widehat{M}$, i.e., $k - \widehat{k} \neq 0$ on Q. Then, by item (1) of Theorem 24.1, system (24.13) is completely controllable. Consider the following optimization problem: given any two contact configurations of the system of rolling bodies, find an admissible motion of the system that steers the first configuration into the second one and such that the path of the contact point in M (or, equivalently, in $\widehat{M}$) was the shortest possible. This geometric problem is stated as the following optimal control one:

$$\dot{q} = u_1 X_1 + u_2 X_2, \qquad q \in Q, \quad u = (u_1, u_2) \in \mathbb{R}^2, \tag{24.21}$$
$$q(0) = q_0, \qquad q(t_1) = q_1, \qquad t_1 \text{ fixed},$$
$$l = \int_0^{t_1} (u_1^2 + u_2^2)^{1/2}\, dt \to \min .$$

Notice that projections of ODE (24.21) to M and $\widehat{M}$ read respectively as

$$\dot{x} = u_1 e_1 + u_2 e_2, \qquad x \in M,$$

and

$$\dot{\widehat{x}} = u_1(\cos\theta\, \widehat{e}_1 + \sin\theta\, \widehat{e}_2) + u_2(-\sin\theta\, \widehat{e}_1 + \cos\theta\, \widehat{e}_2), \qquad \widehat{x} \in \widehat{M},$$

thus the sub-Riemannian length l of the curve $q(t)$ is equal to the Riemannian length of the both curves $x(t)$ and $\widehat{x}(t)$.

As usual, we replace the length l by the action:

$$J = \frac{1}{2} \int_0^{t_1} (u_1^2 + u_2^2)\, dt \to \min,$$

and restrict ourselves to constant velocity curves:

$$u_1^2 + u_2^2 \equiv \text{const} \neq 0.$$

24.5.2 PMP

As we showed in the proof of Theorem 24.1, the vector fields $X_1, \dots, X_5$ form a frame on Q, see (24.14)–(24.16). Denote the corresponding Hamiltonians linear on fibers in T^*Q:

$$g_i(\mu) = \langle \mu, X_i \rangle, \qquad \mu \in T^*Q, \quad i = 1, \dots, 5.$$

Then the Hamiltonian of PMP reads

$$g_u^\nu(\mu) = u_1 g_1(\mu) + u_2 g_2(\mu) + \frac{\nu}{2}(u_1^2 + u_2^2),$$

and the corresponding Hamiltonian system is

$$\dot\mu = u_1 \vec g_1(\mu) + u_2 \vec g_2(\mu), \qquad \mu \in T^*Q. \tag{24.22}$$

24.5.3 Abnormal extremals

Let $\nu = 0$. The maximality condition of PMP implies that

$$g_1(\mu_t) = g_2(\mu_t) \equiv 0 \tag{24.23}$$

along abnormal extremals. Differentiating these equalities by virtue of the Hamiltonian system (24.22), we obtain one more identity:

$$g_3(\mu_t) \equiv 0. \tag{24.24}$$

The next differentiation by virtue of (24.22) yields an identity containing controls:

$$u_1(t) g_4(\mu_t) + u_2(t) g_5(\mu_t) \equiv 0. \tag{24.25}$$

It is natural to expect that conditions (24.22)–(24.25) on abnormal extremals on Q should project to reasonable geometric conditions on M and $\widehat{M}$. This is indeed the case, and now we derive ODEs for projections of abnormal extremals to M and $\widehat{M}$.

According to the splitting of the tangent spaces

$$T_q Q = T_x M \oplus T_{\widehat{x}} \widehat{M} \oplus T_\theta S^1,$$

the cotangent space split as well:

$$T_q^* Q = T_x^* M \oplus T_{\widehat{x}}^* \widehat{M} \oplus T_\theta^* S^1,$$

$$\mu = \lambda + \widehat\lambda + \alpha\, d\theta, \qquad \mu \in T_q^* Q, \ \lambda \in T_x^* M, \ \widehat\lambda \in T_{\widehat{x}}^* \widehat{M}, \ \alpha\, d\theta \in T_\theta^* S^1.$$

Then

$$\begin{aligned} g_1(\mu) &= \langle \mu, X_1 \rangle = \left\langle \lambda + \widehat\lambda + \alpha\, d\theta,\ e_1 + \cos\theta\, \widehat e_1 + \sin\theta\, \widehat e_2 + b_1 \frac{\partial}{\partial\theta} \right\rangle \\ &= h_1(\lambda) + \cos\theta\, \widehat h_1(\widehat\lambda) + \sin\theta\, \widehat h_2(\widehat\lambda) + \alpha b_1, \qquad (24.26) \\ g_2(\mu) &= \langle \mu, X_2 \rangle = \left\langle \lambda + \widehat\lambda + \alpha\, d\theta,\ e_2 - \sin\theta\, \widehat e_1 + \cos\theta\, \widehat e_2 + b_2 \frac{\partial}{\partial\theta} \right\rangle \\ &= h_2(\lambda) - \sin\theta\, \widehat h_1(\widehat\lambda) + \cos\theta\, \widehat h_2(\widehat\lambda) + \alpha b_2, \qquad (24.27) \end{aligned}$$

where $b_1 = -c_1 + \widehat{c}_1 \cos\theta + \widehat{c}_2 \sin\theta$, $b_2 = -c_2 - \widehat{c}_1 \sin\theta + \widehat{c}_2 \cos\theta$,

$$\begin{aligned} g_3(\mu) = \langle \mu, X_3 \rangle &= \left\langle \lambda + \widehat{\lambda} + \alpha d\theta,\ c_1 X_1 + c_2 X_2 + (\widehat{k} - k)\frac{\partial}{\partial\theta} \right\rangle \\ &= c_1 g_1(\mu) + c_2 g_2(\mu) + \alpha(\widehat{k} - k). \end{aligned} \tag{24.28}$$

Then identities (24.23) and (24.24) read as follows:

$$\begin{aligned} &\alpha = 0, \\ &h_1 + \cos\theta\,\widehat{h}_1 + \sin\theta\,\widehat{h}_2 = 0, \\ &h_2 - \sin\theta\,\widehat{h}_1 + \cos\theta\,\widehat{h}_2 = 0. \end{aligned}$$

Under these conditions, taking into account equalities (24.15)–(24.18), we have:

$$\begin{aligned} g_4(\mu) &= \Big\langle \lambda + \widehat{\lambda},\ (X_1 c_1) X_1 + (X_2 c_2) X_2 + c_2 X_3 \\ &\qquad + (X_1(\widehat{k} - k))\frac{\partial}{\partial\theta} + (\widehat{k} - k)\left[X_1, \frac{\partial}{\partial\theta}\right] \Big\rangle \\ &= (\widehat{k} - k)(\sin\theta\,\widehat{h}_1 - \cos\theta\,\widehat{h}_2) = (\widehat{k} - k) h_2, \\ g_5(\mu) &= \Big\langle \lambda + \widehat{\lambda},\ (X_2 c_1) X_1 + (X_2 c_2) X_2 - c_1 X_3 + \\ &\qquad (X_2(\widehat{k} - k))\frac{\partial}{\partial\theta} + (\widehat{k} - k)\left[X_2, \frac{\partial}{\partial\theta}\right] \Big\rangle \\ &= (\widehat{k} - k)(\cos\theta\,\widehat{h}_1 + \sin\theta\,\widehat{h}_2) = -(\widehat{k} - k) h_1. \end{aligned}$$

Then identity (24.25) yields

$$u_1 h_2 - u_2 h_1 = 0.$$

That is, up to reparametrizations of time, abnormal controls satisfy the identities:

$$u_1 = h_1(\lambda), \qquad u_2 = h_2(\lambda). \tag{24.29}$$

In order to write down projections of the Hamiltonian system (24.22) to T^*M and $T^*\widehat{M}$, we decompose the Hamiltonian fields $\vec{g}_1$, $\vec{g}_2$. In view of equalities (24.26), (24.27), we have

$$\begin{aligned} \vec{g}_1 &= \vec{h}_1 + \cos\theta\,\vec{\widehat{h}}_1 + \sin\theta\,\vec{\widehat{h}}_2 + (-\sin\theta\,\widehat{h}_1 + \cos\theta\,\widehat{h}_2)\vec{\theta} + \alpha\vec{a}_1 + a_1\vec{\alpha}, \\ \vec{g}_2 &= \vec{h}_2 - \sin\theta\,\vec{\widehat{h}}_1 + \cos\theta\,\vec{\widehat{h}}_2 + (-\cos\theta\,\widehat{h}_1 - \sin\theta\,\widehat{h}_2)\vec{\theta} + \alpha\vec{a}_2 + a_2\vec{\alpha}. \end{aligned}$$

It follows easily that $\vec{\theta} = -\frac{\partial}{\partial\alpha}$. Since $\alpha = 0$ along abnormal extremals, projection to T^*M of system (24.22) with controls (24.29) reads

$$\dot{\lambda} = h_1\vec{h}_1 + h_2\vec{h}_2 = \vec{H}, \qquad H = \frac{1}{2}(h_1^2 + h_2^2).$$

Consequently, projections $x(t) = \pi(q(t))$ are Riemannian geodesics in M.

Similarly, for projection to $\widehat{M}$ we obtain the equalities

$$u_1 = -\cos\theta\, \widehat{h}_1 - \sin\theta\, \widehat{h}_2, \qquad u_2 = \sin\theta\, \widehat{h}_1 - \cos\theta\, \widehat{h}_2,$$

thus

$$\begin{aligned}\dot{\widehat{\lambda}} &= \left(-\cos\theta\, \widehat{h}_1 - \sin\theta\, \widehat{h}_2\right)\left(\cos\theta\, \vec{\widehat{h}}_1 + \sin\theta\, \vec{\widehat{h}}_2\right) \\ &\qquad + \left(\sin\theta\, \widehat{h}_1 - \cos\theta\, \widehat{h}_2\right)\left(-\sin\theta\, \vec{\widehat{h}}_1 + \cos\theta\, \vec{\widehat{h}}_2\right) \\ &= -\widehat{h}_1\, \vec{\widehat{h}}_1 - \widehat{h}_2\, \vec{\widehat{h}}_2 = -\vec{\widehat{H}}, \qquad \widehat{H} = \frac{1}{2}\left(\widehat{h}_1^2 + \widehat{h}_2^2\right),\end{aligned}$$

i.e., projections $\widehat{x}(t) = \widehat{\pi}(q(t))$ are geodesics in $\widehat{M}$.

We proved the following statement.

Proposition 24.1. *Projections of abnormal extremal curves $x(t) = \pi(q(t))$ and $\widehat{x}(t) = \widehat{\pi}(q(t))$ are Riemannian geodesics respectively in M and $\widehat{M}$.*

Abnormal sub-Riemannian geodesics $q(t)$ are optimal at segments $[0, \tau]$ at which at least one of Riemannian geodesics $x(t)$, $\widehat{x}(t)$ is a length minimizer. In particular, short arcs of abnormal geodesics $q(t)$ are optimal.

24.5.4 Normal extremals

Let $\nu = -1$. The normal extremal controls are determined from the maximality condition of PMP:

$$u_1 = g_1, \qquad u_2 = g_2,$$

and normal extremals are trajectories of the Hamiltonian system

$$\begin{aligned}\dot{\mu} &= \vec{G}(\mu), \qquad \mu \in T^*Q, \\ G &= \frac{1}{2}(g_1^2 + g_2^2).\end{aligned} \tag{24.30}$$

The maximized Hamiltonian G is smooth, thus short arcs of normal extremal trajectories are optimal.

Consider the case where one of the rolling surfaces is the plane: $\widehat{M} = \mathbb{R}^2$. In this case the normal Hamiltonian system (24.30) can be written in a simple form. Choose the following frame on Q:

$$Y_1 = X_1, \quad Y_2 = X_2, \quad Y_3 = \frac{\partial}{\partial\theta}, \quad Y_4 = [Y_1, Y_3], \quad Y_5 = [Y_2, Y_3],$$

and introduce the corresponding linear on fibers Hamiltonians

$$m_i(\mu) = \langle \mu, Y_i \rangle, \qquad i = 1, \dots, 5.$$

Taking into account that $\widehat{c}_1 = \widehat{c}_2 = \widehat{k} = 0$, we compute Lie brackets in this frame:

$$
\begin{aligned}
&[Y_1, Y_2] = c_1 Y_1 + c_2 Y_2 - k Y_3, && \\
&[Y_1, Y_4] = -c_1 Y_5, && [Y_2, Y_4] = -c_2 Y_5, \\
&[Y_1, Y_5] = c_1 Y_4, && [Y_2, Y_5] = c_2 Y_4.
\end{aligned}
$$

Then the normal Hamiltonian system (24.30) reads as follows:

$$
\begin{aligned}
\dot{m}_1 &= -m_2(c_1 m_1 + c_2 m_2 - k m_3), \\
\dot{m}_2 &= m_1(c_1 m_1 + c_2 m_2 - k m_3), \\
\dot{m}_3 &= m_1 m_4 + m_2 m_5, \\
\dot{m}_4 &= -(c_1 m_1 + c_2 m_2) m_5, \\
\dot{m}_5 &= (c_1 m_1 + c_2 m_2) m_4, \\
\dot{q} &= m_1 X_1 + m_2 X_2.
\end{aligned}
$$

Notice that, in addition to the Hamiltonian $G = \dfrac{1}{2}(m_1^2 + m_2^2)$, this system has one more integral: $\rho = (m_4^2 + m_5^2)^{1/2}$. Introduce coordinates on the level surface $G = \dfrac{1}{2}$:

$$
m_1 = \cos\gamma, \quad m_2 = \sin\gamma, \quad m_3 = m, \quad m_4 = \rho\cos(\gamma+\psi), \quad m_5 = \rho\sin(\gamma+\psi).
$$

Then the Hamiltonian system simplifies even more:

$$
\begin{aligned}
\dot{\gamma} &= c_1 \cos\gamma + c_2 \sin\gamma - k m, \\
\dot{m} &= \rho \cos\psi, \\
\dot{\psi} &= k m, \\
\dot{q} &= \cos\gamma\, X_1 + \sin\gamma\, X_2.
\end{aligned}
$$

The case $k = \text{const}$, i.e., the sphere rolling on a plane, is completely integrable. This problem was studied in detail in book [14].

Appendix A

In this Appendix we prove several technical propositions from Chapter 2.

A.1 Homomorphisms and operators in $C^\infty(M)$

Lemma A.1. *On any smooth manifold M, there exists a function $a \in C^\infty(M)$ such that for any $N > 0$ exists a compact $K \Subset M$ for which*

$$a(q) > N \qquad \forall\, q \in M \setminus K.$$

Proof. Let e_k, $k \in \mathbb{N}$, be a partition of unity on M: the functions $e_k \in C^\infty(M)$ have compact supports $\operatorname{supp} e_k \Subset M$, which form a locally finite covering of M, and $\sum_{k=1}^\infty e_k \equiv 1$. Then the function $\sum_{k=1}^\infty k e_k$ can be taken as a. $\square$

Now we recall and prove Proposition 2.1.

Proposition 2.1. *Let $\varphi : C^\infty(M) \to \mathbb{R}$ be a nontrivial homomorphism of algebras. Then there exists a point $q \in M$ such that $\varphi = \widehat{q}$.*

Proof. For the homomorphism $\varphi : C^\infty(M) \to \mathbb{R}$, the set

$$\operatorname{Ker} \varphi = \{f \in C^\infty(M) \mid \varphi f = 0\}$$

is a maximal ideal in $C^\infty(M)$. Further, for any point $q \in M$, the set of functions

$$I_q = \{f \in C^\infty(M) \mid f(q) = 0\}$$

is an ideal in $C^\infty(M)$. To prove the proposition, we show that

$$\operatorname{Ker} \varphi \subset I_q \tag{A.1}$$

for some $q \in M$. Then it follows that $\operatorname{Ker} \varphi = I_q$ and $\varphi = \widehat{q}$.

By contradiction, suppose that $\operatorname{Ker} \varphi \not\subset I_q$ for any $q \in M$. This means that

$$\forall\, q \in M \quad \exists\, b_q \in \operatorname{Ker} \varphi \quad \text{s. t.} \quad b_q(q) \neq 0.$$

Changing if necessary the sign of b_q, we obtain that

$$\forall\, q \in M \quad \exists\, b_q \in \operatorname{Ker} \varphi,\ O_q \subset M \quad \text{s. t.} \quad b_q|_{O_q} > 0. \tag{A.2}$$

Fix a function a given by Lemma A.1. Denote $\varphi(a) = \alpha$, then $\varphi(a-\alpha) = 0$, i.e.,

$$(a - \alpha) \in \operatorname{Ker} \varphi.$$

Moreover,

$$\exists\ K \Subset M \quad \text{s. t.} \quad a(q) - \alpha > 0 \quad \forall\ q \in M \setminus K.$$

Take a finite covering of the compact K by the neighborhoods O_q as in (A.2):

$$K \subset \bigcup_{i=1}^{n} O_{q_i},$$

and let $e_0, e_1, \dots, e_n \in C^\infty(M)$ be a partition of unity subordinated to the covering of M:

$$M \setminus K, O_{q_1}, \dots, O_{q_n}.$$

Then we have a globally defined function on M:

$$c = e_0(a - \alpha) + \sum_{i=1}^{n} e_i b_{q_i} > 0.$$

Since

$$1 = \varphi\left(c \cdot \frac{1}{c}\right) = \varphi(c) \cdot \varphi\left(\frac{1}{c}\right),$$

then

$$\varphi(c) \neq 0.$$

But $c \in \operatorname{Ker} \varphi$, a contradiction. Inclusion (A.1) is proved, and the proposition follows. $\square$

Now we formulate and prove the theorem on regularity properties of composition of operators in $C^\infty(M)$, in particular, for nonautonomous vector fields or flows on M.

Proposition A.1. *Let A_t and B_t be continuous w.r.t. t families of linear continuous operators in $C^\infty(M)$. Then the composition $A_t \circ B_t$ is also continuous w.r.t. t. If in addition the families A_t and B_t are differentiable at $t = t_0$, then the family $A_t \circ B_t$ is also differentiable at $t = t_0$, and its derivative is given by the Leibniz rule:*

$$\left.\frac{d}{dt}\right|_{t_0} (A_t \circ B_t) = \left(\left.\frac{d}{dt}\right|_{t_0} A_t\right) \circ B_{t_0} + A_{t_0} \circ \left(\left.\frac{d}{dt}\right|_{t_0} B_t\right).$$

Proof. To prove the continuity, we have to show that for any $a \in C^\infty(M)$, the following expression tends to zero as $\varepsilon \to 0$:

$$(A_{t+\varepsilon} \circ B_{t+\varepsilon} - A_t \circ B_t)\, a = A_{t+\varepsilon} \circ (B_{t+\varepsilon} - B_t)\, a + (A_{t+\varepsilon} - A_t) \circ B_t a.$$

By continuity of the family A_t, the second term $(A_{t+\varepsilon} - A_t) \circ B_t a \to 0$ as $\varepsilon \to 0$. Since the family B_t is continuous, the set of functions $(B_{t+\varepsilon} - B_t)\, a$ lies in any preassigned neighborhood of zero in $C^\infty(M)$ for sufficiently small ε. For any $\varepsilon_0 > 0$, the family $A_{t+\varepsilon}$, $|\varepsilon| < \varepsilon_0$, is locally bounded, thus equicontinuous by the Banach-Steinhaus theorem. Consequently, $A_{t+\varepsilon} \circ (B_{t+\varepsilon} - B_t)\, a \to 0$ as $\varepsilon \to 0$. Continuity of the family $A_t \circ B_t$ follows.

The differentiability and Leibniz rule follow similarly from the decomposition

$$\frac{1}{\varepsilon}\left(A_{t+\varepsilon} \circ B_{t+\varepsilon} - A_t \circ B_t\right) a = A_{t+\varepsilon} \circ \frac{1}{\varepsilon}\left(B_{t+\varepsilon} - B_t\right) a + \frac{1}{\varepsilon}\left(A_{t+\varepsilon} - A_t\right) \circ B_t a.$$

□

A.2 Remainder term of the chronological exponential

Here we prove estimate (2.13) of the remainder term for the chronological exponential.

Lemma A.2. *For any $t_1 > 0$, complete nonautonomous vector field V_t, compactum $K \Subset M$, and integer $s \geq 0$, there exist $C > 0$ and a compactum $K' \Subset M$, $K \subset K'$, such that*

$$\|P_t a\|_{s,K} \leq C\, e^{C \int_0^t \|V_\tau\|_{s,K'}\, d\tau} \|a\|_{s,K'}, \qquad a \in C^\infty(M), \quad t \in [0, t_1], \tag{A.3}$$

where

$$P_t = \overrightarrow{\exp} \int_0^t V_\tau \, d\tau.$$

Proof. Denote the compact set

$$K_t = \cup \{P_\tau(K) \mid \tau \in [0,t]\}$$

and introduce the function

$$\begin{aligned}
\alpha(t) &= \sup\left\{ \frac{\|P_t a\|_{s,K}}{\|a\|_{s+1,K_t}} \mid a \in C^\infty(M),\ \|a\|_{s+1,K_t} \neq 0 \right\} \\
&= \sup\left\{ \|P_t a\|_{s,K} \mid a \in C^\infty(M),\ \|a\|_{s+1,K_t} = 1 \right\}.
\end{aligned} \tag{A.4}$$

Notice that the function $\alpha(t)$, $t \in [0, t_1]$, is measurable since the supremum in the right-hand side of (A.4) may be taken over only an arbitrary countable dense subset of $C^\infty(M)$. Moreover, in view of inequalities (2.3) and

$$\|a\|_{s,P_t(K)} \leq \|a\|_{s+1,K_t},$$

the function $\alpha(t)$ is bounded on the segment $[0, t_1]$.

As in the definition of the seminorms $\|\cdot\|_{s,K}$ in Section 2.2, we fix a proper embedding $M \subset \mathbb{R}^N$ and vector fields $h_1, \ldots, h_N \in \operatorname{Vec} M$ that span tangent spaces to M.

Let $q_0 \in K$ be a point at which

$$\|P_t a\|_{s,K} = \sup \{|h_{i_l} \circ \dots \circ h_{i_1}(P_t a)(q)| \mid q \in K,\ 1 \le i_1, \dots, i_l \le N,\ 1 \le l \le s\}$$

attains its upper bound, and let $p_a = p_a(x_1, \dots, x_N)$ be the polynomial of degree $\le s$ whose derivatives of order up to and including s at the point $q_t = P_t(q_0)$ coincide with the corresponding derivatives of a at q_t. Then

$$\begin{aligned} &\|P_t a\|_{s,K} = |h_{i_l} \circ \dots \circ h_{i_1}(P_t p_a)(q_0)| \le \|P_t p_a\|_{s,K}, \\ &\|p_a\|_{s,q_t} \le \|a\|_{s,K_t}. \end{aligned} \tag{A.5}$$

In the finite-dimensional space of all real polynomials of degree $\le s$, all norms are equivalent, so there exists a constant $C > 0$ which does not depend on the choice of the polynomial p of degree $\le s$ such that

$$\frac{\|p\|_{s,K_t}}{\|p\|_{s,q_t}} \le C. \tag{A.6}$$

Inequalities (A.5) and (A.6) give the estimate

$$\frac{\|P_t a\|_{s,K}}{\|a\|_{s,K_t}} \le \frac{\|P_t p_a\|_{s,K}}{\|p_a\|_{s,q_t}} \le C \frac{\|P_t p_a\|_{s,K}}{\|p_a\|_{s,K_t}} = C \frac{\|P_t p_a\|_{s,K}}{\|p_a\|_{s+1,K_t}} \le C\alpha(t). \tag{A.7}$$

Since

$$P_t a = a + \int_0^t P_\tau \circ V_\tau a \, d\tau,$$

then

$$\|P_t a\|_{s,K} \le \|a\|_{s,K} + \int_0^t \|P_\tau \circ V_\tau a\|_{s,K} \, d\tau$$

by inequality (A.7) and definition (2.2)

$$\le \|a\|_{s,K} + C \|a\|_{s+1,K_t} \int_0^t \alpha(\tau) \|V_\tau\|_{s,K_t} \, d\tau.$$

Dividing by $\|a\|_{s+1,K_t}$, we arrive at

$$\frac{\|P_t a\|_{s,K}}{\|a\|_{s+1,K_t}} \le 1 + C \int_0^t \alpha(\tau) \|V_\tau\|_{s,K_t} \, d\tau.$$

Thus we obtain the inequality

$$\alpha(t) \le 1 + C \int_0^t \alpha(\tau) \|V_\tau\|_{s,K_t} \, d\tau,$$

from which it follows by Gronwall's lemma that

$$\alpha(t) \le e^{C \int_0^t \|V_\tau\|_{s,K_t} \, d\tau}.$$

Then estimate (A.7) implies that

$$\|P_t a\|_{s,K} \le C \, e^{C \int_0^t \|V_\tau\|_{s,K_t} \, d\tau} \, \|a\|_{s,K_t},$$

and the required inequality (A.3) follows for any compactum $K' \supset K_t$. $\square$

Now we prove estimate (2.13).

Proof. Decomposition (2.11) can be rewritten in the form

$$P_t = S_m(t) + \int\limits_{\Delta_m(t)} \cdots \int P_{\tau_m} \circ V_{\tau_m} \circ \cdots \circ V_{\tau_1}\, d\tau_m \ \ldots\ d\tau_1,$$

where

$$S_m(t) = \mathrm{Id} + \sum_{k=1}^{m-1} \int\limits_{\Delta_k(t)} \cdots \int V_{\tau_k} \circ \cdots \circ V_{\tau_1}\, d\tau_k \ \ldots\ d\tau_1.$$

Then

$$\|(P_t - S_m(t))a\|_{s,K} \leq \int\limits_{\Delta_m(t)} \cdots \int \|P_{\tau_m} \circ V_{\tau_m} \circ \cdots \circ V_{\tau_1} a\|_{s,K}\, d\tau_m \ \ldots\ d\tau_1$$

by Lemma A.2

$$\leq C e^{C \int_0^t \|V_\tau\|_{s,K'}\, d\tau} \int\limits_{\Delta_m(t)} \cdots \int \|V_{\tau_m} \circ \cdots \circ V_{\tau_1} a\|_{s,K'}\, d\tau_m \ \ldots\ d\tau_1.$$

Now we estimate the last integral. By definition (2.2) of seminorms,

$$\begin{aligned}
&\int\limits_{\Delta_m(t)} \cdots \int \|V_{\tau_m} \circ \cdots \circ V_{\tau_1} a\|_{s,K'}\, d\tau_m \ \ldots\ d\tau_1 \\
&\leq \int\limits_{\Delta_m(t)} \cdots \int \|V_{\tau_m}\|_{s,K'} \|V_{\tau_{m-1}}\|_{s+1,K'} \cdots \|V_{\tau_1}\|_{s+m-1,K'}\, \|a\|_{s+m,K'}\, d\tau_m \ \ldots\ d\tau_1 \\
&\leq \|a\|_{s+m,K'} \\
&\qquad \int\limits_{\Delta_m(t)} \cdots \int \|V_{\tau_m}\|_{s+m-1,K'} \|V_{\tau_{m-1}}\|_{s+m-1,K'} \cdots \|V_{\tau_1}\|_{s+m-1,K'}\, d\tau_m \ \ldots\ d\tau_1 \\
&= \|a\|_{s+m,K'}\, \frac{1}{m!} \left(\int_0^t \|V_\tau\|_{s+m-1,K'}\, d\tau \right)^m,
\end{aligned}$$

and estimate (2.13) follows:

$$\|(P_t - S_m(t))\, a\|_{s,K} \leq \frac{C}{m!} e^{C \int_0^t \|V_\tau\|_{s,K'}\, d\tau} \left(\int_0^t \|V_\tau\|_{s+m-1,K'}\, d\tau \right)^m \|a\|_{s+m,K'}.$$

□

Bibliography

[1] A. A. Agrachev, R. V. Gamkrelidze, The exponential representation of flows and the chronological calculus. (Russian) *Mat. sb.*, **107** (1978), 467–532. (English translation: Math. USSR Sb. **35** (1979), 727–785.)

[2] A.A. Agrachev, R.V. Gamkrelidze, and A. V. Sarychev, Local invariants of smooth control systems. *Acta Appl. Math.*, **14** (1989), 191–237.

[3] A.A. Agrachev, A. V. Sarychev, On reduction of a smooth system linear in the control. (Russian) *Mat. sb.*, **130** (1986), No. 1. (English translation: Math. USSR Sb. **58** (1987), 15–30.)

[4] V.I. Arnold, *Mathematical methods of classical mechanics*, Graduate Texts in Mathematics **60**, Springer-Verlag, 1989.

[5] V.I. Arnold, *Ordinary differential equations*, Springer-Verlag, 1992.

[6] V.I. Arnold, A.B. Givental, *Symplectic geometry*, Itogi Nauki, Sovr. probl. matem., Fundamentalnye napravleniya 4 (1985), VINITI, Moscow, 5–139; English translation in Encyclopaedia of Math. Scien., vol. 4, Springer.

[7] R. Bellman, *Introduction to the mathematical theory of control processes*, vol. 1, 2. New York : Academic Press, 1967.

[8] U. Boscain, T. Chambrion, and J.-P. Gauthier, On the $K+P$ problem for a three-level quantum system: Optimality implies resonance, *J. Dynam. Control Systems*, **8** (2002), 547–572.

[9] R.V. Gamkrelidze, *Principles of optimal control theory*, New York : Plenum Publishing Corporation, 1978.

[10] J.-P. Gauthier, I. A. K. Kupka, *Deterministic observation theory and applications*, Cambridge University Press, 2001.

[11] J. Hale, *Ordinary differential equations*, Robert E. Krieger Publishing Company, Inc., 1980.

[12] A. Isidori, *Nonlinear control systems: an introduction*, Springer-Verlag, 1985.

[13] N. Jacobson, *Lie algebras*, John Wiley & Sons, 1962.

[14] V. Jurdjevic, *Geometric control theory*, Cambridge University Press, 1997.

[15] H. Nijmeijer, A. van der Schaft, *Nonlinear dynamical control systems*, Springer-Verlag, 1990.

[16] L. S. Pontryagin, V. G. Boltyanskij, R. V. Gamkrelidze, E. F. Mishchenko, *The mathematical theory of optimal processes*, Oxford, Pergamon Press, 1964.

[17] L. S. Pontrjagin, *Topological groups*, Princeton Univ. Press, Princeton, New Jersey, 1946.

[18] P. H. Rabinowitz, Periodic solutions to Hamiltonian systems, *Commun. Pure Appl. Math.*, **31**, 157–184.

[19] R. T. Rockafellar, *Convex analysis*, Princeton, Princeton University Press, 1972.

[20] W. Rudin, *Functional analysis*, McGraw-Hill, Inc., 1973.

[21] I. E. Segal, R. A. Kunze, *Integrals and operators*, Springer, 1978.

[22] E. D. Sontag, *Mathematical control theory : deterministic finite dimensional systems*, Springer-Verlag, 1990.

[23] H.J. Sussmann, Lie brackets, real analyticity and geometric control, In: *Differential geometric control theory*, Proc. confer. held at Michigan State Technolog. Univ., June 28 – July 2, 1982, Eds.: Roger W. Brockett, R.S. Millman, H.J. Sussmann, Birkhäuser Boston, 1983.

[24] F. W. Warner, *Foundations of differentiable manifolds and Lie groups*, Glenview, Ill. : Scott, Foresman, 1971.

[25] A. Weinstein, Periodic orbits for convex Hamiltonian systems, *Ann. Math.*, **108**, 507–518.

[26] M. I. Zelikin, V. F. Borisov, *Theory of chattering control with applications to astronautics, robotics, economics, and engineering*, Basel: Birkhauser, 1994.

Index

Symbols

$D_u^2 F$ 272
$D_x \Phi$ 5
F^* 130
$L^{\angle}$ 149
L_f 137
P^t 7
$T^* M$ 130
$T_x M$ 3, 4
$[V_1, V_2]$ 13
$\mathrm{Ad}\, B$ 31
$\mathcal{A}_{q_0}$ 11, 55, 122
$\mathcal{A}_{q_0}(t)$ 11, 122
$\mathcal{A}_{q_0}^t$ 122
$\mathcal{O}_{q_0}$ 55
$\mathrm{Diff}\, M$ 3
$\Phi_* V$ 5
$\mathrm{GL}(n)$ 233
$\Lambda^1 M$ 131
$\Lambda^k M$ 134
Λ^Γ 149
$\mathrm{Lie}\, \mathcal{F}$ 58
$\mathrm{Lie}_q\, \mathcal{F}$ 58
$\mathrm{SO}(3)$ 71
$\mathrm{SO}(n)$ 87, 234
$\mathrm{SU}(n)$ 234
$\mathrm{Vec}\, M$ 4
$\mathrm{ad}\, V$ 31
$\widehat{F}$ 16, 134
$\overleftarrow{\exp}$ 33
$\omega_1 \wedge \omega_2$ 133
$\mathrm{Hess}_u\, F$ 271
$\mathrm{ind}_+\, Q$ 275
$\mathrm{ind}_-\, Q$ 275
σ 142
$\mathrm{so}(n)$ 87
$\vec{h}$ 143
$\overleftarrow{\exp}$ 27
$\overrightarrow{\exp}$ 23, 32
$\{a, b\}$ 145
$d\omega$ 136
e^{tV} 29
$e^{t\, \mathrm{ad}\, V}$ 33
i_f 138
s 141

A

attainable set 11, 122

C

Cartan's formula 138
Cauchy's formula 39
chronological exponential, left . . . 27
chronological exponential, right . 23
conjugate point 212
conjugate time 309, 313
connection 351
connection, Levi-Civita 352
control parameter 9
control system 9
control system, affine in control . 39
control system, bilinear 295
control system, completely controllable 41
control system, left-invariant . . . 234
control system, linear 39
control system, linear, Brunovsky normal form 114
control system, linear, Kronecker indices 112
control system, right-invariant . 235
control system, single-input 296
control, bang-bang 184

control, singular 185
controllability, local 295
convex hull 127
cost functional 121
cotangent bundle 130
cotangent bundle, canonical coordinates 131
cotangent bundle, canonical symplectic structure 142
cotangent bundle, trivialization 226
cotangent space 130
critical point, corank 273
curvature, 2-dimensional systems338
curvature, 3-dimensional affine systems 348
curvature, Gaussian 344

D

derivation 18
diffeomorphism 3
differential 5
differential form 131, 134
differential form, exterior differential 136
differential form, integral . . 131, 135
differential form, interior product138
differential form, Lie derivative 137
differential, second 272
distribution 65
distribution, 2-generating 295
distribution, integrable 65
distribution, integral manifold . . . 65
dual basis 129
dual space 129
Dubins car, with angular acceleration control 299
Dynamic Programming 223
dynamical system 4

E

embedding 3
endpoint mapping 269
endpoint mapping, differential . 281
endpoint mapping, second differential 283
equivalence, feedback 106
equivalence, local state 67
exponential 29
exterior k-form 132
exterior product 133
extremal 152
extremal trajectory 152
extremal, abnormal 164
extremal, nice singular 292
extremal, normal 164
extremal, regular 286
extremal, strictly abnormal 294
extremal, totally singular 288

F

family of functionals, regularity properties 21
family of functions, regularity properties 20
family of vector fields, bracket-generating 58, 95
family of vector fields, completely nonholonomic 58
family of vector fields, full-rank . 95
family of vector fields, symmetric55
feedback transformation 106
Filippov's theorem 124
flow 21
Frobenius condition 66
Fuller's phenomenon 306

G

General Position Condition 191

H

Hamilton-Jacobi equation 223
Hamiltonian 143
Hamiltonian system 143
Hessian 271

I

index 275
integral invariant of Poincaré-Cartan 216
isotropic 149

J
Jacobi equation, in moving frame 344
Jacobi equation, regular case . . . 311
Jacobi equation, singular case . . 315
Jacobi identity 31

L
Lagrangian subspaces 149
Lebesgue point 158
Lie algebra 32
Lie algebra, structural constants233
Lie bracket 13
Lie bracket, formula in coordinates 13
Lie group 233
Lie subgroup 235
Lie's theorem 236
Liouville 1-form 141

M
manifold 2
manifold, Riemannian 349
manifolds, diffeomorphic 3
mapping, locally open 272
mapping, proper 3
mapping, smooth 3
module 62
module, finitely generated 63
module, locally finitely generated64
Morse's lemma 275

N
needle-like variations 158

O
ODE 4
ODE, nonautonomous 22
optimal control 122
optimal control problem 122
optimal trajectory 122
optimality condition, Goh . 289, 293
optimality condition, Legendre 285, 292
optimality condition, Legendre, generalized 291, 293
optimality condition, Legendre, generalized, strong 291
optimality condition, Legendre, strong 286
optimality conditions, necessary151, 293
optimality conditions, sufficient218, 221, 320, 327
optimality, geometric 269
optimality, local finite-dimensional 273
optimality, local geometric 273
optimality, strong 320
orbit 55

P
parallel translation 352
point, Poisson stable 100
Poisson bracket 145
Pontryagin Maximum Principle165, 170, 292
problem, linear time-optimal . . . 191
problem, linear-quadratic 203
problem, Riemannian 243
problem, sub-Riemannian . 245, 293
problem, time-optimal 126

R
reachable set 11
reduction 329
relaxation 127
relaxed system 127
Riemannian structure 349
rigid body 71
rigid body, angular velocity 72
rigid body, Euler equations 76

S
space of control parameters 9
state 9
state space 9
submanifold 1
submanifold, immersed 56
switching function 263
symplectic manifold 143
symplectic space 149

symplectic structure 143

T

tangent space 3, 4
tangent vector 4
tautological 1-form 141
topology, $C^\infty(M)$ 19
transversality conditions 165

V

variations formula 35, 36
vector field 4, 21
vector field, autonomous 29
vector field, compatible 98
vector field, complete 7, 23
vector field, conservative 102
vector field, flow 8
vector field, Hamiltonian 143
vector field, Hamiltonian lift . . . 147
vector field, nonautonomous 21
vector field, parallel 352
vector field, Poisson stable 100
vector field, with compact support 23
vector fields, commutator 13