Calculus of Variations and Applications
Lecture Notes
Draft

Andrej Cherkaev and Elena Cherkaev

October 24, 2003

Contents

Part I

Preliminaries

Chapter 1

Introduction

1.1 Preliminary Remarks

Optimization The desire for optimality (perfection) is inherent in humans. The search for extremes inspires mountaineers, scientists, mathematicians, and the rest of the human race. The development of Calculus of Variation was driven by this noble desire. A mathematical technique of minimization of curves was developed in eighteen century to solve the problems of the best possible objects: The minimal surface, the shortest distance, or the trajectory of fastest travel.

In twentieth century, control theory emerged to address the extremal problems in science, engineering, and decision-making. These problems specialize the available the degrees of freedom by the so-called controls; these are constrained functions that can be optimally chosen. Optimal design theory addresses space-dependent analog of control problems. Minimax problems address optimization in a conflict situation or in undetermined environment. A special branch of the theory uses minimization principles to create effective algorithms such as finite element method to computing the solution.

Description of fundamental laws of Nature For centuries, scientists tried to proof that the Universe is rational, symmetric, or optimal in another sense. The attempts were made to formulate laws of natural sciences as extreme problems (variational principles)and to use the variational calculus as a scientific instrument to derive and investigate the motion and equilibria in Nature (Fermat, Lagrange, Gauss, Hamilton, Gibbs..). It was observed by Fermat that light "chooses" the trajectory that minimizes the time of travel, many equilibria correspond to the local minimum of the energy, motion of mechanical systems correspond to stationarity of a functional called the action, etc. In turn, the variational principles link together conservation laws and symmetries.

Does the actual trajectory minimize the action? This question motivated great researcher starting from Leibnitz and Fermat to develop variational methods to justify the Nature's "desire" to choose the most economic way to move, and it caused much heated discussions that involved philosophy and theology. The general principle by Maupertuis proclaims: If there occur some changes in nature, the amount of action necessary for this change must be as small as possible. In a sense, this principle would prove that our world is "the best of all worlds" – the conclusion defended by Fermat, Leibnitz, Maupertuis, and Euler but later ridiculed by Voltaire. It turns out that the action is minimized by short trajectories, but delivers stationary value for long ones, because of violation of so-called Jacobi conditions. This mathematical fact was disappointing for philosophical speculations, "a beautiful conjunction is ruined by an ugly fact." However, the relativity and the notion of the world lines returns the principle of minimization of a quantity at the real trajectory over all other trajectories.

Convenient description of the state of an object No matter do the real trajectories minimize the action or not, the variational methods in physics become an important tool for investigation of motions and equilibria. First, the variational formulation is convenient and economic: Instead of formulation of equations it is enough to write down a single functional that must be optimized at the actual configuration. The equations of the state of the system follow from the optimality requirement. Second, variational approach allows for accounting of symmetries, for invariants of the configuration, and (through the duality) for different differential equations that describe the same configuration in different terms.

There are several ways to describe a shape or a motion. The most explicit way is to describe positions of all points: Sphere is described by the functions $-\sqrt{1-x^2-y^2} \leq z(x,y) \leq \sqrt{1-x^2-y^2}$. The more implicit way is to formulate a differential equation which produces these positions as a solution: The curvature tensor is constant everywhere in a sphere. An even more implicit way is to formulate a variational problem: Sphere is a body with given volume that minimizes its surface area. The minimization of a single quantity produces the "most economic" shape in each point. Such implicit description goes back to Platonic ideals and is opposite to the Aristotelian principle to explicit description/classification of factual events (here, the explicit functions)

New mathematical concepts Working on optimization problems, mathematicians met paradoxes related to absence of optimal solution or its weird behavior; resolving these was useful for the theory itself and resulted in new mathematical development such as weak solutions of differential equations and related functional spaces (Hilbert and Sobolev spaces), various types of convergence of functional sequences, distributions and other limits of function's sequences, and other fundamentals of modern analysis.

Many computational methods as motivated by optimization problems and use the technique of minimization. Methods of search, finite elements, iterative schemes are part of optimization theory. The classical calculus of variation answers the question: What conditions must the minimizer satisfy? while the computational techniques are concern with the question: How to find or approximate the minimizer?

The list of main contributors to the calculus of variations includes the most distinguish mathematicians of the last three centuries such as Leibnitz, Newton, Bernoulli, Euler, Lagrange, Gauss, Jacobi, Hamilton, Hilbert.

History For the rich history of Calculus of variation we refer to such books as [Kline, Boyer].. Here we make several short remarks about the ideas of its development. The story started with the challenge:

> Given two points A and B in a vertical plane, what is the curve traced out by a point acted on only by gravity, which starts at A and reaches B in the shortest time. [1]

The brachistochrone problem was posed by Johann Bernoulli in Acta Eruditorum in June 1696. He introduced the problem as follows:

> *I, Johann Bernoulli, address the most brilliant mathematicians in the world. Nothing is more attractive to intelligent people than an honest, challenging problem, whose possible solution will bestow fame and remain as a lasting monument. Following the example set by Pascal, Fermat, etc., I hope to gain the gratitude of the whole scientific community by placing before the finest mathematicians of our time a problem which will test their methods and the strength of their intellect. If someone communicates to me the solution of the proposed problem, I shall publicly declare him worthy of praise.*

Within a year five solutions were obtained, Newton, Jacob Bernoulli, Leibniz and de L'Hôpital solving the problem in addition to Johann Bernoulli.

The May 1697 publication of Acta Eruditorum contained Leibniz's solution to the brachistochrone problem on page 205, Johann Bernoulli's solution on pages 206 to 211, Jacob Bernoulli's solution on pages 211 to 214, and a Latin translation of Newton's solution on page 223. The solution by de L'Hpital was not published until 1988 when, nearly 300 years later, Jeanne Peiffer presented it as Appendix 1 in [1].

Johann Bernoulli's solution divides the plane into strips and he assumes that the particle follows a straight line in each strip. The path is then piecewise linear. The problem is to determine the angle of the straight line segment in each strip and to do this he appeals to Fermat's principle, namely that light always follows the shortest possible time of travel. If v is the velocity in one strip at angle a to the vertical and u in the velocity in the next strip at angle b to the vertical then, according to the usual sine law v/sin a = u/sin b.

The optimal trajectory turns out to be a cycloid (see Section **??** for the derivation). Cycloid was a well investigated curve in seventeen century. Huygens

[1] Johann Bernoulli was not the first to consider the brachistochrone problem. Galileo in 1638 had studied the problem in 1638 in his famous work Discourse on two new sciences. He correctly concluded that the straight path is not the fastest one, but made an error concluding that an optimal trajectory is a part of a circle.

had shown in 1659, prompted by Pascal's challenge, that the cycloid is the tautochrone of isochrone: The curve for which the time taken by a particle sliding down the curve under uniform gravity to its lowest point is independent of its starting point. Johann Bernoulli ended his solution with the remark: *Before I end I must voice once more the admiration I feel for the unexpected identity of Huygens' tautochrone and my brachistochrone. ... Nature always tends to act in the simplest way, and so it here lets one curve serve two different functions, while under any other hypothesis we should need two curves.*

The methods which the brothers developed to solve the challenge problems they were tossing at each other were put in a general setting by Euler in *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes sive solutio problematis isoperimetrici latissimo sensu accepti* published in 1744. In this work, the English version of the title being *Method for finding plane curves that show some property of maxima and minima*, Euler generalizes the problems studies by the Bernoulli brothers but retains the geometrical approach developed by Johann Bernoulli to solve them. He found what has now come to be known as the Euler-Lagrange differential equation for a function of the maximizing or minimizing function and its derivative.

The idea is to find a function which maximizes or minimizes a certain quantity where the function is constrained to satisfy certain constraints. For example Johann Bernoulli had posed certain geodesic problems to Euler which, like the brachistochrone problem, were of this type. Here the problem was to find curves of minimum length where the curves were constrained to lie on a given surface. Euler, however, commented that his geometrical approach to these problems was not ideal and it only gave necessary conditions that a solution has to satisfy. The question of the existence of a solution was not solved by Euler's contribution.

Lagrange, in 1760, published Essay on a new method of determining the maxima and minima of indefinite integral formulas. It gave an analytic method to attach calculus of variations type problems. In the introduction to the paper Lagrange gives the historical development of the ideas which we have described above but it seems appropriate to end this article by giving what is in effect a summary of the developments in Lagrange's words:-

The first problem of this type [calculus of variations] which mathematicians solved was that of the brachistochrone, or the curve of fastest descent, which Johann Bernoulli proposed towards the end of the last century.

1.2 Properties of the extremals

Every optimization problem has several necessary components. It deals with a set $\mathcal{X}$ of admissible elements x, that can be real or complex numbers, differentiable curves, integrable functions, shapes, people in the town, or ants in the colony. A real-valued function $I(x)$ called objective is put it correspondence to each admissible element. It could be the absolute value of a number, integral of a function over an interval, value of the function at a point, weight or a person, or length of an ant. The goal is to find or characterize the element x_0 called minimizer, such that

$$I(x_0) \leq I(x), \quad \forall x \in \mathcal{X}$$

We denote this element as

$$x_0 = \arg\min_{x \in \mathcal{X}} I(x)$$

and we denote the value $I(x_0$ as

$$I(x_0) = \min_{x \in \mathcal{X}} I(x)$$

Next, we list the basic properties of any extreme problem that are based on the definition of the minimizer.

1. Minimum over larger set is smaller than minimum of the smaller set

   If $\mathcal{X}_1 \supseteq \mathcal{X}_2$, then

   $$\min_{x \in \mathcal{X}_1} F(x) \leq \min_{x \in \mathcal{X}_2} F(x)$$

2. Minimum of a function $F(x)$ is equal to the negative of maximum of $-F(x)$,

   $$\min_{x \in \mathcal{X}} F(x) = -\max_{x \in \mathcal{X}}(-F(x))$$

   This property allows us not to distinguish between minimization and maximization problem: We always can reformulate the maximization problem in the minimization form.

3. Minimum of sum is not smaller than the sum of minima.

   $$\min_x [f(x) + g(x)] \geq \min_x f(x) + \min_x g(x)$$

4. Linearity: If b and $c > 0$ are real numbers, than

   $$\min_x \left(c\, f(x) + b\right) = c \left(\min_x f(x)\right) + b$$

5. The minimizer is invariant to the superposition with any monotonic function. Namely, The minimizer

   $$x_0 = \arg\min_x f(x)$$

   where $f: \ X \to Y \subset R_1$ is also the minimizer of the problem

   $$x_0 = \arg\min_{x \in X} g\left(f(x)\right)$$

   where $g: \ Y \to R_1$ is monotone everywhere on Y.

6. Maximum of several minima is not larger than minimum of several maxima:
$$\max\left\{\min_x f_1(x), \ldots \min_x f_N(x)\right\} \le \min_x f_{\max}(x)$$
where
$$f_{\max}(x) = \max\{f_1(x), \ldots f_N(x)\}$$
7. Minimax theorem
$$\max_y \min_x f(x,y) \le \min_x \max_y f(x,y)$$

The listed properties can be proved by the straightforward use of the definition of the minimizer. We leave the prove to the reader.

1.3 Variational problem

The extremal (variational) problem requires to find an optimal function $u_0(x)$ which can be visualized as a curve (or a surface). The function $u_0(x)$ belongs to a set of admissible functions $\mathcal{U}$: $u \in \mathcal{U}$; it is assumed that $\mathcal{U}$ is a set of differentiable function on the interval $[a,b]$ that is denoted as $C_1[a,b]$. To measure the optimality of a curve, we define a *functional* (a real number) $I(u)$ which may depend on $u(x)$, and its derivative $u'(x)$ as well as on the independent variable x. The examples of variational problems are: The shortest path on a surface, the surface of minimal area, the best approximation by a smooth curve of the experimental data, the most economical strategy, etc.

The classical variational problem is formulated as follows: Find
$$I(u_0) = \min_{u(x)\in\mathcal{U}_b} J(u) \quad \mathcal{U}_b = \{u: \ u \in C_1(a,b), \ u(a) = \alpha, \ u(b) = \beta\} \tag{1.1}$$
where $x \in [a,b]$, $u_0(x)$ is an unknown function called the *minimizer*, the boundary values of u are fixed, $J(u)$ is the functional of the type
$$J(u) = \int_a^b F(x, u(x), u'(x))dx. \tag{1.2}$$
F is a function of three arguments, $x, u(x), u'(x)$, called *Lagrangian*, and it is assumed that the integral in (1.2) exists.

The value of the objective functional $I(u)$ (also called the cost functional) is a real number. Since real numbers are ordered, one can compare functionals $J(u_1), J(u_2), \ldots$ of different admissible functions $u_1, u_2, \ldots$, and build *minimizing sequences* of functions
$$u_1, \ u_2, \ldots, u_n, \ldots$$
with the property:
$$I(u_1) \ge I(u_2) \ge \ldots \ge I(u_n) \ldots$$

The limit u_0 of a minimizing sequence (if it exists) is called the *minimizer*; it delivers the minimum of I

$$I(u_0) \leq I(u) \quad \forall u \in \mathcal{U} \tag{1.3}$$

The minimizing sequence can always be built independently of the existence of the minimizer.

Generalization The formulated problem can be generalized in several ways.

- The minimizer and an admissible function can be a vector-function; the functional may depend of higher derivatives, and be of a more general form such as the ratio of two integrals.
- The integration can be performed over a spacial domain instead of the interval $[a, b]$; this domain may be completely or partly unknown and should be determined together with the minimizer.
- The problem may be constrained in several ways: The isoperimetric problem asks for the minimum of $I(u)$ if the value of another functional $I_r(u)$ is fixed. Example: find a domain of maximal area enclosed by a curve of a fixed length. The restricted problem asks for the minimum of $I(u_1, \ldots u_n)$ if a function(s) $\phi(u_1, \ldots u_n)$ is fixed everywhere. Example: The problem of geodesics: the shortest distance between two points on a surface. In this problem, the path must belong to the surface in everywhere.

Outline of the methods There are several groups of methods aimed to find the minimizer of an extremal problem.

1. **Methods of sufficient conditions.** These methods directly establish the inequality $I(u_0) \leq I(u), \forall u \in \mathcal{U}$. These rigorous methods are applicable to a small variety of problems, and the results are logically perfect. To establish the above inequality, the methods of convexity are commonly used. The method often requires a guess of the global minimizer u_0 and is applicable to relatively simple extremal problems.

2. **Methods of necessary conditions (variational methods).** Using these methods, we establish necessary conditions for $u(x)$ to provide a local minimum. In other words, the conditions tell that there is no other curve $u + \delta u$ that is (i) sufficiently close to the chosen curve u (that is assuming $||\delta u||$ is infinitesimal), (ii) satisfies the same boundary conditions, and (iii) corresponds to a smaller value $I(u+\delta u) < I(u)$ of the objective functional. The closeness of two compared curves allows for a relative simple form of the resulting *variational conditions* of optimality; on the other hand it restricts the generality of the obtained conditions.

   Variational methods yield to only necessary conditions of optimality because it is assumed that the compared trajectories are close to each other in a sense; they detect *locally optimal* curves provided that the assumptions are correct. On the other hand, variational methods are regular and robust; they are applicable to a great variety of extremal problems called *variational problems.* Necessary conditions are the true workhorses of extremal problem theory, while exact sufficient conditions are rare and remarkable exceptions.

3. **Direct optimization methods** These are aimed to building the minimizing sequence $\{u^s\}$ and provide a sequence of better solutions. Generally, the convergence to the true solution may not be required, but it is guaranteed that the solutions are improved on each step of the procedure: $I(u^s) \leq I(u^{s-1})$ for all s. These method require no a priori assumption of the dependence of functional on the minimizer, but only the possibility to compare a project with an improved one and chose the best of two. Of course, additional assumptions help to optimize the search but it can be conducted without these. The method can be applied to even discontinuous functionals.

	Global methods	Variational methods	Algorithmic search
Objectives	Search for the global minimum	Search for a local minimum	An improvement of existing solution
Means	Sufficient conditions	Necessary conditions	Algorithms of sequential improvement
Tools	Inequalities, Fixed point methods	Analysis of features of optimal trajectories	Gradient-type search
Existence of solution	Guaranteed	Not guaranteed	Not discussed
Applicability	Special problems	Large class of problems	Universal

Table 1.1: Approaches to variational problems

There are many books that expound the calculus of variations, including [**?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**].

Chapter 2

Geometric problems and Sufficient conditions

2.1 Convexity

The best source for the theory of convexity is probably the book [?].

2.1.1 Definitions and inequalities

Convexity is the most important and general feature of a function allowing for establishing inequalities. We start with definitions.

Definition 2.1.1 The set Ω in R_n is convex, if the following property holds. If any two points x_1 and x_2 belong to the set Ω, all points x_h with coordinates $x_h = \lambda x_1 + (1-\lambda)x_2$ belong to Ω

The interior of an ellipsoid or paraboloid are convex sets, the crescent is not convex. Convex sets are simply connected (do not have holes). The whole space R_n is a convex set, any linear hyperplane is also a convex set. The intersection of two convex sets is also a convex set, but the union of two convex sets may not be convex.

Next, we can define a convex function.

Definition 2.1.2 Consider a scalar function $f : \Omega \to R_1$ $\Omega \subset R_n$ of vector argument. Function F is called *convex* if it possesses the property

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda)f(x_2) \quad \forall x_1, x_2 \in R_n, \quad \forall \lambda \in [0,1] \tag{2.1}$$

Geometrically, the property (2.1) states that the graph of the convex function lies below the chord.

Figure 2.1: Basic property of convex function: real caption

Figure 2.2: Graph of nonconvex function $f(x) = \exp(-|x|)$

Example 2.1.1 Function $f(x) = x^2$ is convex. Indeed, $f(\lambda x_1 + (1-\lambda)x_2)$ can be represented as follows

$$(\lambda x_1 + (1-\lambda)x_2)^2 = \lambda(x_1)^2 + (1-\lambda)(x_2)^2 - C$$

where $C = \lambda(1-\lambda)(x_1 - x_2)^2 \geq 0$ is nonnegative. Therefore, (2.1) is true.

Properties of convex functions One can easily show (try!) that the function is convex if and only if

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2} \quad \forall x_1, x_2 \in R_n.$$

The convex function is differentiable almost everywhere. If it has second derivatives, the Hessian $He(f, x)$ is nonnegative everywhere

$$He(f, x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix} \geq 0.$$

Particularly, the convex function of one variable has the nonnegative second derivative:

$$f''(x) \geq 0 \quad \forall x \in R_1. \tag{2.2}$$

Convexity is a global property. If the inequality (2.2) is violated at one point, the function may be nonconvex everywhere. Consider, for example, $f(x) = \exp(-|x|)$. Its second derivative is positive everywhere, $f'' = \exp(-|x|)$ except $x = 0$ where it does not exist. This function is not convex, because

$$f(0) = 1 > \frac{1}{2}(f(x) + f(-x)) = \exp(-|x|) \quad \forall x \in R.$$

Jensen's inequality The definition (2.1) is equivalent to the so-called Jensen's inequality

$$f(x) \leq \frac{1}{N}\sum_{i=1}^{N} f(x + \zeta_i) \quad \forall \zeta_i : \sum_{i=1}^{N} \zeta_i = 0 \tag{2.3}$$

for any $x \in \Omega$. (Show the equivalence!)

Jensen's inequality enables us to define convexity in a point: The function f is convex at the point x if (2.3) holds.

Example 2.1.2 Function $f(x) = x^4 - x^2$ is convex if $x \not\in [-\sqrt{3}, \sqrt{3}]$. Notice that the inequality $f''(x) \geq 0$ holds in a smaller interval $x \not\in [-1, 1]$. At the intervals $[-\sqrt{3}, -1]$ and $[1, \sqrt{3}]$ the second derivative of F is positive, but F is not convex.

Integral form of Jensen inequality Increasing the number N of vectors ζ_i in (2.3), we find the integral form of Jensen inequality:

Function $F(z)$ is convex if and only if the inequality holds

$$F(z) \leq \frac{1}{b-a} \int_a^b F(z + \theta(x))dx \tag{2.4}$$

where

$$\int_a^b \theta(x)dx = 0 \tag{2.5}$$

and all integrals exist.

Remark 2.1.1 (Stability to perturbations) The integral form of the Jensen's inequality can be interpreted as follows: The minimum of an integral of a convex function corresponds to a constant minimizer. No perturbation with zero mean value can increase the functional.

Another interpretation is: The average of a convex function is larger then the function of an averaged argument.

Example 2.1.3 Assume that $F(u) = u^2$. We have

$$0 \leq \frac{1}{b-a} \int_a^b (z + \theta(x))^2 dx = z^2 + \frac{2z}{b-a} \int_a^b \theta(x)dx + \frac{1}{b-a} \int_a^b \theta(x)^2 dx$$

The second integral in the right-hand side is zero because of (2.5), the third integral is nonnegative. The required inequality

$$z^2 \leq \frac{1}{b-a} \int_a^b (z + \theta(x))^2 dx$$

(see (2.4) follows.

Next, we illustrate the use of convexity for solution of optimization problems. Being global property, convexity allow for establishing the most general between the optimal trajectory and any other trajectory.

2.1.2 Minimal distance at a plane, cone, and sphere

Let us start with the simplest problem with an intuitively expected solution: Find the minimal distance between the points (a, α) and (b, β) on a plane.

Consider any piece-wise differentiable path $x(t), y(t)$, $t \in [0.1]$ between these points. We set

$$x(0) = a, \quad x(1) = b, \quad y(0) = \alpha, \quad y(1) = \beta$$

The length of the path is

$$L(x,y) = \int_0^1 \sqrt{(x')^2 + (y')^2} dx$$

(We need the piece-wise differentiability of $x(t)$ and $y(t)$ to be able define the length of the pass) We have in mind to compare the path with the straight line (which we might expect to be a solution); therefore, we assume the representation

$$x(t) = a + t(b-a) + \int_0^t \psi_1(t)dt, \quad y(t) = \alpha + t(\beta - \alpha) + \int_o^t \psi_2(t)dt$$

the terms dependent on ϕ and ψ define the deviation from the straight path. The deviation in the beginning and in the end of the trajectory is zero, therefore we require

$$\int_0^1 \psi_1(t)dt = 0 \quad \int_0^1 \psi_2(t)dt = 0; \tag{2.6}$$

We prove that the deviation are identically zero at the optimal trajectory.

First, we rewrite the functional L in the introduced notations

$$L(\psi_1, \psi_2) = \int_0^1 \sqrt{((b-a) + \psi_1(t))^2 + ((\beta - \alpha) + \psi_2(t))^2} dx$$

where the Lagrangian $W((\psi_1, \psi_2)$ is

$$W((\psi_1, \psi_2) = \sqrt{((b-a) + \psi_1(t))^2 + ((\beta - \alpha) + \psi_2(t))^2}$$

and we use expressions for the derivatives x', y':

$$x' = (b-a) + \psi_1(t), \quad y' = (\beta - \alpha) + \psi_2(t).$$

The Lagrangian $W((\psi_1, \psi_2)$ is a convex function of its arguments ψ_1, ψ_2. Indeed, it is twice differentiable with respect to them and the Hessian He is

$$He(W) = \begin{pmatrix} y^2(x^2+y^2)^{-\frac{3}{2}} & xy(x^2+y^2)^{-\frac{3}{2}} \\ xy(x^2+y^2)^{-\frac{3}{2}} & x^2(x^2+y^2)^{-\frac{3}{2}} \end{pmatrix}$$

where $x = (b-a) + \psi_1(t)$ and $y = (\beta - \alpha) + \psi_2(t)$. The eigenvalues of the Hessian are equal to 0 and $(x^2 + y^2)^{-\frac{1}{2}}$ respectively, and therefore it is nonnegative defined (as the reader can easily check, the graph of $W((\psi_1, \psi_2)$ is a cone).

Due to Jensen's inequality in integral form, the convexity of the Lagrangian and the boundary conditions (2.6) lead to the relation

$$L(\psi_1, \psi_2) \geq L(0,0) = \int_0^1 \sqrt{(b-a)^2 + (\beta - \alpha)^2} dx$$

and to the minimizer $\psi_1 = 0, \psi_2 = 0$.

Thus we prove that the straight line corresponds to the shortest distance between two points. Notice that (1) we compare all differentiable trajectories no matter how far away from the straight line are they, and (2) we used our correct guess of the minimizer (the straight line) to compose the Lagrangian. These features are typical for the global optimization.

Geodesic on a cone Consider the problem of shortest path between two points of a cone, assuming that the path should lie on the conical surface. This problem is a simplest example of geodesics, the problem of the shortest path on a surface discussed below in Section **??**.

Because of simplicity of the cone's shape, the problem can be solved by pure geometrical means. Firstly, we show that it exists a ray on a cone that does not intersect with the geodesics between any two point if none of then coincide with the vertex. If this is not the case, than a geodesics makes a whole spiral around the cone. This cannot be because one can shorten the line replacing spiral part of a geodesics by an interval if a ray.

Now, let us cut the cone along this ray and straighten the surface: It becomes a wedge of a plane with the geodesics lying entirely inside the wedge. Obviously, the straighten does not change the length of a path. The coordinates of any point of the wedge can be characterized by a pair r, θ where $r > 0$ is the distance from the vertex and θ, $0 \leq \theta \leq \Theta$ is the angle counted from the cut. Parameter Θ characterizes the cone itself.

The problem is reduced to a problem of a shortest path between two points that lies within a wedge. Its solution depends on the angle Θ of the wedge. If this angle is smaller that π, $\Theta < \pi$, the optimal path is a straight line

$$r = A \tan\theta + B \sec\theta \tag{2.7}$$

One can observe that the $r(\theta)$ is a monotonic function that passes through two positive values, therefore $r(\theta) > 0$ – the path never goes through the origin. This is a remarkable geometric result: *no geodesics passes through the vertex on a cone if* $\Theta < \pi$*!*: There always is a shorter path around the vertex.

At the other hand, if $\Theta > \pi$, then a family of the geodesics will path through the vertex and consist of two straight intervals. This happens if $\theta > \pi$. Notice that in this case the original cone, when cut, becomes a wedge with the angle larger than 2π and consist of at least two overtopping sheets.

Distance on a sphere: Columbus problem Consider the problem of geodesics on a sphere. Let us prove that a geodesics is a part of the great circle.

Suppose that geodesics is a different curve, or that it exists an arc that is a part of the geodesics but does not coincide with the arc of the great circle. This arc can be replaced with its mirror image – the reflection in the plane that passes through the ends of the arc and the center of the sphere. The reflected curve has the same length of the path and it lies on the sphere, therefore the new path remains a geodesics.

At the other hand, the new path is broken in two points, and therefore cannot be the shortest path. Indeed, consider a part of the path in an infinitesimal circle around the point of breakage and fix the points A and B where the path crosses that circle. This path can be shorten by a arc of a great circle that passes through the points A and B. To illustrate this part, it is enough to imagine a human-size scale on Earth: The infinitesimal part of the round surface becomes flat and obviously the shortest path correspond to a straight line and not to a zigzag line with an angle.

The same consideration shows that the length of geodesics is no larger than π times the radius of the sphere or it is shorter than the great semicircle. Indeed, if the length of geodesics is larger than the great semicircle one can fix two opposite points – the poles of the sphere – on the path and turn around the axis the part of geodesics that passes through these points. The new path lies of the sphere, has the same length as the original one, and is broken at the poles, thereby its length is not minimal.

To summarize *geodesics on a sphere is a part of the great circle that joins the starting and end points and which length is less that a half of the equator.*

Remark 2.1.2 This geometric consideration, when algebraically developed and generalized to larger class of extremal problems, yields to the so-called Jacobi test, see below Section 6.3. The Jacobi test is violated if the length of geodesics is larger than π times the radius of the sphere.

The argument that the solution to the problem of shortest distance on a sphere bifurcates when its length exceeds a half of the great circle was in fact famously used by Columbus who argued that the shortest way to India passes through the Western route. As we know, Columbus wasn't be able to prove or disprove the conjecture because he bumped into American continent discovering New World for better and for worst.

2.1.3 Minimal surface

A three-dimensional generalization of the geodesics is the problem of the minimal surface that is the surface of minimal area stretched on a given contour. If the contour is plane, the solution is obvious: the minimal surface is a plane. The proof is quite similar to the above proof of the minimal distance on the plane.

In general, the contour can be any closed curve in three-dimensional space; the corresponding surface can be very complicated, and nonunique. It may contain several smooth branches with nontrivial topology (see the pictures). The example of such surface is provided by a soap film stretched on a contour made from a wire: the surface forces naturally minimize the area of the film. Theory of minimal surfaces is actively developing area, see the books [**?**, **?**].

In contrast with the complexity of a minimal surface in the large scale, caused by the complexity of the supporting contour, the local feature of any minimal surface is simple; we show that any smooth segment of the minimal surface has zero mean curvature.

We prove the result using an infinitesimal (variational) approach. Let S be an optimal surface, and s_0 be a regular point of it. Assume that S is a smooth surface in the neighborhood of s_o and introduce a local Cartesian coordinate system ξ_1, ξ_2, Z so oriented that the normal to the surface at a point s_0 coincides with the axes Z. The equation of the optimal surface can locally be represented as

$$Z = D + A\xi_1^2 + 2C\xi_1\xi_2 + B\xi_2^2 + o(\xi_1^2, \xi_2^2) = 0$$

Here, the linear with respect to ξ_1 and ξ_2 terms vanish because of orientation of Z-axis. In cylindrical coordinates r, θ, Z, the equation of the surface $F(r,\theta)$ becomes

$$0 \le r \le \epsilon, \quad \pi \le \theta \le \pi,$$

and

$$F(r,\theta) = D + a\,r^2 + b\,r^2\cos(2\theta + \theta_0) + o(r^2) \tag{2.8}$$

Consider now a cylindrical ϵ-neighborhood of s_0 – a part $r \le \epsilon$ of the surface inside an infinite cylinder with the cental axes Z. The equation of the contour , – the intersection of S with the cylinder $r = \epsilon$ – is

$$,\,(\theta) = F(r,\theta)|_{r=\epsilon} = D + \epsilon^2 a + \epsilon^2 b\cos(2\theta + \theta_0) + o(\epsilon^2) \tag{2.9}$$

If the area of the whole surface is minimal, its area inside contour , is minimal among all surfaces that passes through the same contour. Otherwise, the surface could be locally changed without violation of continuity so that its area would be smaller.

In other words, the coefficients D, a, b, θ_0 of the equation (2.8) for an admissible surface should be chosen to minimize its area, subject to restrictions following from (2.9): The parameters b, θ_0 and $D + \epsilon^2 a$ are fixed. This leaves only one degree of freedom – parameter a – in an admissible smooth surface. Let us show that the optimal surface corresponds to $a = 0$.

We observe, as in the previous problem, that the surface area

$$A = \int_0^{2\pi}\int_0^{\epsilon}\left(\sqrt{1 + \left(\frac{\partial F}{\partial r}\right)^2 + \left(\frac{1}{r}\frac{\partial F}{\partial \theta}\right)^2}\right) r\,dr d\theta$$

is a strictly convex and even function of a (which can be checked by substitution of (2.9) into the formula and direct calculation of the second derivative). This implies that the minimum is unique and correspond to $a = 0$.

Another way is to use the approximation based on smallness of ϵ. The calculation of the integral must be performed up to ϵ^3, and we have

$$A = \pi\epsilon^2 + \frac{1}{2}\int_0^{2\pi}\int_0^{\epsilon}\left(\left(\frac{\partial F}{\partial r}\right)^2 + \left(\frac{1}{r}\frac{\partial F}{\partial \theta}\right)^2\right) r\,dr d\theta + o(\epsilon^3).$$

After substitution of the expression for F from (2.8) into this formula and calculation, we find that

$$A = \pi\epsilon^2 + \frac{8}{3}\pi\epsilon^3(a^2 + b^2) + 0(\epsilon^3)$$

The minimum of A corresponds to $a = 0$ as stated. Geometrically, the result means that the mean curvature of a minimal surface is zero in any regular point. The minimal surface area

$$A_{\min} = \pi\epsilon^2 + \frac{8}{3}\pi\epsilon^3 b^2 + 0(\epsilon^3)$$

depends only on the total variation $2b = (\max, \ - \min,)$ of , as expected.

In addition, notice that the minimal area between all surfaces enclosed in a cylinder that do not need to pass through a fixed contour is equal to the area $\pi\epsilon^2$ of a circle and corresponds to a flat contour $b = 0$, as expected.

Proof by symmetry Another proof does not involve direct calculation of the surface. We only states that the minimal surface S locally is entirely determined by the infinitesimal contour , . Therefore, a transform of the coordinate system that keeps the contour unchanged cannot change the minimal surface inside it. Observe, that the infinitesimal contour (2.9) is invariant to transform

$$Z' = -Z + 2(D + \epsilon^2 a), \quad r' = r, \quad \theta' = \theta + 90^\circ. \tag{2.10}$$

that consists of reverse of the direction of Z axes, shift along Z, and rotation on 90° across this axes. The minimal surface (2.8) must be invariant to this transform as well, which again gives $a = 0$.

Remark 2.1.3 This proof assumes uniqueness of the minimal surface.

Thin film model The equation of the minimal surface can be deduced from the model of a thin film as well. Assume that the surface of the film shrinks by the inner tangent forces inside each infinitesimal element of it, and there are no bending forces generated that is forces normal to the surface. The tangent forces at a point depend only on local curvatures at this point.

Separate again the cylindrical neighborhood and replace the influence of the rest of the surface by the tangential forces applied to the surface at each point of the contour. Consider conditions or equilibrium of these forces and the inner tangent forces in the film. First, we argue that the average force applied to the contour is zero. This force must be directed along the z-axes, because the contour is invariant to rotation on 180° degree around this axes. If the average force (that depends only on the geometry) had a perpendicular to z component, this component would change its sign. The z-component of the average force applied to the contour is zero too, by the virtue of invariance of the transform (2.10). By the equilibrium condition, the average z-component of the tangent force inside the surface element must be zero as well. Look of the representation (2.8) of the surface: The average over the area force depends on a and b: $F = F(a, b, \theta_0)$. The force is in fact independent of θ_0, because of symmetry; The dependence on b is even, because the change of sign of b corresponds to 90° rotation of the contour that leaves the force unchanged.

The dependence on a is odd, because the change of the direction of the force correspond to change of the sign of a.

$$F = \text{constant}(\theta_0), \quad F(a,b) = F(a,-b) = -F(-a,b) \quad \forall \theta_0, a, b.$$

Therefore, zero average force corresponds to $a = 0$, as stated.

The direction of average along the contour and over the surface forces cannot depend on b because the 180° degree rotation of the contour leaves is invariant, therefore the force remains invariant, too.

2.1.4 Shortest path around an obstacle: Convex envelope

A helpful tool in the theory of extremal problem is the convex envelope. Here, we introduce the convex envelope of a finite set in a plane as the solution of a variational problem about the minimal path around an obstacle. The problem is to find the shortest closed contour that contains finite not necessarily connected domain Ω inside. This path is called the convex envelope of the set Ω.

Definition 2.1.3 (Convex envelope of a set) The convex envelope $\mathcal{C}\Omega$ of a finite closed set Ω is the minimal of the sets that (i) contain Ω inside, $\mathcal{C}\Omega \supset \Omega$ and (ii) is convex.

We argue that the minimal path , is convex, that is every straight line intersects its boundary not more than twice. Indeed, if a component is not convex, we may replace a part of it with a straight interval that lies outside of , thus finding another path , $'$ that encircles a larger set but has a smaller perimeter. Perimeter of a convex set is decreased only when the encircled set , is lessen.

Also, the strictly convex (not straight) part of the path coincides with the boundary of Ω. Otherwise, the length of this boundary can be decreased by replacing an arc of it with the chord that lies completely outside of Ω.

We demonstrated that a convex envelope consists of at most two types of lines: the boundary of Ω and straight lines (shortcuts). The convex envelope of a convex set coincide with it, and the convex envelope of the of the set of finite number of points is a convex polygon that is supported by some of the points and contains the rest of them inside.

Properties of the convex envelope The following properties are geometrically obvious and the formal proofs of then are left to the interested reader.

1. Envelope cannot be further expanded.

$$\mathcal{C}(\mathcal{C}(\Omega)) = \mathcal{C}(\Omega)$$

2. Conjunction property:

$$\mathcal{C}(\Omega_1 \cup \Omega_2) \supseteq \mathcal{C}(\Omega_1) \cup \mathcal{C}(\Omega_2)$$

3. Absorbtion property: If $\Omega_1 \subset \Omega_2$ then

$$\mathcal{C}(\Omega_1 \cup \Omega_2) = \mathcal{C}(\Omega_2)$$

4. Monotonicity: If $\Omega_1 \subset \Omega_2$ then

$$\mathcal{C}(\Omega_2) \subseteq \mathcal{C}(\Omega_1)$$

Shortest trajectory in a plane with an obstacle
Find the shortest path $p(A, B, \Omega)$ between two points A and B on a plane if a bounded connected region (an obstacle) Ω in a plane between them cannot be crossed.

- First, split a plane into two semiplanes by a straight line that passes through the connecting points A and B.

- If the interval between A and B does not connect inner points of Ω, this interval is the shortest pass. In this case, the constraint (the presence of the obstacle) is *inactive*, $p(A, B, \Omega) = \|A - B\|$ independently of Ω.

- If the interval between A and B connects inner points of Ω, the constraint becomes *active*. In this case, obstacle Ω is divided into two parts Ω_+ and Ω_- that lie in the upper and the lower semiplanes, respectively, and have the common boundary along the divide – an interval ∂_0; ∂_0 lies inside the original obstacle Ω.

  Because of the connectedness of the obstacle, the shortest path lies entirely either in the upper or lower semiplane, but not in both; otherwise, the path would intersect ∂_0. We separately determine the shortest path in the upper and lower semiplanes and choose the shortest of them.

- Consider the upper semiplane. Notice that points A and B lie on the boundary of the convex envelope $\mathcal{C}(\Omega_+, A, B)$ of the set Ω and the connecting points A and B.

  The shortest path in the upper semiplane $p_+(A, B, \Omega)$ coincides with the upper component of the boundary of $\mathcal{C}(\Omega_+, A, B)$, the component that does not contains ∂_0. It consists of two straight lines that pass through the initial and final points of the trajectory and are tangents to the obstacle, and a part that passes along the boundary of the convex envelope $\mathcal{C}\Omega$ of the obstacle only.

- The path in the lower semiplane is considered similarly. Points A and B lie on the boundary of the convex envelope $\mathcal{C}(\Omega_-, A, B)$. Similarly to the shortest path in the upper semiplane, the shortest path in the lower semiplane $p_-(A, B, \Omega)$ coincides with the lower boundary of $\mathcal{C}(\Omega_-, A, B)$.

- The optimal trajectory is the one of the two pathes $p_+(A, B, \Omega)$ and $p_-(A, B, \Omega)$; the one with smaller length.

> Analytical methods cannot tell which of these two trajectories is shorter, because this would require comparing of non-close-by trajectories; a straight calculation is needed.

If there is more than one obstacle, the number of the competing trajectories quickly raises.

Convex envelope supported at a curve Consider a slightly different problem: Find the shortest way between two points around the obstacle assuming that the these points lie on a curve that passes through the obstacle on the opposite sides of it. The points are free to move along the curve it this would decrease the length of the path. Comparing with the previous problem, we asking in addition where the points A and B are located. The position of the points depends on the shape of the obstacle and the the curve, but it is easy to establish the conditions that must be satisfied at optimal location.

Problem: Show that an optimal location of the point A is either on the point of intersection of the line and an obstacle, or the optimal trajectory $p_-(A, B, \Omega)$ has a straight component near the point A and this component is perpendicular to the line at the point A.

Lost tourists

Finally, we consider a variation of the theme of convex envelope, the problem of the lost tourists. Crossing a plain, tourists have lost their way in a mist. Suddenly, they find a pole with a message that reads: "A straight road is a mile away from that pole." The tourists need to find the road; they are shortsighted in the mist: They can can see the road only when they step on it. What is the shortest way to the road even if the road is most inauspiciously located?

The initial guess would suggest to go straight for a mile in a direction, then turn 90°, and go around along the one-mile-radius circumference. This route meets any straight line that is located at the one mile distance from the central point. The length of this route is $1 + 2\pi \approx 7.283185$ miles.

However, a detailed consideration shows that this strategy is not optimal. Indeed, there is no need to intersect each straight line (the road) at the point of the circle but at any point and the route does not need to be closed. Any route that starts and ends at two points A and B at a tangent to a circle and goes around the circle intersects all other tangents to that circle. In other words, the convex envelope of the route includes a unit circle. The problem becomes: Find the curve that begins and ends at a tangent AB to the unit circle, such that (i) its convex envelope contains a circle and (ii) its length plus the distance $0A$ from the middle of this circle to one end of the curve is minimal.

The optimal trajectory consists of an straight interval OA that joints the central point O with a point A outside of the circle C and the convex envelope (ACB) stretched on the two points A and B and circle C.

The boundary of the convex envelope is either straight or coincide with the circle. More exactly, it consists of two straight intervals $A\,A_1$ supported by the point A and a point A_1 at the circumference and $A\,A_1$ supported by the end point B and a point B_1 of circumference. These intervals are tangent to the circumference at the points A_1 and B_1, respectively. Finally, line AB touches the circumference a point V.

Calculation The length L of the trajectory is

$$L = \mathcal{L}(O\,A) + \mathcal{L}(A\,A_1) + \mathcal{L}(A_1\,B_1) + \mathcal{L}(B_1\,B)$$

where $\mathcal{L}$ is the length of the corresponding component. These components are but straight lines and an circle's arch; the problem is thus parameterized. To compute the trajectory, we introduce two angles α and $-\beta$ from the point V there the line AB touches the circle. Because of symmetry, the points A_1 and B_1 correspond to the angles 2α and -2β, respectively, and we compute

$$\begin{aligned} \mathcal{L}(O\,A) &= \frac{1}{\cos\alpha}, \quad \mathcal{L}(A\,A_1) = \tan\alpha \\ \mathcal{L}(A_1\,B_1) &= 2\pi - 2\alpha - 2\beta, \quad \mathcal{L}(B_1\,B) = -\tan\beta, \end{aligned}$$

plug these expressions into the expression for L, solve the conditions $\frac{dL}{d\alpha} = 0$ and $\frac{dL}{d\beta} = 0$, and find optimal angles:

$$\alpha = \frac{\pi}{6}, \quad \beta = -\frac{\pi}{4},$$

The minimal length L equal to $L = \frac{7}{6}\pi + \sqrt{3} + 1 = 6.397242$.

Solution without calculation One could find solution to the problem without any trigonometry but with a bit of geometric imagination. Consider the mirror image C_m of the circle C assuming that the mirror is located at the tangent AB. Assume that the optimal route goes around that image instead of original circle; this assumption evidently does not change the length of the route. This new route consists of three pieces instead of four: The straight line $O\,A'_m$ that passes through the point O and is tangent to the circumference C_m, the part $A'_m\,B'_m$ of this circumference, and the straight line $B'_m B$ that passes through a point B on the line and is tangent to the circumference C_m.

The right triangle $O_m A'_m O$ has the hypothenuse $O'O$ equal to two and the side $O_m A'_m$ equal to one; the length of remaining side OA' equals to $\sqrt{3}$ and the angle $O_m O\,A'_m$ is $\frac{\pi}{3}$. The line $B'_m B$ is perpendicular to AB, therefore its length equals one. Finally, the angle of the arch $A'_m B'_m$ equals to $\frac{7}{6}\pi$. Summing up, we again obtain $L = \frac{7}{6}\pi + \sqrt{3} + 1$.

Generalization The generalization of the concept of convex envelope to the three-dimensional (or multidimensional) sets is apparent. The problem asks for set of minimal surface area that contains a given closed finite set. The solution is again given by the convex envelope, definition (2.1.4) is applicable for the similar reasons.

Consider the three-dimensional analog of the problem 2.1.4 assuming in addition that the obstacle Ω is convex. Repeating the arguments for the plane problem, we conclude that the optimal trajectory belongs to the convex envelope $\mathcal{C}(\Omega, A, B)$. The envelope is itself a convex surface and therefore the problem is reduced to geodesics on the convex set – the envelope $\mathcal{C}(\Omega, A, B)$. The variational analysis of this problem allows to disqualify as optimal all (or almost all) trajectories on the convex envelope one by comparing near-by trajectories that touch the obstacle in close-by points.

If the additional assumption of convexity of obstacle is lifted, the problem becomes much more complex because the passes through "tunnels" and in folds in the surface of Ω should be accounted for. If at least one of the points A or B lies inside the convex envelope of a nonconvex obstacle, the minimal path partly goes inside the convex envelope $\mathcal{C}\Omega$ as well. We leave this for the interested reader.

2.1.5 Formalism of convex envelopes

The notion of convex envelope can be transformed from sets to functions. A graph of any function $y = f(x)$ divides the space into two sets, and the convex envelope of a function is the convex envelope of the set $y > f(x)$. It the function is not defined for all $x \in R^n$ (like $\log x$ is defined only for $x \geq 0$), we extend the definition of a function assigning the improper value $+\infty$ to function of in all undefined values arguments.

Definition 2.1.4 (Convex envelope of a function) The convex envelope $\mathcal{C}f(x)$ of a function $f : R^n \to R^1$ is the maximal of the functions $g(x)$ that (i) do not surpass $f(x)$ everywhere, $g(x) \leq f(x), \forall x$ and (ii) is convex.

The Jensen's inequality produces the following definition of the convex envelope:

Definition 2.1.5 The *convex envelope* $\mathcal{C}F(v)$ is a solution to the following minimal problem:

$$\mathcal{C}F(\boldsymbol{v}) = \inf_{\boldsymbol{\xi}} \frac{1}{l} \int_0^l F(\boldsymbol{v} + \boldsymbol{\xi}) dx \quad \forall\, \boldsymbol{\xi} : \int_0^l \boldsymbol{\xi} dx = 0. \tag{2.11}$$

This definition determines the convex envelope as the minimum of all parallel secant hyperplanes that intersect the graph of F; it is based on Jensen's inequality (**??**).

To compute the convex envelope $\mathcal{C}F$ one can use the Carathéodory theorem (see [**?**, **?**]). It states that the argument $\boldsymbol{\xi}(x) = [\xi_1(x), \dots, \xi_n(x)]$ that minimizes the right-hand side of (2.11) takes no more than $n + 1$ different values. This theorem refers to the obvious geometrical fact that the convex envelope consists of the supporting hyperplanes to the graph $F(\xi_1, \dots, \xi_n)$. Each of these hyperplanes is supported by no more than $(n + 1)$ arbitrary points.

The Carathéodory theorem allows us to replace the integral in the right-hand side of the definition of $\mathcal{C}F$ by the sum of $n+1$ terms; the definition (2.11) becomes:

$$\mathcal{C}F(\boldsymbol{v}) = \min_{m_i \in M} \min_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}} \left\{ \sum_{i=1}^{n+1} m_i F(\boldsymbol{v} + \boldsymbol{\xi}_i) \right\}, \tag{2.12}$$

where

$$M = \left\{ m_i : \quad m_i \geq 0, \quad \sum_{i=1}^{n+1} m_i = 1 \right\} \tag{2.13}$$

and

$$\boldsymbol{\Xi} = \left\{ \boldsymbol{\xi}_i : \quad \sum_{i=1}^{n+1} m_i \boldsymbol{\xi}_i = 0 \right\}. \tag{2.14}$$

The convex envelope $\mathcal{C}F(\boldsymbol{v})$ of a function $F(\boldsymbol{v})$ at a point $\boldsymbol{v}$ coincides with either the function $F(\boldsymbol{v})$ or the hyperplane that touches the graph of the function F. The hyperplane remains below the graph of F except at the tangent points where they coincide.

The position of the supporting hyperplane generally varies with the point $\boldsymbol{v}$. A convex envelope of F can be supported by fewer than $n+1$ points; in this case several of the parameters m_i are zero.

On the other hand, the convex envelope is the greatest convex function that does not exceed $F(\boldsymbol{v})$ in any point $\boldsymbol{v}$ [?]:

$$\mathcal{C}F(\boldsymbol{v}) = \max \phi(\boldsymbol{v}) : \ \phi(\boldsymbol{v}) \leq F(\boldsymbol{v}) \ \forall \boldsymbol{v} \quad \text{and } \phi(\boldsymbol{v}) \text{ is convex.} \tag{2.15}$$

Example 2.1.4 Obviously, the convex envelope of a convex function coincides with the function itself, so all m_i but m_1 are zero in (2.12) and $m_1 = 1$; the parameter $\boldsymbol{\xi}_1$ is zero because of the restriction (2.14).

The convex envelope of a "two-well" function,

$$\Phi(\boldsymbol{v}) = \min \left\{ F_1(\boldsymbol{v}), \ F_2(\boldsymbol{v}) \right\}, \tag{2.16}$$

where $F_1, \ F_2$ are convex functions of $\boldsymbol{v}$, either coincides with one of the functions $F_1, \ F_2$ or is supported by no more than two points for every $\boldsymbol{v}$; supporting points belong to different wells. In this case, formulas (2.12)–(2.14) for the convex envelope are reduced to

$$\mathcal{C}\Phi(\boldsymbol{v}) = \min_{m, \boldsymbol{\xi}} \left\{ m F_1(\boldsymbol{v} - (1-m)\boldsymbol{\xi}) + (1-m) F_2(\boldsymbol{v} + m\boldsymbol{\xi}) \right\}. \tag{2.17}$$

Indeed, the convex envelope touches the graphs of the convex functions F_1 and F_2 in no more than one point. Call the coordinates of the touching points $\boldsymbol{v} + \boldsymbol{\xi}_1$ and $\boldsymbol{v} + \boldsymbol{\xi}_2$, respectively. The restrictions (2.14) become $m_1 \boldsymbol{\xi}_1 + m_2 \boldsymbol{\xi}_2 = 0$, $m_1 + m_2 = 1$. It implies the representations $\boldsymbol{\xi}_1 = -(1-m)\boldsymbol{\xi}$ and $\boldsymbol{\xi}_2 = m\boldsymbol{\xi}$.

Example 2.1.5 Consider the special case of the two-well function,

$$F(v_1, v_2) = \begin{cases} 0 & \text{if } v_1^2 + v_2^2 = 0, \\ 1 + v_1^2 + v_2^2 & \text{if } v_1^2 + v_2^2 \neq 0. \end{cases} \quad (2.18)$$

The convex envelope of F is equal to

$$\mathcal{C}F(v_1, v_2) = \begin{cases} 2\sqrt{v_1^2 + v_2^2} & \text{if } v_1^2 + v_2^2 \leq 1, \\ 1 + v_1^2 + v_2^2 & \text{if } v_1^2 + v_2^2 > 1. \end{cases} \quad (2.19)$$

Here the envelope is a cone if it does not coincide with F and a paraboloid if it coincides with F.

Indeed, the graph of the function $F(v_1, v_2)$ is rotationally symmetric in the plane $v_1,\ v_2$; therefore, the convex envelope is symmetric as well: $\mathcal{C}F(v_1, v_2) = f(\sqrt{v_1^2 + v_2^2})$. The convex envelope $\mathcal{C}F(\boldsymbol{v})$ is supported by the point $\boldsymbol{v} - (1-m)\boldsymbol{\xi} = \mathbf{0}$ and by a point $\boldsymbol{v} + m\boldsymbol{\xi} = \boldsymbol{v}^0$ on the paraboloid $\phi(\boldsymbol{v}) = 1 + v_1^2 + v_2^2$. We have

$$\boldsymbol{v}^0 = \frac{1}{1-m}\boldsymbol{v}$$

and

$$\mathcal{C}F(\boldsymbol{v}) = \min_m \left\{ (1-m)\phi\left(\frac{1}{1-m}\boldsymbol{v}\right)\right\}. \quad (2.20)$$

The calculation of the minimum gives (2.19).

Example 2.1.6 Consider the nonconvex function $F(v)$ used in Example **??**:

$$F(v) = \min\{(v-1)^2, (v+1)^2\}.$$

It is easy to see that the convex envelope $\mathcal{C}F$ is

$$\mathcal{C}F(v) = \begin{cases} (v+1)^2 & \text{if } v \leq -1, \\ 0 & \text{if } v \in (-1, 1), \\ (v-1)^2 & \text{if } v \geq 1. \end{cases}$$

Example 2.1.7 Compute convex envelope for a more general two-well function:

$$F(v) = \min\{(av)^2, (bv+1)^2\}.$$

The envelope $\mathcal{C}F_n(v)$ coincides with either the graph of the original function or the linear function $l(v) = A\,v + B$ that touches the original graph in two points (as it is predicted by the Carathéodory theorem; in this example $n = 1$). This function can be found as the common tangent $l(v)$ to both convex branches (wells) of $F(v)$:

$$\begin{cases} l(v) = av_1^2 + 2av_1(v - v_1), \\ l(v) = (bv_2^2 + 1) + 2bv_2(v - v_2), \end{cases} \quad (2.21)$$

where v_1 and v_2 belong to the corresponding branches of F_p:

$$\begin{cases} l(v_1) = av_1^2, \\ l(v_2) = bv_2^2 + 1. \end{cases} \quad (2.22)$$

Solving this system for $v,\ v_1,\ v_2$ we find the coordinates of the supporting points

$$v_1 = \sqrt{\frac{b}{a(a-b)}}, \quad v_2 = \sqrt{\frac{a}{b(a-b)}}, \tag{2.23}$$

and we calculate the convex envelope:

$$\mathcal{C}F(v) = \begin{cases} av^2 & \text{if } |v| < v_1, \\ 2v\sqrt{\frac{ab}{a-b}} - \frac{b}{a-b} & \text{if } v \in [v_1,\ v_2], \\ 1 + bv^2 & \text{if } |v| < v_2 \end{cases} \tag{2.24}$$

that linearly depends on v in the region of nonconvexity of F.

Hessian of Convex Envelope We mention here a property of the convex envelope that we will use later. If the convex envelope $\mathcal{C}F(\boldsymbol{v})$ does not coincide with $F(\boldsymbol{v})$ for some $\boldsymbol{v} = \boldsymbol{v}_n$, then $\mathcal{C}F(\boldsymbol{v}_n)$ is convex, but not strongly convex. At these points the Hessian $He(F) = \frac{\partial^2}{\partial v_i \partial v_j} F(\boldsymbol{v})$ is semipositive; it satisfies the relations

$$He(\mathcal{C}F(\boldsymbol{v})) \geq 0, \quad \det H(\mathcal{C}F(\boldsymbol{v})) = 0 \quad \text{if } \mathcal{C}F < F, \tag{2.25}$$

which say that $He(\mathcal{C}F)$ is a nonnegative degenerate matrix. These relations can be used to compute $\mathcal{C}F(\boldsymbol{v})$.

Remark 2.1.4 (Convex envelope as second conjugate) We may as well compute convex envelope in more regular way as a second conjugate of the original function as described later in Section 8.3.

Proof that

(i) The global minimum of a function coincides with the minimum of its convex envelope.

(ii) Convex envelope of a function does not have minima that are local but not global.

Convex envelope are used below in Chapter 7 to address ill-posed variational problems.

2.2 Symmetrization

An interesting geometric method, symmetrization, is based on convexity inequality; it allows for solution of several isoperimetric problem. The detailed discussion can be found in the books by Blaschke [], Pólya and Szegö [**?**]

The idea of symmetrization Consider a plane finite domain Ω and a straight line A. The transformation of Ω is called a symmetrization with respect to A if it moves each interval that crosses Ω and is orthogonal to A parallel to itself so that the middle of the interval belongs to A.

One can easily see that the symmetrization of a polygon is a polygon with equal or larger number of angles than the original one.

2.2.1 Symmetrization of a triangle

Let us prove that unilateral triangle has minimal perimeter among all triangles with given area.

Consider an arbitrary nonunilateral triangle ABC and apply symmetrization to it. Generally, the symmetrization transforms a triangle into a quadrangle; the triangle remains a triangle only if the axis of symmetrization is orthogonal to one of the side. In this case, an arbitrary triangle becomes an isosceles triangle, the base a and the hight h remain unchanged. This implies that symmetrization leaves the area A of the triangle unchanged.

Let us show that symmetrization decreases the perimeter. Let the coordinates of the vertexes be

$$A = (a, 0), \quad B = (-a, 0), \quad C = (c, h)$$

and let the axes of symmetrization be the Y axes. After symmetrization, the coordinates A and B remain the same, and the vertex C moves to $C' = (0, h)$.

The sum of the two sides' lengths equal to

$$L = \sqrt{h^2 + (a - c)^2} + \sqrt{h^2 + (a + c)^2}$$

becomes

$$L_S = 2\sqrt{h^2 + a^2}.$$

We prove that

$$L_S \leq L \tag{2.26}$$

and the equality sign corresponds to the only case $c = 0$.

Consider the length of a side as a function of c:

$$L(c) = \sqrt{h^2 + (a + c)^2}$$

the function f is strictly convex since

$$L''(c) = \frac{h^2}{(h^2 + (a + c)^2)^{\frac{3}{2}}} > 0$$

and is even. The inequality of convexity (13.21) implies that

$$L(c) + L(-c) \geq 2f(0)$$

that is, the inequality (2.26)

If the obtained rectangle is not unilateral, the symmetrization procedure can be repeated, using one of the equal legs as the base. The new triangle has the same area and smaller perimeter.

Consider now a sequence of symmetrizations applied to an arbitrary triangle. On each step, symmetrization preserves one side, makes two other sides equal to each other, and decreases their total length. The area of the triangle is preserved, its perimeter decreases and is obviously bounded from below, say by zero. Therefore the sequence of symmetrizations is a monotone bounded sequence and it must have a unique stable point: A triangle that is stable against symmetrization. This is of course a unilateral triangle. We have proved the theorem:

Theorem 2.2.1 Among all triangles with equal area, the unilateral triangle has the smallest perimeter.

2.2.2 Symmetrization of quadrangle and circle

Symmetrization of a quadrangle Let is apply symmetrization to an arbitrary quadrangle, requiring that the quadrangle remains quadrangle after the symmetrization.

At the first step, we have to perform symmetrization orthogonal to one of two diagonals. The resulting quadrangle has two pairs of neighboring sides of equal lengths. At the second step, we symmetrize orthogonally to the other diagonal, the resulting figure is a rhombus of equal area but of smaller perimeter than the original quadrangle.

Now we may start a two-steps sequence of symmetrizations. Firstly, we transform the rhombus into a rectangle using the side as an axis of symmetrization. Secondly, we transform the rectangle back to rhombus, using the diagonal as the axis of symmetrization. The obtained rhombus has smaller ratio of the larger diagonal to the smaller one (compute the change of this ratio!) and the smaller perimeter, but its area stays unchanged. This monotonic sequence has a stable point. The stable point is the square, which enables us to formulate the next theorem.

Theorem 2.2.2 Among all quadrangles with equal area, the square has the smallest perimeter.

Circle Symmetrization can be applied to an arbitrary finite bounded domain $F(x, y) \geq 0$ with the boundary $F(x, y) = 0$. For definiteness let assume that the y-axis is the axis of symmetrization.

Dissect the plane by a family $\{y_k\}$ of equidistant parallel lines

$$y_0, \quad y_1 = y_0 + \Delta, \quad y_2 = y_0 + 2\Delta, \ldots, \quad y_N = y_0 + N\Delta$$

. Assume that this division covers the figure $F(x, y) = 0$ and that the number N is arbitrary large so that the distance between two neighboring parallel lines is infinitesimal.

An infinitesimal part of the domain $F(x, y) = 0$ located between two close-by parallel lines can be approximated by a trapezoid. Symmetrization replaces this trapezoid by a equilateral trapezoid of equal area, parallel sides of equal length, but with smaller total length of the non-parallel sides (show this!). We can formulate

Theorem 2.2.3 The total area of the symmetrized domain remains constant, but its perimeter (equal to the sum of the lengths of the sides of the trapezoids) decreases.

Now consider the sequence of symmetrization with variable axis. The sequence of the transformed figures tends to a circle: The only figure that is stable against any symmetrization. Indeed, this sequence tends to its unique stable point, and the circle is that point.

We came to the theorem

Theorem 2.2.4 Among all plane domains with equal finite area, the circle has the smallest perimeter.

Geometric proof of the theorem An independent geometric proof of the theorem is elegant and does not require any infinitesimal operation. However, we need to assume the existence of an optimal shape which we do not need to do in the previous consideration.

The proof requires the following steps:

1. We show that the optimal domain is convex. If it is not convex, we pass to the convex envelope increasing the area and decreasing the perimeter at the same time.

2. We cut the optimal domain by a straight line so that both parts have the same area. This is always possible by moving a line across the domain and keeping it parallel to itself. The two cut parts must have the same perimeter too, otherwise the perimeter could be decreased replacing the part with larger perimeter with the mirror image of the part with smaller perimeter. The replacement of one of the domain with the mirror image of the second one changes neither the area nor the perimeter.

3. Consider the half of the optimal shape with the straight base. Choose an arbitrary point C on it surface and connect it with the ends of the base by two straight intervals. The domain is thus divided into two outer shapes and a triangle.

4. We may change the length of the base without changing the perimeter. This change keeps the areas of two outer domains constant but varies the angle and by the area of the triangle. The maximal area of the triangle corresponds to the angle C opposite the base being equal $90°$. Indeed, by the geometric theorem the area A equals to

$$A = \frac{1}{2} a\, b \, \sin C$$

where the lengths a and b of the intervals are constant to the motion and the angle arbitrary varies. The maximal area corresponds to $C = 90°$.

5. Because the point C was arbitrarily chosen, the angle between any point of the surface and the base is equal $90°$. The figure must be a circle: the set of points from which an interval (diameter) is visible on a right angle.

2.2.3 Dido problem

Probably the first extremal problem known from the antic time is the Dido problem. The problem is based on a passage from Virgil's Aeneid (cited from []):

> "The Kingdom you see is Carthage, the Tyrians, the town of Agenor;
> "But the country around is Libya, no folk to meet in war.
> Dido, who left the city of Tyre to escape her brother,
> Rules here – a long a labyrinthine tale of wrong
> Is hers, but I will touch on its salient points in order
> ...
> Dido, in great disquiet, organized her friends for escape.
> They met together, all those who harshly hated the tyrant
> Or keenly feared him: they seized some ships which chanced to be ready
> ...
> They came to this spot, where to-day you can behold the mighty
> Battlements and the rising citadel of New Carthage,
> And purchased a site, which was named 'Bull's Hide' after the bargain
> By which they should get as much land as they could enclose with a bull's
> hide."

According to the legend, the Trojans arrived in the North African shore of Mediterranean Sea after the defeat by Greeks. Here, their leader, wise queen Dido, purchased from the local tribe a piece land on the shore "that can be covered by the Bull's hide." Sophisticated Trojans had a much more advanced technology than the locals; in particular, they knew how to use sharp knifes to cut hides into thin strips (and they knew some math, too!). So, they made a long leather rope out of the hide and encircled by it enough land to build the Carthage who later become a mighty rival of Rome.

The extremal problems that Dido brilliantly solved was: Given a curve of a given length (the rope) and a straight line (the sea shore), encircle the domain of maximal area (place for future Carthage). This problem, known as *Dido problem*, inspirited many generations by its cleverness; it influenced the development of theory of extremal problems, demonstrated usefulness of mathematics, and accustomed people to respect political leaders able to use brains instead of brutal force.

Dido problem can be solved by symmetrization together with the following trick: Assume that the seashore is a mirror and consider the domain Ω of the enclosed land and its mirror image; obviously, the perimeter and area of Ω is twice larger than the perimeter and area of the enclosed domain, respectively. The symmetrization tells that Ω is a circle; thereby, the answer to Dido problem is a semicircle with the shore as a diameter and the rope as a semi-circumference.

The reference of how to use Maple to work on Dido problem:

http://www.mapleapps.com/powertools/engineeringmath/html/Section

2.2.4 Formalism of symmetrization

The considered symmetrization of a plane domain can be formalized as following: Assume for simplicity that the boundary of the set F is y-simple: The set $F(x, y) \geq 0$ described as

$$f_-(x) \leq y \leq f_+(x), \quad a \leq x \leq b$$

The area A of the domain is equal to

$$A = \int_a^b (f_+ - f_-) dx \tag{2.27}$$

and the perimeter P is

$$P = \int_a^b \left(\sqrt{1 + (f'_+)^2} + \sqrt{1 + (f'_-)^2} \right) dx$$

The symmetrized domain is described as

$$-\frac{1}{2}\left(f_+(x) - f_-(x)\right) \leq y \leq \frac{1}{2}\left(f_+(x) - f_-(x)\right), \quad a \leq x \leq b$$

Its area A_S of the symmetrized domain is obviously given by the formula (2.27) and its perimeter P_S is

$$P_S = 2 \int_a^b \sqrt{1 + \frac{1}{4}(f'_+ - f'_-)^2}\, dx$$

If remains to prove that $P_S \leq P$ or

$$\int_a^b \left(\sqrt{1 + (f'_+)^2} + \sqrt{1 + (f'_-)^2} - 2\sqrt{1 + \frac{1}{4}(f'_+ - f'_-)^2} \right) dx \geq 0$$

We show that the integrant is nonnegative in each point. Starting with the inequality

$$\sqrt{1 + (f'_+)^2} + \sqrt{1 + (f'_-)^2} \geq 2\sqrt{1 + \frac{1}{4}(f'_+ - f'_-)^2}$$

we square its left- and right-hand sides, cancel equal terms, and obtain an equivalent inequality

$$\sqrt{\left(1 + (f'_+)^2\right)\left(1 + (f'_-)^2\right)} \geq 1 - f'_+ f'_-$$

If the right-hand side is negative, the inequality is true, otherwise square it one more time and obtain the true equivalent inequality

$$(f'_+ + f'_-)^2$$

The result is proved.

3D symmetrization

Consider a bounded body

$$F(x, y, z) \geq 0$$

in three-dimensional space with the boundary

$$F(x, y, z) = 0.$$

Dissect it by a family of equidistant parallel planes

$$z = z_0, z = z_0 + \Delta, z = z_0 + 2\Delta, \ldots, z = z_0 + N\Delta.$$

Replace a part of the body located between two planes by a conical surface, replacing each closed contour $F(x, y, z_0 + k\Delta) = 0$ by the circle of equal area, all centered at the z-axis

$$x^2 + y^2 = r_k^2, \quad \text{where } \pi r_k^2 = \text{Area of } F(x, y, z_0 + k\Delta)$$

Doing this, we obtain a body of revolution defined by the curve $r(z)$ that revolves around the z-axis.

We can show (do it yourself or look into [?]) that this transformation (symmetrization by Schwartz) (i) Conserves the volume of the body and (ii) decreases its surface area.

Particularly, consider the domain bounded by the plane $z = 0$ and a non-negative surface $z = u(x, y) \geq 0$ such that $u(x, y) = 0$ if $(x, y) \in \,,\ = \partial\Omega$. The symmetrization

1. Replaces the base Ω with a circle of equal area:

$$\Omega_S = \text{ A circle: } |\Omega| = |\Omega_S|$$

2. Conserves the volume:

$$\int_\Omega u\, dx\, dy \quad \text{is stable to symmetrization}$$

3. Decreases the surface area:

$$\int_\Omega \sqrt{1 + (\nabla u)^2}\, dx\, dy \quad \text{decreases by symmetrization} \tag{2.28}$$

Using symmetrization, we may deduct some inequalities for the functionals different from the volume or the area. For example, assuming that $u(x, y) \ll 1$, we notice that (2.28) implies the decrease of the Dirichlet integral:

$$\int_\Omega (\nabla u)^2\, dx\, dy \quad \text{decreases by symmetrization}$$

Extremal property of the sphere As in two-dimensional case, one applies the series of symmetrization around all axes, look into the resulting stable point and arrive at the theorem:

Theorem 2.2.5 Among all three-dimensional bodies with equal finite volume, the sphere has the smallest surface area.

Limits of the method

The method of symmetrization operates with special type of functionals (area, perimeter, volume).

It cannot handle any additional constraints besides the fixed area, such the requirement that a part of the boundary stays unchanged. In particular, it does not preserve the number of edges in polygons of more than fourth order.

2.2.5 Summary

The sufficient conditions are the most elegant statements in the theory extremal problems. In these methods, the guessed optimal solution is directly compared with all admissible solutions; thus the global optimum of the functional is proven. By its nature, a sufficient conditions technique is irregular and the area of its applicability is limited.

Symmetrization shows that is many problem a symmetric solution is better than a nonsymmetric one. This principle is reflected in an intuitive preference to symmetric designs which are often considered to be more elegant or beautiful that nonsymmetric ones.

2.3 Problems

1. Use Jensen inequality to prove the relation between arithmetic and harmonic means:

$$\frac{a_1 + \ldots + a_N}{N} \geq (a_1 \cdot \ldots \cdot a_N)^{\frac{1}{N}} \quad \forall a_1 \geq 0, \ldots a_N \geq 0$$

2. Describe the area of a symmetrized ellipse.

Part II

Calculus of Variations: One variable

Chapter 3

Stationarity

Since, however, the rules ⟨for isoperimetric curves (or, in our words, extremal problems)⟩ were not sufficiently general, the famous Euler undertook the task of reducing all such investigations to a general method which he gave in the work "Essay on a new method of determining the maxima and minima of indefinite integral formulas"; an original work in which the profound science of the calculus shines through. Even so, while the method is ingenious and rich, one must admit that it is not as simple as one might hope in a work of pure analysis.

In "Essay on a new method of determining the maxima and minima of indefinite integral formulas", by Lagrange, 1760

3.1 Derivation of Euler equation

The technique was developed by Euler, who also introduced the name "Calculus of variations" in 1766. The method is based on an analysis of infinitesimal variations of a minimizing curve.

The main scheme of the variational method is as follows: Assume that the optimal curve $u(x)$ exist among smooth (twice-differentiable curves), $u \in C_2[a,b]$. Compare the optimal curve with close-by trajectories $u(x)+\delta u(x)$, where $\delta u(x)$ is small in some sense. Using the smallness of δu, we simplify the comparison, deriving necessary conditions for the optimal trajectory $u(x)$ Variational methods yield to only necessary conditions of optimality because it is assumed that the compared trajectories are close to each other; on the other hand, they are applicable to a great variety of extremal problems called *variational problems*.

3.1.1 Euler equation (Optimality conditions)

Consider the problem called *the simplest problem of the calculus of variations*

$$\min_u I(u), \quad I(u) = \int_0^1 F(x,u,u')dx, \quad u(0) = a_0, \ u(1) = a_1. \tag{3.1}$$

where F is twice differentiable function of its three arguments. We suppose that function $u_0 = u_0(x)$ is a minimizer and replace u_0 with a test function $u_0 + \delta u$, assuming that the norm$\|\delta u\|$ of the variation δu is infinitesimal. The test function $u_0 + \delta u$ satisfies the same boundary conditions as u_0. If indeed u_0 is a minimizer, the increment of the cost $\delta I(u_0) = I(u_0 + \delta u) - I(u_0)$ is nonnegative:

$$\delta I(u_0) = \int_0^1 (F(x, u_0 + \delta u, (u_0 + \delta u)') - F(x, u_0, u_0'))dx \geq 0. \tag{3.2}$$

If δu is not specified, the equation (3.2) is not too informative. However, it allows to find a minimizer if it can be simplified due to a particular form of the variation. Calculus of variations suggests a set of tests that differ by various assumed form of variations δu and corresponding form of (3.2).

Euler–Lagrange Equations The simplest variational condition (the Euler–Lagrange equation) is derived assuming that the variation δu is infinitesimal small and localized:

$$\delta u = \begin{cases} \rho(x) & \text{if} \quad x \in [x_0, x_0 + \varepsilon], \\ 0 & \text{if} \quad x \text{ is outside of } [x_0, x_0 + \varepsilon]. \end{cases} \tag{3.3}$$

Here $\rho(x)$ is a continuous function that vanishes at points x_0 and $x_0 + \varepsilon$ and is constrained as follows:

$$|\rho(x)| < \varepsilon, \quad |\rho'(x)| < \varepsilon \quad \forall x. \tag{3.4}$$

Linearizing (3.2) with respect to ε and collecting linear terms, we rewrite it as

$$\delta I(u_0) = \varepsilon \left(\int_0^1 \left(\frac{\partial F}{\partial u}(\delta u) + \underline{\frac{\partial F}{\partial u'}(\delta u)'} \right) dx \right) + o(\varepsilon) \geq 0. \tag{3.5}$$

Integration by parts of the underlined term in (3.5) gives

$$\int_0^1 \frac{\partial F}{\partial u'}(\delta u)' dx = \int_0^1 \left(-\frac{d}{dx}\frac{\partial F}{\partial u'}(\delta u) \right) dx + \left. \frac{\partial F}{\partial u'}\delta u \right|_{x=0}^{x=1}$$

and we obtain

$$0 \leq \delta I(u_0) = \varepsilon \int_0^1 S(u, u', x)\delta u \, dx + \left. \frac{\partial F}{\partial u'}\delta u \right|_{x=0}^{x=1} + o(\varepsilon), \tag{3.6}$$

where

$$S(u, u', x) = -\frac{d}{dx}\frac{\partial F}{\partial u'} + \frac{\partial F}{\partial u}. \tag{3.7}$$

The nonintegral term in the right-hand side of (3.6) is zero, because the boundary values of u are prescribed $u(0) = a_0$ and $u(1) = a_1$; therefore their variations $\delta u|_{x=0}$ and $\delta u|_{x=1}$ equal zero,

$$\delta u|_{x=0} = 0, \quad \delta u|_{x=1} = 0$$

Due to the arbitrariness of δu, we conclude that any differentiable minimizer u_0 of the simplest variational problem solves the boundary value problem

$$S(x,u,u') = \frac{d}{dx}\frac{\partial F}{\partial u'} - \frac{\partial F}{\partial u} = 0 \quad \forall x \in (0,1); \quad u(0) = u_0, \quad u(1) = u_1, \quad (3.8)$$

and the corresponding boundary conditions, called the *Euler–Lagrange equation*. The Euler–Lagrange equation is also called the *stationary condition* since it expresses stationarity of the variation. Indirectly, we assume in this derivation that u_0 is a twice differentiable function of x. Indeed, the left-hand side of equation (3.8) can be rewritten as

$$S(x,u,u') = \frac{\partial^2 F}{\partial u'^2}u'' + \frac{\partial^2 F}{\partial u'\partial u}u' + \frac{\partial^2 F}{\partial u'\partial x} - \frac{\partial F}{\partial u} \quad (3.9)$$

using the chain rule.

Example 3.1.1 Compute the Euler equation for the problem

$$I = \min_{u(x)} \int_0^1 \left(\frac{1}{2}(u')^2 + \frac{1}{2}u^2\right) dx \quad u(0) = 1,\ u(1) = c$$

We compute $\frac{\partial L}{\partial u'} = u', \quad \frac{\partial L}{\partial u} = u$ and the Euler equation becomes

$$u'' - u = 0 \text{ in } (0,1), \quad u(0) = 1,\ u(1) = c.$$

The minimizer $u_0(x)$ is

$$u_0(x) = \cosh(x) - \coth(1)\sinh(x)$$

Remark 3.1.1 The stationarity test alone does not allow to conclude whether u_0 is a true minimizer or even to conclude that a solution to (3.8) exists. For example, the function u that *maximizes* $I(u)$ satisfies the same Euler–Lagrange equation. The tests that distinguish minimal trajectory from other stationary trajectories are discussed in Chapter 6.

Remark 3.1.2 In many application, we consider a broken extremals that do not have the second derivative at some points. In this cases, it is more convenient to understand the Euler equation in the *weak sense*, or replace it with the integral identity

$$\int_0^1 \left(\frac{\partial F}{\partial u}v + \frac{\partial F}{\partial u'}v'\right) dx = 0 \quad \forall v(x) \in V \quad (3.10)$$

that must be satisfied for all differentiable functions v that vanish at the ends of the interval:

$$\mathcal{V} = \{v(x) : v(x) \in C_1[0,1], \quad v(0) = x(1) = 0\}.$$

The reader notices that the arbitrary "trial function" v is but the variation δu.

The definition of the weak solution naturally arise from the variational formulation that does not check the behavior of the minimizer in each point but in each infinitesimal interval. The minimizer can change its values at a several points, or even at a set of zero measure without alternation the objective functional. In ambiguous cases, one should specify in what sense (Riemann, Lebesgue) the integral is defined and change the definition of variation accordingly.

3.1.2 First integrals: Three special cases

In several cases, the Euler equation (3.8) can be integrated at least once. These are the cases when Lagrangian $F(x,u,u')$ does not depend on one of arguments. Below, we investigate them.

Lagrangian is independent of u' Assume that $F = F(x,u)$, and the minimization problem is

$$J(u) = \int_0^1 F(x,u)dx \tag{3.11}$$

In this case, the variation does not involve integration by parts, and the minimizer does not need to be continuous. Euler equation (3.8) becomes an algebraic relation for u

$$\frac{\partial F}{\partial u} = 0 \tag{3.12}$$

Curve $u(x)$ is determined in each point independently of neighboring points. The boundary conditions in (3.8) are satisfied by jumps of the extremal $u(x)$ in the end points; these conditions do not affect the objective functional at all.

Example 3.1.2 Consider the problem

$$\min_{u(x)} J(u), \quad J(u) = \int_0^1 (u - \sin x)^2 dx, \quad u(0) = 1;\; u(1) = 0.$$

The minimal value $J(u_0) = 0$ corresponds to the discontinuous minimizer

$$u(x) = \begin{cases} \sin x & \text{if } \ 0 \le x \le 1 \\ 1 & \text{if } \ x = 0 \\ 0 & \text{if } x = 1 \end{cases}$$

Formally, the discontinuous minimizer contradicts the assumption posed when the Euler equation were derived. To be consistent, we need to repeat the derivation of the necessary condition for the problem (3.11) without any assumption on the continuity of the minimizer. This derivation is quite obvious.

Lagrangian is independent of u If Lagrangian is independent on u, $F = F(x,u')$, Euler equation (3.8) can be integrated once:

$$\frac{\partial F}{\partial u'} = \text{ constant} \tag{3.13}$$

The first order differential equation (3.13) for u is the *first integral* of the problem; it defines a quantity that stays constant everywhere along the optimal trajectory. To find the optimal trajectory, it remains to integrate the first order equation (3.13) and determine the constants of integration from the boundary conditions.

Example 3.1.3 Consider the problem

$$\min_{u(x)} J(u), \quad J(u) = \int_0^1 (u' - \cos x)^2 dx, \quad u(0) = 1; \; u(1) = 0.$$

The first integral is

$$\frac{\partial F}{\partial u'} = u'(x) - \cos x = C$$

Integrating, we find the minimizer,

$$u(x) = -\sin x + Cx + C_1.$$

The constants C and C_1 are found from and the boundary conditions:

$$C_1 = 1, \quad C = -1 - \sin 1,$$

minimizer u_0 and the cost of the problem become, respectively

$$u_0(x) = \sin x - x - \sin 1 \quad J(u_0) = \int_0^1 x^2 dx = \frac{1}{3}$$

Notice that the Lagrangian in the example (3.1.2) is the square of difference between the minimizer u and function $\sin x$, and the Lagrangian in the example (3.1.3) is the square of difference of their derivatives. In the problem (3.1.2), the minimizer coincides with $\sin x$, and jumps to the prescribed boundary values. The minimizer u in the example (3.1.3) does not coincide with $\sin x$ at any interval. The difference between these two examples is that in the last problem the derivative of the minimizer must exist everywhere. Formally, the discontinuity of the minimizer would leave the derivative formally undefined. More important, that an approximation of a derivative to a discontinuous function would grow fast in the proximity of the point of discontinuity, this growth would increase the objective functional, and therefore it is nonoptimal. We deal with such problems below in Chapter 7.

Lagrangian is independent of x If $F = F(u, u')$, equation (3.8) has the first integral:

$$W(u, u') = \text{ constant} \tag{3.14}$$

where

$$W(u, u') = u' \frac{\partial F}{\partial u'} - F$$

Indeed, compute the x-derivative of $W(u, u')$ which must be equal to zero by virtue of (3.14):

$$\frac{d}{dx}W(u,u') =$$
$$\left[u''\frac{\partial F}{\partial u'} + u'\left(\frac{\partial^2 F}{\partial u' \partial u}u' + \frac{\partial^2 F}{\partial u'^2}u''\right)\right] - \frac{\partial F}{\partial u}u' - \frac{\partial F}{\partial u'}u'' = 0$$

where the expression in square brackets is the derivative of the first term of $W(u, u')$. Cancelling the equal terms, we bring this equation to the form

$$u'\left(\frac{\partial^2 F}{\partial u'^2}u'' + \frac{\partial^2 F}{\partial u' \partial u}u' - \frac{\partial F}{\partial u}\right) = 0 \tag{3.15}$$

The expression in parenthesis coincide with the left-hand-side term $S(x, u, u')$ of the Euler equation in the form (3.9), simplified for the considered case (F is independent of x, $F = F(u, u')$).

Example 3.1.4 Consider the Lagrangian

$$F = \frac{1}{2}\left[(u')^2 - \omega^2 u^2\right]$$

The Euler equation is

$$u'' + \omega^2 u = 0$$

The first integral is

$$W = (u')^2 + \omega^2 u^2 = C^2 = \text{ constant}$$

Let us immediately check the constancy of the first integral. The solution u of the Euler equation is equal

$$u = A\cos(\omega x) + B\sin(\omega x)$$

where A and B are constants. Substituting the solution into the expression for the first integral, we compute

$$W = (u')^2 + \omega^2 u^2 = \left[-A\omega\sin(cx) + B\omega\cos(\omega x)\right]^2$$
$$+\omega^2\left[A\cos(\omega x) + B\sin(\omega x)\right]^2 = \omega^2(A^2 + B^2)$$

We have shown that W is constant at the optimal trajectory. In mechanical application, W is the whole energy of the oscillator.

Instead of solving the Euler equation, we may solve the first-order equation $W = 0$ obtaining the same solution.

Later we discuss the methods to regularly find first integrals of Euler equations for more general variational problems.

3.1.3 Variational problem as the limit of a vector problem

Consider a finite-dimensional approximation of the simplest variational problem

$$\min_{u(x)} I(u), \quad I(u) = \int_a^b F(x, u, u')dx$$

Assume in addition that the minimizer belongs to the class of piece-wise constant functions $\mathcal{U}_N$:

$$\bar{u}(x) \in \mathcal{U}_N, \quad \text{if } \bar{u}(x) = u_i \quad \forall x \in \left[a + \frac{i}{N}(b-a)\right]$$

A function $\bar{u}$ in $\mathcal{U}_N$ is defined by an N-dimensional vector $\{u_1, \dots u_N\}$.

Reformulation the variational problem, we replace the derivative $u'(x)$ with a finite difference $\mathrm{Diff}\,(u_i)$ where the operator Diff is defined at sequences $\mathcal{U}_N$ as follows

$$\mathrm{Diff}\,(u_i) = \frac{1}{\Delta}(u_i - u_{i-1}), \quad \Delta = \frac{b-a}{N}; \tag{3.16}$$

when $N \to \infty$, this operator tends to the derivative.

The variational problem is replaced with the finite-dimensional optimization problem:

$$\min_{u_1,\dots,u_{N-1}} I_N \quad I_N = \sum_{i=1}^{N} F_i(u_i, \mathrm{Diff}\,(u_i)), \quad \mathrm{Diff}\,(z_i) = \frac{1}{\Delta}(z_i - z_{i-1}) \tag{3.17}$$

Compute the stationary conditions for the minimum of $I_N(u)$

$$\frac{\partial I_N}{\partial u_i} = 0, \quad i = 1.\dots, N.$$

Notice that only two terms, F_i and F_{i+1}, in the above sum depend on u_i: the first depends on u_i directly and also through the operator $\mathrm{Diff}\,(u_i)$, and the second– only through $\mathrm{Diff}\,(u_i)$:

$$\begin{aligned}
\frac{dF_i}{du_i} &= \frac{\partial F_i}{\partial u_i} + \frac{\partial F_i}{\partial \mathrm{Diff}\,(u_i)}\frac{1}{\Delta}, \\
\frac{dF_{i+1}}{du_i} &= -\frac{\partial F_{i+1}}{\partial \mathrm{Diff}\,(u_i)}\frac{1}{\Delta}. \\
\frac{dF_k}{du_i} &= 0 \quad k \neq i, k \neq i+1
\end{aligned}$$

Therefore, the stationary condition with respect to u_i has the form

$$\frac{\partial I_N}{\partial u_i} = \frac{\partial F_i}{\partial u_i} + \frac{1}{\Delta}\left(\frac{\partial F_i}{\partial \mathrm{Diff}\,(u_i)} - \frac{\partial F_{i+1}}{\partial \mathrm{Diff}\,(u_{i+1})}\right) = 0$$

or, recalling the definition (3.16) of Diff-operator, the form

$$\frac{\partial I_N}{\partial u_i} = \frac{\partial F_i}{\partial u_i} - \mathrm{Diff}\left(\frac{\partial F_{i+1}}{\partial \mathrm{Diff}\,(u_{i+1})}\right) = 0.$$

The initial and the final point u_0 and u_N enter the difference scheme only once, therefore the optimality conditions are different. They are, respectively,

$$\frac{\partial F_{N+1}}{\partial \mathrm{Diff}\,(u_{N+1})} = 0; \quad \frac{\partial F_o}{\partial \mathrm{Diff}\,(u_0)} = 0.$$

Formally passing to the limit $N \to \infty$, Diff $\to \frac{d}{dx}$, we simply replace the index $(_i)$ with a continuous variable x, vector of values $\{u_k\}$ of the piece-wise constant function with the continuous function $u(x)$, difference operator Diff with the derivative $\frac{d}{dx}$; then

$$\sum_{i=1}^{N} F_i(u_i, \mathrm{Diff}\, u_i) \to \int_a^b F(x, u, u')dx.$$

and

$$\frac{\partial F_i}{\partial u_i} - \mathrm{Diff}\left(\frac{\partial F_{i+1}}{\partial \mathrm{Diff}\,(u_{i+1})}\right) \to \frac{\partial F}{\partial u} - \frac{d}{d\,x}\frac{\partial F}{\partial u'}$$

The conditions for the end points become the natural variational conditions:

$$\frac{\partial F}{\partial u'(0)} = 0, \quad \frac{\partial F}{\partial u'(T)} = 0,$$

Remarks on existence of a differentiable minimizer

So far, we followed the formal scheme of necessary conditions, thereby tacitly assuming that all derivatives of the Lagrangian exist, the increment of the functional is correctly represented by the first term of its power expansion, and the limit of the sequence of finite-dimensional problems exist and does not depend on the partition $\{x_1, \ldots x_N\}$ if only $|x_k - x_{k-1}| \to 0$ for all k. We also indirectly assume that the Euler equation has at least one solution consistent with boundary conditions.

If all the made assumptions are correct, we obtain a curve that might be a minimizer because it cannot be disproved by the stationary test. In other terms, we find that is there is no other close-by classical curve that correspond to a smaller value of the functional. This statement about the optimality seems to be rather weak but this is exactly what the calculus of variation can give us. On the other hand, the variational conditions are universal and, being appropriately used and supplemented by other conditions, lead to a very detailed description of the extremal as we show later in the course.

Remark on differentiability Freshet and Chateaux derivatives.

In this text, we do not fully discuss the assumptions restricting ourself with remarks and references to more detailed sources.

Remark on convergence In the above procedure, we assume that the limits of the components of the vector $\{u_k\}$ represent values of a smooth function in the close-by points $x_1, \ldots, x_N$. At the other hand, u_k are solutions of optimization problems with the coefficients that slowly vary with the number k. We need to answer the question whether the solution of a minimization problem tends to is a differentiable function of x; that is whether the limit

$$\lim_{k\to\infty} \frac{u_k - u_{k-1}}{x_k - x_{k-1}}$$

exists and this is not always the case. We address this question later in Chapter 7

3.2 Boundary terms

3.2.1 Boundary conditions and Weierstrass-Erdman condition

Variational conditions and natural conditions In some variational problems, the condition $u(b) = u_b$ on one or both ends of extremal can be not specified. Also, the objective functional may contain terms defined on the boundary only in which case the problem becomes

$$\min_{u(x):u(a)=u_a} I(u), \quad I(u) = \int_a^b F(x,u,u')dx + f(u(b)) \tag{3.18}$$

The Euler equation for the problem remain the same $S(x,u,u') = 0$ but this time it must be supplemented by a *variational boundary condition* that is derived from the requirement of the stationarity of the minimizer with respect to variation of the boundary term. This term is

$$\delta u \frac{\partial F}{\partial u'} + \delta u \frac{\partial f}{\partial u}$$

The first term comes from the integration by part in the derivation of Euler equation (see (**??**)) and the second is the variation of the out-of-integral term in the objective functional (3.18) The stationarity condition with respect to the variation of $\delta u(b)$

$$\frac{\partial F}{\partial u'}|_{x=b} + \frac{\partial f}{\partial u}|_{x=b} = 0 \tag{3.19}$$

expresses the boundary condition for the extremal $u(x)$ at the endpoint $x = b$. Similar condition can be derived for the point $x = a$ if the value in this point is not prescribed.

Example 3.2.1 Minimize the functional

$$I(u) = \min_u \int_0^1 \frac{1}{2}(u')^2 dx + Au(1), \quad u(0) = 0$$

Here, we want to minimize the endpoint value and we do not want the trajectory be too steep. The Euler equation $u'' = 0$ must be integrated with boundary conditions $u(0) = 0$ and (see (3.19)) $u'(1) + A = 0$ The extremal is a straight line, $u = -Ax$. The cost of the problem is $I = -\frac{1}{2}A^2$.

If no out-of-integral terms are presented, the condition becomes

$$\frac{\partial F}{\partial u'}|_{x=b} = 0 \tag{3.20}$$

and it is called the *natural boundary condition.*

Example 3.2.2 The natural boundary condition for the problem with the Lagrangian $L = (u')^2 + \phi(x, u)$ is $u'|_{x=b} = 0$

Broken extremal and the Weierstrass-Erdman condition The classical derivation of the Euler equation requires the existence of all second partials of F, and the solution u of the second-order differential equation is required to be twice-differentiable.

In many cases of interest, the Lagrangian is only piece-wise twice differentiable; in this case, the extremal consists of several curves – solutions of the Euler equation that are computed at the intervals of smoothness of the Lagrangian. The question is: How to join these pieces together?

We always assume that the extremal u is differentiable everywhere so that the first derivative u' exists at all point of the trajectory. But the derivative u'itself does not need to be continuous to solve Euler equation: Only the differentiability of $\frac{\partial F}{\partial u'}$ is needed to ensure the exitance of the term $\frac{d}{dx}\frac{\partial F}{\partial u'}$ in the Euler equation. This requirement on differentiability of an optimal trajectory is yields to the *Weierstrass-Erdman condition on broken extremal.*

> At any point of the optimal trajectory, the Weierstrass-Erdman condition must be satisfied:
>
> $$\left[\frac{\partial F}{\partial u'}\right]_-^+ = 0 \quad \text{along the optimal trajectory } u(x) \tag{3.21}$$
>
> Here $[z]_-^+ = z_+ - z_-$ denotes the jump of the variable z.

Example 3.2.3 (Broken extremal) Consider the Lagrangian

$$F = \frac{1}{2}a(x)(u')^2 + \frac{1}{2}u^2, \quad a(x) = \begin{cases} a_1 & \text{if } x \in [0, x_*) \\ a_2 & \text{if } x \in [x_*, 1) \end{cases}$$

where x_* is point in $(0, 1)$.

The Euler equation that is hold everywhere in $(0, 1)$ except of the point x_*,

$$\frac{d}{dx}[a_1 u'] - bu = 0 \quad \text{if } x \in [0, x_*)$$
$$\frac{d}{dx}[a_2 u'] - bu = 0 \quad \text{if } x \in [x_*, 1),$$

At $x = x_*$, the Weierstrass-Erdman condition holds,

$$a_1(u')(x_* - 0) = a_2(u')(x_* + 0).$$

The derivative u' itself is discontinuous; its jump is determined by the jump in coefficients:

$$u'(x_* + 0) = \left(\frac{a_1}{a_2}\right) u'(x_* - 0)$$

This condition, together with the Euler equation and boundary conditions allows for determination of the optimal trajectory.

3.2.2 Non-fixed interval. Transversality condition

Free boundary Consider now the case when the interval (a, b) is not fixed, and the end point is to be chosen to minimize the functional. Suppose first that no conditions on the end point are imposed. We compute the difference between two functionals

$$\delta I = \int_a^{b+\delta x} F(x, u + \delta u, u' + \delta u')dx - \int_a^b F(x, u, u')dx$$

The linear terms of the difference are

$$\delta I = A_x \delta x + A_u \delta u$$

where A_x is the increment due to variation of the interval when u keeps its stationary value, and A_u is the increment due to variation $\delta_x u = \frac{du}{dx}\delta x$ of u when the interval keeps its stationary value. Let us compute these quantities. We have

$$A_x = \int_b^{b+\delta x} F(x, u, u')dx + \left.\frac{\partial F}{\partial u'}\right|_{x=b} \delta_x u$$

where $\delta_x u$ is the variation of u due to variation of the point b. It is equal

$$\delta_x u = u(b + \delta x) - u(b) = u'|_{x=b}\delta x.$$

Substituting this into expression for A_x and rounding to $o(\delta x)$, we obtain

$$A_x = \delta x \left(F(x, u, u') - \frac{\partial F}{\partial u'} u'\right)_{x=b}.$$

The increment's part A_u is computed in a standard manner

$$A_u = \int_a^b \delta u \left(S(x, u, u')\right) dx + \left.\frac{\partial F}{\partial u'}\right|_{x=b} \delta u|_{x=b}$$

where S is the differential expression of the Euler equation and δu is the variation of the trajectory (that is independent of the variation of x. Because of arbitrariness of δx and δu, we conclude that

$$A_x = \left(F(b, u(b), u'(b)) - u' \frac{\partial F}{\partial u'}\right)_{x=b} = 0 \qquad (3.22)$$

and

$$\left.\frac{\partial F}{\partial u'}\right|_{x=b} = 0 \tag{3.23}$$

at the unknown end of the trajectory. Equation (3.22) together with boundary conditions determine boundary values of u and the length of interval of integration, while the equation $S(x, u, u') = 0 \in (a, b)$ states that the Euler equation is satisfied along the optimal trajectory. The differential equation for extremal an extra boundary condition (3.22) to satisfy, but is also has an additional degree of freedom: the non-fixed length of the interval of integration.

Notice that the condition at the unknown end has the same form as the first integral of the problem in the case when $F(u, u')$ is independent of x. This shows that the condition (3.22) cannot be satisfied at an isolated point of the trajectory, unless the Lagrangian explicitly depends on x.

Check the next example

Example 3.2.4 Consider the problem

$$\min_{u(x),s} \int_0^s \left(\frac{1}{2}u'^2 - u + \frac{3}{2}\right) dx \quad u(0) = 0.$$

Euler equation $u'' + 1 = 0$ and the condition $u(0) = 0$ produces the solution

$$u = -\frac{1}{2}x^2 + Ax, \quad u' = x + A$$

where A is a constant. The conditions at the unknown point s are

$$\frac{\partial F}{\partial u'} = s + A = 0 \quad \text{or } A = -s$$

(condition (3.23)) and

$$A_x = F(s, u(s), u'(s)) - u'\frac{\partial F}{\partial u'} = -\frac{1}{2}(s + A)^2 + \frac{1}{2}s^2 - As - \frac{3}{2} = \frac{3}{2}s^2 - \frac{3}{2} = 0$$

(condition (3.22)). Solving for $A = -s$, we obtain $s = 1$ and $u = \frac{1}{2}x^2 - x$.

3.2.3 Extremal broken at the unknown point

Combining the techniques, we may address the problem of en extremal broken in an unknown point. The position of this point is determined from the minimization requirement. Assume that Lagrangian has the form

$$F(x, u, u') = \begin{matrix} F_-(x, u, u') & \text{if } x \in (a, \xi) \\ F_+(x, u, u') & \text{if } x \in (\xi, b) \end{matrix}$$

where ξ is an unknown point in the interval (a, b) of the integration. The Euler equation is

$$S_F(u) = \begin{matrix} S_{F_-}(u) & \text{if } x \in (a, \xi) \\ S_{F_+}(u) & \text{if } x \in (\xi, b) \end{matrix}$$

The stationarity conditions at the unknown point ξ are

$$\frac{\partial F_+}{\partial u'} = \frac{\partial F_0}{\partial u'} \tag{3.24}$$

(the stationarity of the trajectory) and

$$F_+(u) - u'_+ \frac{\partial F_+}{\partial u'} = F_-(u) - u'_- \frac{\partial F_-}{\partial u'} \tag{3.25}$$

(the stationarity of the position of the transit point). They are derived by the same procedure as the conditions at the end point. The variation δx of the transit point increases the first part of the trajectory and and increases the second part, $\delta x = \delta x_+ = -\delta x_-$ which explains the structure of the stationary conditions.

In particular, if the Lagrangian is independent of x, the condition (3.25) express the constancy of the first integral (**??**) at the point ξ.

Example 3.2.5 Consider the problem with Lagrangian

$$F(x,u,u') = \begin{array}{ll} a_+u'^2 + b_+u^2 & \text{if } x \in (a,\xi) \\ a_-u'^2 & \text{if } x \in (\xi,b) \end{array}$$

and boundary conditions

$$u(a) = 0, \quad u(b) = 1$$

The Euler equation is

$$S(F,u) = \begin{array}{ll} a_+u'' - b_-u = 0 & \text{if } x \in (a,\xi) \\ a_-u'' = 0 & \text{if } x \in (\xi,b) \end{array}$$

The solution to this equation that satisfies the boundary conditions is

$$\begin{array}{ll} u_+(x) = C_1 \sinh\left(\sqrt{\frac{b_+}{a_+}}(x-a)\right) & \text{if } x \in (a,\xi) \\ u_-(x) = C_2(x-b) + 1 & \text{if } x \in (\xi,b) \end{array};$$

it depends on three constants ξ, C_1, and C_2 (Notice that the coefficient a_- does not enter the Euler equations). These constants are determined from the thee remaining conditions at the unknown point ξ which express
(1) continuity of the extremal

$$u_+(\xi) = u_-(\xi),$$

(2) the Weierstrass-Endmann condition

$$a_+u'_+(\xi) = a_-u'_-(\xi),$$

(3) and the transversality condition

$$-a_+(u'_+(\xi))^2 + b_+u(\xi)^2 = -a_-(u'_-(\xi))^2.$$

Let us analyze them. The transversality condition is the simplest one because it states the equality of two first integral. It is simplified to

$$C_1^2 b_+ = C_2^2 a_-$$

From the condition (2), we have

$$C_1 \sqrt{a_+ b_+} \cosh q = C_2, \quad \text{where } q = \sqrt{\frac{b_+}{a_+}}(\xi - a)$$

Together with the previous condition and the definition of q, it allows for determination of ξ:

$$\cosh q = \sqrt{a_+ a_-}, \quad \Rightarrow \quad \xi = a + \frac{a_+}{b_+} \cosh^{-1} \sqrt{a_+ a_-}$$

Finally, we define constants C_1 and C_2 from the continuity condition:

$$C_1 \sinh q = 1 + C_2(\xi - b)$$

and the transversality condition as

$$C_1 = \frac{\sqrt{a_-}}{\sqrt{a_-} \sinh q - \sqrt{b_+}(\xi - b)}, \quad C_2 = \frac{\sqrt{b_+}}{\sqrt{a_-} \sinh q - \sqrt{b_+}(\xi - b)},$$

3.3 Several minimizers

3.3.1 Euler equations and first integrals

The Euler equation can be naturally generalized to the problem with the vector-valued minimizer

$$I(u) = \min_u \int_a^b F(x, u, u')dx, \tag{3.26}$$

where x is a point in the interval $[a, b]$ and $u = (u_1(x), \dots, u_n(x)]$ is a vector function. We suppose that F is a twice differentiable function of its arguments.

Let us compute the variation $\delta I(u)$ equal to $I(u + \delta u) - I(u)$, assuming that the variation of the extremal and its derivative is small and localized. To compute the Lagrangian at the perturbed trajectory $u+\delta u$, we use the expansion

$$F(x, u + \delta u, u' + \delta u') = F(x, u, u') + \sum_{i=1}^n \frac{\partial F}{\partial u_i}\delta u_i + \sum_{i=1}^n \frac{\partial F}{\partial u_i'}\delta u_i'$$

We can perform n independent variation of each component of vector u applying variations $\delta_i u = (0, \dots, \delta u_i \dots, 0)$. The increment of the objective functional should be zero for each of these variation, otherwise the functional can be decrease by one of them. But the stationary condition for any of considered variation coincide with the one-minimizer case.

$$\delta_i I(u) = \int_a^b \left(\delta u_i \frac{\partial F}{\partial u_i} + \delta u_i' \frac{\partial F}{\partial u_i'} \right) dx \geq 0 \quad i = 1, \dots, n.$$

Proceeding as before, we obtain the system of differential equations of the order $2n$,

$$\frac{d}{dx}\frac{\partial F}{\partial u_i'} - \frac{\partial F}{\partial u_i} = 0, \quad i = 1, \dots n \tag{3.27}$$

and the boundary terms

$$\sum_{i=1}^{n} \frac{\partial F}{\partial u_i'}\delta u_i \Bigg|_{x=a}^{x=b} = 0 \tag{3.28}$$

Remark 3.3.1 The vector form of the system (3.27),

$$S(F,u) = \frac{d}{dx}\frac{\partial F}{\partial u'} - \frac{\partial F}{\partial u} = 0, \qquad \delta u^T \frac{\partial F}{\partial u'}\Bigg|_{x=a}^{x=b} = 0 \tag{3.29}$$

is identical to the scalar Euler equation. This system is but an algebraic definition of differentiation on a vector argument u.

Example 3.3.1 Consider the problem with the integrand

$$F = \frac{1}{2}u_1'^2 + \frac{1}{2}u_2'^2 - u_1 u_2' + \frac{1}{2}u_1^2 \tag{3.30}$$

The system of stationarity conditions is computed to be

$$\begin{aligned} \frac{d}{dx}\frac{\partial F}{\partial u_1'} - \frac{\partial F}{\partial u_1} &= \quad u_1'' + u_2' - u_1 = 0 \\ \frac{d}{dx}\frac{\partial F}{\partial u_2'} - \frac{\partial F}{\partial u_2} &= \quad (u_2' + u_1)' = 0. \end{aligned}$$

If consists of two differential equations of second order for two unknowns $u_1(x)$ and $u_2(x)$.

First integrals The earlier mentioned first integrals can be derived for the vector problem as well.

1. If F is independent of u_k', then one of the Euler equations degenerates into algebraic relation:

   $$\frac{\partial F}{\partial u_k} = 0$$

   and the order of the system (3.27) decreases by two. The variable $u_k(x)$ can be a discontinuous function of x in an optimal solution. Since the Lagrangian is independent of u_k', the jumps in $u_k(x)$ may occur along the optimal trajectory.

2. If F is independent of u_k, the first integral exists:

   $$\frac{\partial F}{\partial u_k'} = \text{constant}$$

For instance, the second equation in Example 3.3.1 can be integrated and replaced by

$$u_2' + u_1 = \text{ constant}$$

3. Finally, if F is independent of x, $F = F(\boldsymbol{u}, \boldsymbol{u}')$ then a first integral exist

$$W = u'^T \frac{\partial F}{\partial u'} - F = \text{constant} \tag{3.31}$$

Here

$$u'^T \frac{\partial F}{\partial u'} = \sum_{i=1}^{n} u_i' \cdot \frac{\partial F}{\partial u_i'}$$

For the Lagrangian in Example 3.3.1, this first integral is computed to be

$$\begin{aligned} W &= u_1^2 + u_2(u_2 - u_1) - \left(\frac{1}{2}u_1'^2 + \frac{1}{2}u_2'^2 - u_1 u_2' + \frac{1}{2}u_1^2\right) \\ &= \frac{1}{2}\left(u_1'^2 + u_2'^2 - u_1^2\right) = \text{ constant} \end{aligned}$$

Clearly, these three cases do not exhaust all possible first integrals for vector case; one can hope to find new invariants for instance by changing the variables. The theory of first integrals will be discussed later in Sections 8.1 and 8.2.

Transversality and Weierstrass-Erdmann conditions These conditions are quite analogous to the scalar case and their derivation is straightforward. We simply listen here these conditions.

The expression $\frac{\partial F}{\partial u_i'}$ remain continuous at every point of an optimal trajectory, including the points where u_i is discontinuous.

If the end point of the trajectory is unknown, the condition

$$u^T \frac{\partial F}{\partial u'} - F = 0$$

at the end point is satisfied.

3.3.2 Variational boundary conditions

The first variation corresponds to solution of the equation (3.28) which must produce $2n$ boundary conditions for the Euler equations (3.27). If the values of all minimizers are prescribed at the end points,

$$u_i(a) = u_i^a, \quad u_i(b) = u_i^b,$$

then the equation (3.28) is satisfied, because all variations are zero. If the values of several components of $u(a)$ or $u(b)$ are not given, the corresponding natural boundary conditions must hold:

$$\text{Either } \left.\frac{\partial F}{\partial u_i'}\right|_{x=a,b} = 0 \quad \text{or } \delta u|_{x=a,b} = 0 \tag{3.32}$$

Therefore, the variational problem for a vector minimizer can be solved with a number of boundary requirements that does not surpass n for both ends of the interval. The missing boundary conditions will be supplemented by the requirement of optimality of the trajectory.

Consider a more general case when p boundary conditions of the form

$$\beta_k(u_1, \ldots, u_n) = 0 \quad k = 1, \ldots, p < n \tag{3.33}$$

are prescribed and the end point $x = b$ (the other end is considered similarly). We need to find $n - p$ supplementary variational constraints at this point that together with (3.33) form n boundary conditions for the Euler equation (3.28). The conditions (3.33) are satisfied at all perturbed trajectories, $\beta_k(u_1 + \delta u_1, \ldots, u_n + \delta u_n) = 0$, therefore the variations δu_i are bounded by a linear system

$$\frac{\partial \beta_k}{\partial u_1}\delta u_1 + \ldots + \frac{\partial \beta_k}{\partial u_n}\delta u_n, \quad k = 1, \ldots, p$$

or, in the matrix form $P\delta u = 0$, where P is the $p \times n$ matrix with the elements $P_{ki} = \frac{\partial \beta_k}{\partial u_i}$. This system has n unknowns and $p < n$ equations, and it is satisfied when the unknowns are expressed through $(n-p)$-dimensional arbitrary vector v as follows $\delta u = Qv$. Here, $(n-p) \times n$ matrix Q is supplementary to P; it is computed solving the matrix equation $PQ = 0$. This representation, substituted into second equation of (3.29), gives the missing boundary conditions

$$Q^T \; \delta u^T \frac{\partial F}{\partial u'}\bigg|_{x=b} = 0$$

(Here, use used the arbitrariness of potential vector v).

Example 3.3.2 Consider again the variational problem with the Lagrangian (3.30) assuming that the following boundary conditions are prescribed

$$u_1(a) = 1, \quad \beta(u_1(b), u_2(b)) = u_1^2(b) + u_2^2(b) = 1$$

Find the complementary variational boundary conditions. At the point $x = a$, the variation δu_1 is zero, and δu_2 is arbitrary. The variational condition is

$$\frac{\partial F}{\partial u_2'}\bigg|_{x=a} = u_2'(a) - u_1(a) = 0$$

or, since $u_1(a) = 1$, $u_2'(a) = 1$

At the point $x = b$, the variations δu_1 and δu_2 are connected by the relation

$$\frac{\partial \beta}{\partial u_1}\delta u_1 + \frac{\partial \beta}{\partial u_2}\delta u_2 = 2u_1\delta u_1 + 2u_2\delta u_2 = 0$$

which implies the representation ($\delta u = Qv$)

$$\delta u_1 = -u_2 v, \quad \delta u_2 = u_1 v$$

where v is an arbitrary potential. The variational condition at $x = b$ becomes

$$\left(-\frac{\partial F}{\partial u_1'} u_2 + \frac{\partial F}{\partial u_2'} u_1 \right)_{x=b} = \left(-u_1' u_2 + (u_2' - u_1) u_1 \right)_{x=b} v = 0 \quad \forall v$$

or

$$-u_1' u_2 + u_1 u_2' - u_1^2 \big|_{x=b} = 0.$$

We end up with four boundary conditions:

$$\begin{array}{ll} u_1(a) = 1, & u_1^2(b) + u_2^2(b) = 1, \\ u_2'(a) = 1, & u_1(b) u_2'(b) - u_1(b)' u_2(b) - u_1(b)^2 = 0. \end{array}$$

The conditions in the second raw are the variational conditions.

3.3.3 Lagrangian dependent on higher derivatives

Consider a more general type variational problem with the Lagrangian that depends on the minimizer and its first and second derivative,

$$J = \int_a^b F(x, u, u', u'') dx$$

The Euler equation is derived similarly to the simplest case: The variation of the goal functional is

$$\delta J = \int_a^b \left(\frac{\partial F}{\partial u} \delta u + \frac{\partial F}{\partial u'} \delta u' + \frac{\partial F}{\partial u''} \delta u'' \right) dx$$

Integrating by parts the second term and twice the third term, we obtain

$$\delta J = \int_a^b \left(\frac{\partial F}{\partial u} - \frac{d}{dx} \frac{\partial F}{\partial u'} + \frac{d^2}{dx^2} \frac{\partial F}{\partial u''} \right) \delta u \, dx + \left[\frac{\partial F}{\partial u'} \delta u + \frac{\partial F}{\partial u''} \delta u' - \frac{d}{dx} \frac{\partial F}{\partial u''} \delta u \right]_{x=a}^{x=b} \tag{3.34}$$

The stationarity condition becomes the fourth-order differential equation

$$\frac{d^2}{dx^2} \frac{\partial F}{\partial u''} - \frac{d}{dx} \frac{\partial F}{\partial u'} + \frac{\partial F}{\partial u} \tag{3.35}$$

supplemented by two natural boundary conditions on each end,

$$\delta u' \frac{\partial F}{\partial u''} = 0, \quad \delta u \left[\frac{\partial F}{\partial u'} - \frac{d}{dx} \frac{\partial F}{\partial u''} \right] = 0 \quad \text{at } x = a \text{ and } x = b \tag{3.36}$$

or by the correspondent main conditions posed on the minimizer u and its derivative u' at the end points.

Example 3.3.3 The equilibrium of an elastic bending beam correspond to the solution of the variational problem

$$\min_{w(x)} \int_0^L \left(\frac{1}{2}(E(x)w'')^2 - q(x)w\right)dx \tag{3.37}$$

where $w(x)$ is the deflection of the point x of the beam, $E(x)$ is the elastic stiffness of the material that can vary with x, $q(x)$ is the load that bends the beam. Any of the following kinematic boundary conditions can be considered at each end of the beam.
(1) A clamped end: $w(a) = 0, \quad w'(a) = 0$
(2) a simply supported end $w(a) = 0$.
(3) a free end (no kinematic conditions).

Let us find equation for equilibrium and the missing boundary conditions in the second and third case. The Euler equation (3.35) becomes

$$(Ew'')'' - q = 0 \quad \in (a, b)$$

The equations (3.36) become

$$\delta u'(Eu'') = 0, \quad \delta u\left((Ew'')'\right) = 0$$

In the case (2) (simply supported end), the complementary variational boundary condition is $Eu'' = 0$, it expresses vanishing of the bending momentum at the simply supported end. In the case (3), the variational conditions are $Eu'' = 0$ and $(Ew'')' = 0$; the last expresses vanishing of the bending force at the free end (the the bending momentum vanishes here as well).

Generalization Similarly, the stationary equations for Lagrangian $F\left(x, u, u', \ldots, u^{(n)}\right)$ dependent on first n derivatives of u is

$$\sum_{k=1}^{n} (-1)^k \frac{d^k}{dx^k} \frac{\partial F}{\partial u^{(k)}} + \frac{\partial F}{\partial u} = 0$$

In this formula, u can be replaced by a vector, if the case of several minimizers is considered.

Chapter 4

Immediate Applications

4.1 Geometric optics and Geodesics

4.1.1 Geometric optics problem. Snell's law

A half of century before the calculus of variation was invented, Fermat suggested that light propagates along the trajectory which minimizes the time of travel between the source with coordinates (a, A) and the observer with coordinates (b, B). This principle implies, that light travels along straight lines when the medium is homogeneous and along curved trajectories in an inhomogeneous medium in which the speed $v(x, y)$ of light depends on the position. The exactly same problem – minimization of the travel's time – can be formulated as the best route for a cross-country runner; the speed depends on the type of the terrains the runner crosses and is a function of the position. This problem is called the problem of geometric optic.

In order to formulate the problem of geometric optics, consider a trajectory in a plane, call the coordinates of the initial and final point of the trajectory (a, A) and (b, B), respectively, assuming that $a < b$ and call the optimal trajectory $y(x)$ thereby assuming that the optimal route is a graph of a function. The time T of travel can be found from the relation $v = \frac{ds}{dt}$ where $ds = \sqrt{1 + y'^2}dx$ is the infinitesimal length along the trajectory $y(x)$, or

$$dt = \frac{ds}{v(x, y)} = \frac{\sqrt{1 + y'^2}}{v}dx$$

where $ds = \sqrt{1 + y'^2}dx$ is the differential of the path. From this, we immediately find that

$$T = \int_a^b dt = \int_a^b \frac{\sqrt{1 + y'^2}}{v}dx$$

Let us consider minimization of T by the trajectory assuming that the medium is layered and the speed $v(y) = \frac{1}{\psi(y)}$ of travel varies only along the

y axes. The corresponding variational problem has the Lagrangian

$$F(y, y') = \psi(y)\sqrt{1 + y'^2}.$$

This problem allows for the first integral, (see above)

$$\psi(y)\frac{y'^2}{\sqrt{1 + y'^2}} - \psi(y)\sqrt{1 + y'^2} = c$$

or

$$\psi(y) = -c\sqrt{1 + y'^2} \tag{4.1}$$

Solving for y', we obtain the equation with separated variables

$$\frac{dy}{dx} = \pm\frac{\sqrt{c^2\psi^2(y) - 1}}{c}$$

with the solution

$$x = \pm\Phi(u) = \int \frac{c\,dy}{\sqrt{\psi^2(y) - c^2}} \tag{4.2}$$

Notice that equation (4.1) allows for a geometric interpretation: Derivative y' defines the angle α of inclination of the optimal trajectory, $y' = \tan\alpha$. In terms of α, the equation (4.1) assumes the form

$$\psi(y)\ \cos\alpha = c \tag{4.3}$$

which shows that the angle of the optimal trajectory varies with the speed $v = \frac{1}{\psi}$ of the signal in the media. The optimal trajectory is bent and directed into the domain where the speed is higher.

Snell's law of refraction

Assume that the speed of the signal in medium is piecewise constant; it changes when $y = y_0$ and the speed v jumps from v_+ to v_-, as it happens on the boundary between air and water,

$$v(y) = \begin{cases} v_+ & \text{if } y > y_0 \\ v_- & \text{if } y < y_0 \end{cases}$$

Let us find what happens with an optimal trajectory. Weierstrass-Erdman condition are written in the form

$$\left[v\frac{y'}{\sqrt{1 + y'^2}}\right]_-^+ = 0$$

Recall that $y' = \tan\alpha$ where α is the angle of inclination of the trajectory to the axis OX, then $\frac{y'}{\sqrt{1+y'^2}} = \sin\alpha$ and we arrive at the refraction law called Snell's law of refraction

$$v_+ \sin\alpha_+ = v_- \sin\alpha_-$$

4.1.2 Brachistochrone

Problem of the Brachistochrone is probably the most famous problem of classical calculus of variation; it is the problem this discipline start with. In 1696 Bernoulli put forward a challenge to all mathematicians asking to solve the problem: Find the curve of the fastest descent (brachistochrone), the trajectory that allows a mass that slides along it without tension under force of gravity to reach the destination point in a minimal time.

The problem was formulated by Galileo in "Besedy i math. dokazatelstvca" check it!

To formulate the problem, we use the law of conservation of the total energy – the sum of the potential and kinetic energy is constant in any time instance:

$$\frac{1}{2}mv^2 + mgy = \text{ constant}$$

where $y(x)$ is the vertical coordinate of the sought curve. From this relation, we express the speed v as a function of u

$$v = \sqrt{C - gy}$$

thus reducing the problem to a special case of geometric optics. (Of course the founding fathers of the calculus of variations did not have the luxury of reducing the problem to something simpler because it was the first and only real variational problem known to the time)

Applying the formula (4.1), we obtain

$$\frac{1}{\sqrt{C - gy}} = \sqrt{1 + y'^2}$$

and

$$x = \int \frac{\sqrt{y - y_0}dy}{\sqrt{2a - (y - y_0)}}$$

To compute the quadrature, we substitute

$$y = y_0 + 2a\sin^2\frac{\theta}{2},$$

then

$$x = 2a\int \sin^2\frac{\theta}{2}d\theta = a(\theta - \sin\theta) + x_0$$

To summarize, the optimal trajectory is

$$\begin{aligned} x &= x_0 + a(\theta - \sin\theta), \\ y &= y_0 + a(1 - \cos\theta), \end{aligned} \tag{4.4}$$

We recognize the equation of the cycloid in (4.4). Recall that cycloid is a curve generated by a motion of a fixed point on a circumference of the radius a which rolls on the given line $y - y_0$.

Remark 4.1.1 The obtained solution was formulated in a strange for modern mathematics terms: "Brachistochrone is isochrone." Isochrone was another name for the cycloid; the name refers to a remarkable property of it found shortly before the discovery of brachistochrone: The period of oscillation of a heavy mass that slides along a cycloid is independent of its magnitude. We will prove this property below in Example ??.

Remark 4.1.2 Notice that brachistochrone is in fact solution to the problem of optimal design: the trajectory must be chosen by a designer to minimize the goal (time of travel).

4.1.3 Minimal surface of revolution

Another classical example of design problem solved by variational methods is the problem of minimal surface that was discussed in Chapter 1. Here, we formulate is for the surface of revolution: Minimize the area of the surface of revolution supported by two circles. According to the calculus, the area J of the surface is

$$J = \pi \int_0^a y\sqrt{1+y'^2}\, dx$$

This problem is again a special case of the geometric optic, corresponding to $\psi(y) = y$. Equation (4.2) becomes

$$x = \pm\Phi(u) = \int \frac{dy}{\sqrt{c^2y^2-1}} = \frac{1}{C}\cosh^{-1}(Cy)$$

and we find

$$y(x) = \frac{1}{C}\cosh\left(C(x-x_0)\right) + c_1$$

Assume for clarity that the surface is supported by two equal circle parted symmetric to OX axis; the equation (13.21) becomes

$$Cy = \cosh\left(Cx\right)$$

The family of extremal lies inside the triangle $\frac{|x|}{y} \leq 2/3$. Analysis of this formula reveals unexpected features: The solution may be either unique, or has two different solutions (in which case, the one with smaller value of the objective functional must be selected) or it may not have solutions at all. The last case looks strange because the problem of minimal area obviously has a solution.

The defect in our consideration is the following: We tacitly assumed that the minimal surface of revolution is a differentiable curve with finite tangent y' to the axis of revolution. There is another solution: Two circles and an infinitesimal bar between them. The objective functional is

$$I_0 = \pi(R_1^2 + R_2^2).$$

The minimizer (the Goldschmidt solution) is a distribution

$$y = -R_1\delta(x-a) + R_2(\delta(x-b)$$

where $\delta(x)$ is the delta-function. Obviously, this minimizer does not belong to the presumed class of twice-differentiable functions.

From geometrical perspective, the problem should be correctly reformulated as the problem for the best parametric curve $[x(t), y(t)]$ then $y' = \tan\alpha$ where α is the angle of inclination to OX axis. The equation (4.3) that takes the form

$$y \cos\alpha = C$$

admits either regular solution $C \neq 0$ and $y = C \sec\alpha$, $C \neq 0$ which yields to the catenoid (13.21), or the singular solution $C = 0$ and either $y = 0$ or $\alpha = \frac{\pi}{2}$ which yield to Goldschmidt solution.

Geometric optics suggests a physical interpretation of the result: The problem of minimal surface is formally identical to the problem of the quickest path between two equally distanced from OX-axis points, if the speed $v = 1/y$ is inverse proportional to the distance to the axis OX. The optimal path between the two close-by points lies along the arch of catenoid $\cosh(z)$ that passes through the given end points. In order to cover the distance quicker, the path sags toward the OX-axis where the speed is larger.

The optimal path between two far-away points is different: The particle goes straight to the OX-axis where the speed is infinite, than transports instantly (infinitely fast) to the closest to the destination point at the axis, and goes straight to the destination. This "Harry Potter Transportation Strategy" is optimal when two supporting circles are sufficiently far away from each other.

In spite of these clarifications, the concern still remain because geometric explanation is not always available. We need a formal analysis of the discontinuous solution and δ-function-type derivative of an extremal. The analytical tests that are able to detect such unexpected unbounded solutions in a regular manner are discussed later, in Chapter 7.

4.1.4 Geodesics on an explicitly given surface

The problem of shortest on a surface path between two points on this surface is called the problem of geodesics. We dealt with it in the Introduction. Now we are able to formulate it as a variational problem

$$I = \min_{s(t)} \int_{t_0}^{t_1} ds$$

where $s(t)$ is the arch on a surface, and t is a parameter. Depending on the used representation of the surface, the problem can be formulated in several ways.

Geodesics on an explicitly given surface Assume that the surface is given by an explicit relation $z = \psi(x, y)$ and the geodesics is an spacial curve which coordinates are given by an explicit formula $[x, y(x), \psi(x, y(x))]$. The unknown function $y(x)$ is the projection of geodesics on XY plane. In this case, the infinitesimal distance ds along the surface can be found from Pythagorean relation

$ds^2 = dx^2 + dy^2 + dz^2$ where

$$dy = y' dx, \quad dz = \left(\frac{\partial \psi}{\partial y} y' + \frac{\partial \psi}{\partial x} \right) dx.$$

The Lagrangian – an infinitesimal length ds becomes

$$ds = \sqrt{1 + y'^2 + \left(\frac{\partial \psi}{\partial x} + \frac{\partial \psi}{\partial y} y' \right)^2} \, dx$$

Check that The Euler equation for $y(x)$ is:

$$C \frac{d^2 y}{dx^2} = A_3 \left(\frac{dy}{dx} \right)^3 - A_2 \left(\frac{dy}{dx} \right)^2 + A_1 \left(\frac{dy}{dx} \right) - A_0$$

where

$$C = 1 + \left(\frac{\partial \psi}{\partial x} \right)^2 + \left(\frac{\partial \psi}{\partial y} \right)^2$$

is the square of the surface area and

$$\begin{aligned} A_0 &= \frac{\partial \psi}{\partial y} \frac{\partial^2 \psi}{\partial x^2}, \qquad & A_1 &= \frac{\partial \psi}{\partial x} \frac{\partial^2 \psi}{\partial x^2} - 2 \frac{\partial \psi}{\partial y} \frac{\partial^2 \psi}{\partial x \partial y}, \\ A_3 &= \frac{\partial \psi}{\partial x} \frac{\partial^2 \psi}{\partial y^2}, \qquad & A_2 &= \frac{\partial \psi}{\partial y} \frac{\partial^2 \psi}{\partial y^2} - 2 \frac{\partial \psi}{\partial x} \frac{\partial^2 \psi}{\partial x \partial y}. \end{aligned}$$

When $\psi =$ constant(x) or $\psi =$ constant(x), the equation becomes ..
Problems: Find geodesics on cone, hyperboloid, paraboloid.

Geodesics on the sphere In some problem, it is natural to use curved coordinate frame: Find the path of minimal length on a unit sphere D between two points at this sphere. In spherical coordinates, the positions the two points are ϕ_0, θ_0 and ϕ_1, θ_1 where ϕ is the latitude and θ is the longitude. The infinitesimal distance ds is found from Pythagorean triangle:

$$ds^2 = \sin^2 \theta (d\phi)^2 + (d\theta)^2$$

Assuming that $\phi = \phi(\theta)$ we have $d\phi = \phi' d\theta$ and

$$D = \min_{\phi(\theta)} \int_{\theta_0}^{\theta_1} \sqrt{(\phi')^2 \sin^2 \theta + 1} \, d\theta, \quad \phi(\theta_0) = \phi_0, \ \phi(\theta_1) = \phi_1$$

The Lagrangian is independent of ϕ; there exist the first integral (see (13.21))

$$\frac{\phi' \sin^2 \theta}{\sqrt{(\phi')^2 \sin^2 \theta + 1}} = c$$

Solve for ϕ':

$$\phi' = \frac{d\phi}{d\theta} = \frac{c}{\sin\theta\sqrt{\sin^2\theta - c^2}}$$

and integrate

$$\phi(\theta) = \phi_0 + c\int_{\theta_0}^{\theta} \frac{d\theta}{\sin\theta\sqrt{\sin^2\theta - c^2}}$$

To define c, we use the condition $\phi(\theta_1) = \phi_1$.

Proof the geodesics is a great circle.

Remark 4.1.3 A geometric proof was discussed earlier in ??

Geodesics through the metric tensor Properties of geodesics characterize the surface, or, more generally, a manifold in a metric space. For example, geodesics are unique in simple-connected spaces with negative curvatures; in spaces with positive curvatures there may be two or more geodesics that joint two points and one has the choose the shortest path, using *Calculus of variation in the large* that utilizes topological methods to investigate extremals on manifolds, see Leng, Rashevsky, Milior. Geodesics naturally determine the tensor of curvature in space; in general relativity, the curvature of light rays which represented by the geodesics allows for physical interpretation of the curved time-space continuum. These problems are beyond the scope of this book.

Here we only derive the equations of geodesics through the metric tensor of a surface. Suppose that x_1, x_2 are the coordinates on the surface, similar to the coordinates θ, ϕ on a sphere. We start with generalization of Pythagorean theorem in curved coordinates:

$$ds^2 = g_{ij}dx_i dx_j$$

where g_{ij} is called the metric tensor. The problem of geodesics is: Minimize the path

$$\int ds = \int \sqrt{g_{ij}\dot{x}_i\dot{x}_j}dt.$$

The Euler equation for the problem,

$$\left(\frac{d}{dt}\frac{\partial}{\partial\dot{x}_i} - \frac{\partial}{\partial x_i}\right)\sqrt{g_{ij}\dot{x}_i\dot{x}_j} = 0$$

can be transformed to the form

$$\frac{d^2x_k}{ds^2} + ,^{k}_{ij}\frac{dx^i}{ds}\frac{dx^j}{ds}, \quad k = 1, 2$$

where $,^{k}_{ij}$ is the Christoffel symbol defined as

$$,^{k}_{ij} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x_j} + \frac{\partial g_{jk}}{\partial x_i} - \frac{\partial g_{ij}}{\partial x_k}\right)$$

Examples

4.2 Approximations with penalties

Consider the problem of approximation of a function by another one with better smoothness or other favorable properties. For example, we may want to approximate the noisy experimental curve by a smooth one, or approximate a curve with a block-type piece-wise constant curve. The following method is used for approximations: A variational problem is formulated to minimize the integral of the square of the difference of the approximating function and the approximate plus a penalty imposed the approximate for being non-smooth or having its non-zero variation. The approximate compromises the closeness to the approximating curve and the smoothness properties. Here we consider several problem of the best approximation.

4.2.1 Approximation with penalized growth rate

The problem of the best approximation of the given function $h(x)$ by function $u(x)$ with a limited growth rate results a variational problem

$$\min_u J(u), \quad J(u) = \int_a^b \frac{1}{2}\left(\alpha u'^{\,2} + (h-u)^2\right) dx \tag{4.5}$$

Here, $\alpha \geq 0$, the first term of the integrant represents the penalty for growth and the second term describes the quality of approximation: the closeness of the original and the approximating curve. The approximation depends on the parameter α: When $\alpha \to 0$, the approximation coincides with $h(x)$ and when When $\alpha \to \infty$, the approximation is a constant curve, equal to the mean value of $h(x)$.

The equation for the approximate (Euler equation of (4.8)) is

$$\alpha u'' - u = h, \quad u'(a) = u'(b) = 0$$

Here, the natural boundary conditions are assumed since there is no reason to assign special values of the approximation curve at the ends of the interval.

Integrating the Euler equation, we find

$$u(x) = \frac{1}{b-a}\int_a^b G(x,y)h(y)dy$$

where Green's function $G(x,y)$ (see next section) is

$$G(x,y) =$$

4.2.2 About Green's function

Green's function for approximations with quadratic penalty The solution of a linear boundary value problem is most conveniently done by the Green's function. Here we remind this technique.

Consider the linear differential equation with the differential operator L

$$L(x)u(x) = f(x) \quad x \in [a,b], \quad B_a(u,u')|_{x=a} = 0, \quad B_b(u,u')|_{x=a} = 0. \tag{4.6}$$

an arbitrary external excitation $f(x)$ and homogeneous boundary conditions $B_a(u,u')|_{x=a} = 0$ and $B_b(u,u')|_{x=a} = 0$. For example, the problem (**??**) corresponds to

$$L(x)\,u = \left(\alpha^2 \frac{d^2}{dx^2} - 1\right) u, \quad B_a(u,u') = u', \quad B_b(u,u') = u'$$

To solve the equation means to invert the dependence between u and f, that is to find the linear operator

$$u = L^{-1} f$$

In order to solve the problem (4.6) one solves first the problem for a single concentrated load $\delta(x-\xi)$ applied at the point $x = \xi$

$$L(x)g(x,\xi) = \delta(x-\xi), \quad B_a(g,g')|_{x=a} = 0 \quad B_b(g,g')|_{x=a} = 0$$

This problem is usually simpler than (4.6). The solution $g(x,\xi)$ is called the Green's function, it depends on the point of the applied excitation ξ as well as of the point x where the solution is evaluated. Formally, the Green's function can be expressed as

$$g(x,\xi) = L(x)^{-1}\delta(x-\xi) \tag{4.7}$$

Then, we use the identity

$$f(x) = \int_a^b f(x)\delta(x-\xi)d\xi$$

(essentially, the definition of the delta-function) to find the solution of (4.6). We multiply both sides of (4.7) by $f(\xi)$ and integrate over ξ from a to b, obtaining

$$\int_a^b g(x,\xi)f(\xi)d\xi = L^{-1}\left(\int_a^b f(\xi)\delta(x-\xi)d\xi\right) = L^{-1}f(x) = u(x).$$

Notice that operator $L = L(x)$ is independent of ξ therefore we can move L^{-1} out of the integral over ξ.

Thus, we obtain the solution,

$$u(x) = \int_a^b g(x,\xi)f(\xi)d\xi$$

that expresses $u(x)$ as a linear mapping of $f(x,\xi)$ with the kernel $G(x,\xi)$. The finite-dimentional version of this solution is the matrix equation for the vector u.

Green's function for approximation at an interval For the problem (**??**), the problem for the Green's function is

$$\left(\alpha^2 \frac{d^2}{dx^2} - 1\right) g(x, \xi) = \delta(x - \xi), \quad u'(a) = u'(b) = 0$$

At the intervals $x \in [a, \xi)$ and $x \in (\xi, b]$ the solution is

$$g(x, \xi) = \begin{cases} g_-(x, \xi) = A_1 \cosh\left(\frac{x-a}{\alpha}\right) & \text{if } x \in [a, \xi) \\ g_+(x, \xi) = A_2 \cosh\left(\frac{x-b}{\alpha}\right) & \text{if } x \in (\xi, b] \end{cases}$$

This solution satisfies the differential equation for all $x \neq \xi$ and the boundary conditions. At the point of application of the concentrated force $x = \xi$, the conditions hold

$$g_+(\xi, \xi) = g_-(\xi, \xi); \quad \frac{d}{dx} g_+(x, \xi)\bigg|_{x=\xi} - \frac{d}{dx} g_-(x, \xi)\bigg|_{x=\xi} = 1$$

that express the continuity of $u(x)$ and the unit jump of the derivative $u'(x)$. These allow for determination of the constants

$$A_1 = \alpha \frac{\cosh\left(\frac{\xi-b}{\alpha}\right)}{\sinh\left(\frac{b-a}{\alpha}\right)} \quad A_2 = -\alpha \frac{\cosh\left(\frac{\xi-a}{\alpha}\right)}{\sinh\left(\frac{b-a}{\alpha}\right)}$$

which completes the calculation.

Green's function for approximation in R_1 The formulas for the Green's function are simpler when the approximation of an integrable in R_1 function $f(x)$ is performed over the whole real axes, or when $a \to -\infty$ and $b \to \infty$. In this case, the boundary terms $u'(a) = u'(b) = 0$ are3 replaced by requirement that the approximation u is finite,

$$u(x) < \infty \quad \text{when } x \to \pm\infty$$

In this case, the Green's function is

$$g(x, \xi) = \frac{1}{2\alpha} e^{\frac{|x-\xi|}{\alpha}}$$

One easily check that it satisfies the differential equation, boundary conditions, and continuity and jump conditions at $x = \xi$.

The best approximation becomes simply an average

$$u(x) = \frac{1}{2\alpha} \int_{-\infty}^{\infty} f(\xi) e^{\frac{|x-\xi|}{\alpha}} d\xi$$

with the exponential kernel $e^{\frac{x-\xi}{\alpha}}$.

4.2.3 Approximation with penalized smoothness

The problem of smooth approximation is similarly addressed but the penalization functional is differently defined. This time the approximate is penalized for being different form a straight line by the integral of the square of the second derivative u''. The resulting variational problem reads

$$\min_u J(u), \quad J(u) = \int_a^b \frac{1}{2}\left(\alpha(u'')^2 + (h-u)^2\right) dx \tag{4.8}$$

Here, $\alpha \geq 0$, the first term of the integrant represents the penalty for non-smoothness and the second term describes the closeness of the original curve and the approximate. When $\alpha \to 0$, the approximation coincides with $h(x)$ and when $\alpha \to \infty$, the approximation is a straight line closest to h.

The equation for the approximate (Euler equation of (4.8)) is

$$\alpha u^{IV} + u = h, \quad u''(a) = u''(b) = 0, \quad u'''(a) = u'''(b) = 0,$$

Here, the natural boundary conditions are assumed since there is no reason to assign special values of the approximation curve at the ends of the interval.

Integrating the Euler equation, we find

$$u(x) = \frac{1}{b-a}\int_a^b G(x,y)h(y)dy$$

where Green's function $G(x,y)$ (see []) is

$$G(x,y) =$$

4.2.4 Approximation with penalized total variation

This approximation penalizes the function for its total variation. The total variation $T(f)$ of a function u is defined as

$$T(u) = \int_a^b |u'(x)|dx$$

For a monotonic function u one evaluates the integral and finds that

$$T(u) = \max_{x\in[a,b]} u(x) - \min_{x\in[a,b]} u(x)$$

If $u(x)$ has finite number N of intervals L_k of monotonicity then the total variation is

$$T(u) = \sum_k^N \left(\max_{x\in L_k} u(x) - \min_{x\in L_k} u(x)\right)$$

The variational problem with total-variation penalty has the form

$$\min_u J(u), \quad J(u) = \int_a^b \frac{1}{2}\left(\gamma|u'| + (h-u)^2\right) dx \tag{4.9}$$

Here, $\alpha \geq 0$, the first term of the integrant represents the total-variation penalty and the second term describes the closeness of the original curve and the approximate. When $\alpha \to 0$, the approximation coincides with $h(x)$ and when $\alpha \to \infty$, the approximation becomes constant equal to mean value of h.

The formal application of Euler equation:

$$(\gamma \text{sign}\,(u'))' + u = h, \quad \text{sign}\,(u'(a)) = \text{sign}\,(u'(b)) = 0 \tag{4.10}$$

is not too helpful because it requires the differentiation of a discontinuous function sign ; besides, the Lagrangian (4.9) is not a twice-differential function of u' as it is required in the procedure of derivation of the Euler equation.

Let us reformulate the problem in a *regularized form*, noticing that

$$\int_a^b |u'(x)| dx = \lim_{\epsilon \to +0} \int_a^b \sqrt{u'(x)^2 + +\epsilon^2} dx$$

and replacing the former by the later in the problem (4.9). We fix $\epsilon > 0$, derive the necessary conditions and analyze them assuming that $\epsilon \to +0$. The Euler equation is more regular,

$$(k(u', \epsilon))' = h - u + O(\epsilon), \quad u'(a) = u'(b) = 0 \tag{4.11}$$

where

$$k(u', \epsilon) = \frac{u'}{(|u'|^2 + \epsilon^2)^{\frac{1}{2}}}$$

Remark 4.2.1 Scale ϵ

The term $k(u', \epsilon)$ is ϵ- close to one outside of the $\sqrt{\epsilon}$-neighborhood of zero, $|u' \geq \sqrt{\epsilon}$, $k \in (0, \epsilon^{\frac{3}{2}})$. Inside this neighborhood, is unbounded $k \in [0, \epsilon^{\frac{2}{2}})$.

The stationary condition (4.11) is satisfied (up to the order of ϵ) in one of two ways. When $u = f$ and $|u'| = |f'| > \epsilon$, the first term $(k(u', \epsilon))'$ is of the order smaller than ϵ and it does not influence the condition. Indeed, $k(u', \epsilon)$ is approximately equal to one no matter what the value of $|u'| \geq \epsilon$ is. When $u \approx$ constant and $|u'| \leq \epsilon$, the first term is extremely sensitive to the variation of u' and it can take any value; in particular, it can compensate the second term $u - f$ of the equality. This rough analysis shows that in the limit $\epsilon \to 0$, the stationary condition (4.11) is satisfied either when $u(x)$ is a constant, $u' = 0$, or when $u(x)$ coincides with $h(x)$.

$$u(x) = h(x) \quad \text{or } u'(x) = 0, \quad \forall x \in [a, b]$$

The approximation cuts the maxima and minima of the approximating function.

Let us find the cutting points. For simplicity in notations we assume that the function u monotonically increases at $[a, b]$. The approximation u is also a monotonically increasing function, $u' \geq 0$ that either coincides with $h(x)$ or stays constant cutting the maximum and the minimum of $h(x)$:

$$u(x) = \begin{cases} h(\alpha) & \text{if } x \in [a, \alpha] \\ h(x) & \text{if } x \in [\alpha, \beta] \\ h(\beta) & \text{if } x \in [\beta, b] \end{cases}$$

The cost of the problem

$$J = \frac{\gamma}{2}\left[\int_a^\alpha (h(x)-h(\alpha))^2 dx + \int_\beta^b (h(x)-h(\beta))^2 dx\right] + h(\beta) - h(\alpha)$$

depends on two unknown parameters, α and β, the coordinates on the cuts. They are found by straight differentiation. The equation for α is

$$\frac{dJ}{d\alpha} = \gamma\left[\frac{1}{2}(h(x)-h(\alpha))^2|_{x=a} + h'(a)\int_a^\alpha (h(x)-h(\alpha))dx\right] - h'(a) = 0$$

or, noticing the cut point α is not a stationary point, $h'(a) \neq 0$

$$\int_a^\alpha [h(x)-h(\alpha)]dx = \frac{1}{\gamma}$$

the equation for β is similar:

$$\int_\beta^b [h(x)-h(\beta)]dx = \frac{1}{\gamma}$$

Notice that the extremal is broken; regular variational method based the Euler equation is not effective. These irregular problems will be discussed later in Chapter 7.

4.3 Lagrangian mechanics

Leibnitz and Mautoperie suggested that any motion of a system of particles minimizes a functional of action; later Lagrange came up with the exact definition of that action: the functional that has the Newtonian laws of motion as its Euler equation. The question whether the action reaches the true minimum is more complicated: Generally, it does not; Nature is more sophisticated and diverse than it was expected. We will show that the true motion of particles settles for a local minimum or even a saddle pint of action' each stationary point of the functional correspond to a motion with Newtonian forces. As a result of realizability of local minima, there are many ways of motion and multiple equilibria of particle system which make our world so beautiful and unexpected (the picture of the rock). The variational principles remain the abstract and economic way to describe Nature but one should be careful in proclaiming the ultimate goal of Universe.

4.3.1 Stationary Action Principle

Lagrange observed that the second Newton's law for the motion of a particle,

$$m\ddot{x} = f(x)$$

can be viewed as the Euler equation to the variational problem

$$\min_{x(t)} \int_{t_0}^{t_f} \left(\frac{1}{2}m\dot{x}^2 - V(x)\right) dx$$

where V is the negative of antiderivative (potential) of the force f.

$$V = -\int f(x)dx$$

The minimizing quantity – the difference between kinetic and potential energy – is called *action*; The Newton equation for a particle is the Euler equations.

In the stated form, the principle is applicable to any system of free interacting particles; one just need to specify the form of potential energy to obtain the Newtonian motion.

Example 4.3.1 (Central forces) For example, the problem of celestial mechanics deals with system bounded by gravitational forces f_{ij} acting between any pair of masses m_i and m_j and equal to

$$f_{ij} = \gamma \frac{m_i m_j}{|r_i - r_j|^3}(r_i - r_j)$$

where vectors r_i define coordinates of the masses m_i as follows $r_i = (x_i, y_i, z_j)$. The corresponding potential V for the n-masses system is

$$V = -\frac{1}{2}\sum_{i,j}^{N} \gamma \frac{m_i m_j}{|r_i - r_j|}$$

where γ is Newtonian gravitational constant. The kinetic energy T is the sum of kinetic energies of the particles

$$T = \frac{1}{2}\sum_{i}^{N} m_i \dot{r}_i^2$$

The motion corresponds to the stationary value to the Lagrangian $L = T - V$, or the system of N vectorial Euler equations

$$m_i \ddot{r}_i - \sum_{j}^{N} \gamma \frac{m_i m_j}{|r_i - r_j|^3}(r_i - r_j) = 0$$

for N vector-function $r_i(t)$.

Since the Lagrangian is independent of time t, the first integral (13.21) exist

$$T + V = \text{ constant}$$

which corresponds to the conservation of the whole energy of the system.

Later in Section 13.21, we will find other first integrals of this system and comment about properties of its solution.

Example 4.3.2 (Spring-mass system) Consider the sequence of masses Consider the sequence of masses lying on an axis with coordinates $m_1, \ldots, m_n$ lying on an axis with coordinates $x_1, \ldots, x_n$ joined by the sequence of springs between two sequential masses. Each spring generate force f_i proportional to $x_i - x_{i+1}$ where $x_i - x_{i+1} - l_i$ is the distance between the masses and l_i correspond to the resting spring.

Let us derive the equations of motion of this system. The kinetic energy T of the system is equal to the sum of kinetic energies of the masses,

$$T = \frac{1}{2}m(\dot{x_1} + \ldots + \dot{x_n})$$

the potential energy V is the sum of energies of all springs, or

$$V = \frac{1}{2}C_1(x_2 - x_1)^2 + \ldots + \frac{1}{2}C_{n-1}(x_n - x_{n-1})^2$$

The Lagrangian $L = T - V$ correspond to n differential equations

$$\begin{aligned} m_1\ddot{x}_1 &+ C_1(x_1 - x_2) = 0 \\ m_2\ddot{x}_2 &+ C_2(x_2 - x_3) - C_1(x_1 - x_2) = 0 \\ \ldots & \quad \ldots \\ m_n\ddot{x}_n &- C_{n-1}(x_{n-1} - x_n) = 0 \end{aligned}$$

or in vector form

$$M\ddot{x} = P^T C P x$$

where $x = (m_1, \ldots, x_n)$ is the vector of displacements, M is the $n \times n$ diagonal matrix of masses, V is the $(n-1) \times (n-1)$ diagonal matrix of stiffness, and P is the $n \times (n-1)$ matrix that shows the operation of difference,

$$M = \begin{pmatrix} m_1 & 0 & \ldots & 0 \\ 0 & m_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & m_n \end{pmatrix}, \quad C = \begin{pmatrix} C_1 & 0 & \ldots & 0 \\ 0 & C_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & C_{n-1} \end{pmatrix}, \quad P = \begin{pmatrix} 1 & -1 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix},$$

When the masses and the springs are identical, $m_1 = \ldots = m_n = m$ and $C_1 = \ldots = C_{n-1} = C$, the system simplifies to

$$\begin{aligned} m_1\ddot{x}_1 &+ C_1(x_1 - x_2) = 0 \\ m_2\ddot{x}_2 &+ C_2(x_2 - x_3) - C_1(x_1 - x_2) = 0 \\ \ldots & \quad \ldots \\ m_n\ddot{x}_n &- C_{n-1}(x_{n-1} - x_n) = 0 \end{aligned}$$

or in vector form,

$$\ddot{x} + \frac{C}{m}P_2 x = 0$$

where P_2 is the $n \times n$ matrix of second differences,

$$P_2 = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 \end{pmatrix},$$

4.3.2 Generalized coordinates

The Lagrangian concept allows for obtaining equations of motion of a constrained system. In this case, the kinetic and potential energy must be defined as a function of *generalized coordinates* that describes degrees of freedom of motion consistent with the constraints. The constraints are be accounted either by Lagrange multipliers or directly, by introducing *generalized coordinates.* If a particle can move along a surface, one can introduced coordinates on this surface and allow the motion only along these coordinates.

The particles can move along the generalized coordinates q_i. Their number corresponds to the allowed degrees of freedom. The position x allowed by constraints becomes $x(q)$. The speed $\dot{x}$ becomes a linear form of $\dot{q}$

$$\dot{x} = \sum \left(\frac{\partial x}{\partial q_i} \dot{q}_i \right)$$

For example, a particle can move along the circle of the radius R, the generalized coordinate will be an angle θ which determines the position $x_1 = R\cos\theta$, $x_2 = R\sin\theta$ at this circle and its speed becomes

$$\dot{x}_1 = -R\dot{\theta}\sin\theta, \quad x\dot{x}_2 = R\dot{\theta}\cos\theta$$

This system has only one degree of freedom, because fixation of one parameter θ completely defines the position of a point.

When the motion is written in terms of generalized coordinates, the constraints are automatically satisfied. Let us trace equations of Lagrangian mechanics in the generalized coordinates. It is needed to represent the potential and kinetic energies in these terms. The potential energy $V(x)$ is straightly rewritten as $W(q) = V(x(q))$ and the kinetic energy $T(\dot{x}) = \sum_i m_i \dot{x}_i^2$ becomes a quadratic form of derivatives of generalized coordinates $\dot{\boldsymbol{q}}$

$$T(\dot{x}) = \sum_i m_i x_i^2 = \dot{\boldsymbol{q}}^T R(\boldsymbol{q}) \dot{\boldsymbol{q}}$$

where the symmetric nonnegative matrix R is equal to

$$R = \{R_{ij}\}, \quad R_{ij} = \left(\frac{\partial T}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}}{\partial q_i} \right)^T \left(\frac{\partial T}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}}{\partial q_j} \right)$$

Notice that $T_q(\dot{q})$ is a homogeneous quadratic function of $\dot{q}$, $T_q(k\dot{q}) = k^2 T_q(\dot{q})$ and therefore

$$\frac{\partial}{\partial \dot{\boldsymbol{q}}} T_q(\boldsymbol{q}, \dot{\boldsymbol{q}}) \cdot \dot{\boldsymbol{q}} = 2T_q(\boldsymbol{q}, \dot{\boldsymbol{q}}) \tag{4.12}$$

the variational problem that correspond to minimal action with respect to generalized coordinates becomes

$$\min_{\boldsymbol{q}} \int_{t_0}^{t_1} (T_q - V_q) dt \tag{4.13}$$

Because potential energy V does not depend on $\dot{q}$, the Euler equations have the form

$$\frac{d}{dt}\frac{\partial T_q}{\partial \dot{\boldsymbol{q}}} - \frac{\partial}{\partial \boldsymbol{q}}(T_q - V_q) = 0 \tag{4.14}$$

which is similar to the form of unrestricted motion.

The analogy can be continued. When the Lagrangian is independent of t the system is called *conservative*, In this case, the Euler equation assumes the first integral in the form (use (4.12))

$$\dot{\boldsymbol{q}}\frac{\partial T_q}{\partial \dot{\boldsymbol{q}}} - (T_q - V_q) = T_q + V_q = \text{constant}(t) \tag{4.15}$$

The quantity $\Pi = T_q + V_q$ is called the whole energy of a mechanical system; it is preserved along the trajectory.

The generalized coordinates help to formulate differential equations of motion of constrained system. Consider several examples

Example 4.3.3 (Isochrone) Consider a motion of a heavy mass along the cycloid:

$$x = \theta - \cos\theta, \quad y = \sin\theta$$

To derive the equation of motion, we write down the kinetic T and potential V energy of the mass m, using $q = \theta$ as a generalized coordinate. We have

$$T = \frac{1}{2} m\dot{x}^2 + \dot{y}^2 = m(1 + \sin\theta)\dot{\theta}^2$$

and $V = my = -m\sin\theta$.

The Lagrangian

$$L = T - V = m(1 + \sin\theta)\dot{\theta}^2 + m\sin\theta$$

allows to derive Euler equation

$$S(\theta, \dot{\theta}) = \frac{d}{dt}\left((1 + \sin\theta)\frac{d\theta}{dt}\right) - \cos\theta = 0.$$

which solution is

$$\theta(t) = \arccos(C_1 \sin t + C_2 \cos t)$$

where C_1 and C_2 are constant of integration. One can check that $\theta(t)$ is 2π-periodic for all values of C_1 and C_2. This explains the name "isochrone" given to the cycloid before it was found that this curve is also the brachistochrone (see Section **??**)

Example 4.3.4 (Winding around a circle) Describe the motion of a mass m tied to a cylinder of radius R by a rope that winds around it when the mass evolves around the cylinder. Assume that the thickness of the rope is negligible small comparing with the radius R, and neglect the gravity.

It is convenient to use the polar coordinate system with the center at the center of the cylinder. Let us compose the Lagrangian. The potential energy is zero, and the kinetic energy is

$$L = T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$$
$$= \frac{1}{2}m\left(\dot{r}\cos\theta - r\dot{\theta}\sin\theta\right)^2 + \frac{1}{2}m\left(\dot{r}\sin\theta + r\dot{\theta}\cos\theta\right)^2$$
$$= \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right)$$

The coordinates $r(t)$ and $\theta(t)$ are algebraically connected by Pythagorean relation $R^2 + l(t)^2 = r(t)^2$ at each time instance t. Here $l(t)$ is the part of the rope that is not winded yet; it is expressed through the angle $\theta(t)$ and the initial length l_0 of the rope, $l(t) = l_0 - R\theta(t)$. We obtain

$$(l_0 - R\theta(t))^2 = r(t)^2 - R^2 \quad \forall t \in [0, t_{\text{final}}],$$

and observe that the time of winding t_{final} is finite. The trajectory $r(\theta)$ is a spiral.

The obtained relation allows for linking of $\dot{r}$ and $\dot{\theta}$. We differentiate it and obtain

$$r\dot{r} = -R(l_0 - R\theta(t))\dot{\theta} = -R(\sqrt{r^2 - R^2}\dot{\theta}$$

or

$$\dot{\theta} = -\frac{\dot{l}}{R} = -\frac{r\dot{r}}{R\sqrt{r^2 - R^2}}$$

The Lagrangian becomes

$$L(r, \dot{r}) = \frac{1}{2}m\dot{r}^2\left(1 + \frac{r^4}{R^2(r^2 - R^2)}\right)$$

Its first integral

$$\frac{1}{2}m\dot{r}^2\left(1 + \frac{r^4}{R^2(r^2 - R^2)}\right) = C$$

shows the dependence of the speed $\dot{r}$ on the coordinate r. It can be integrated in a quadratures, leading to the solution

$$t(r) = C_1\int_{r_0}^{r}\sqrt{\frac{r^2 - R^2}{r^4 + R^2r^2 - R^4}}dx$$

The two constants r_0 and C_1 are determined from the initial conditions.

The first integral allows us to visualize the trajectory by plotting $\dot{r}$ versus r. Such graph is called the phase portrait of the trajectory.

4.3.3 More examples: Two degrees of freedom.

Example 4.3.5 (Move through a funnel) Consider the motion of a heavy particle through a vertical funnel. The axisymmetric funnel is described by the equation $z = \phi(r)$ in cylindrical coordinate system. The potential energy of the particle is proportional to z, $V = -mgz = -mg\phi(r)$ The kinetic energy is

$$T = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2\right)$$

or, accounting that the point moves along the funnel,

$$T = \frac{1}{2}m\left((1+\phi'^2)\dot{r}^2 + r^2\dot{\theta}^2\right).$$

The Lagrangian

$$L = T - V = \frac{1}{2}m\left((1+\phi'^2)\dot{r}^2 + r^2\dot{\theta}^2\right) + mg\phi(r)$$

is independent of the time t and the angle θ, therefore two first integrals exist:

$$\frac{\partial L}{\partial \dot{\theta}} = \mu \quad \Rightarrow \; \dot{\theta} = \frac{\mu}{r^2}$$

and

$$T + V = \frac{1}{2}m\left((1+\phi'^2)\dot{r}^2 + r^2\dot{\theta}^2\right) - mg\phi(r) = \Pi$$

The second can be simplified by excluding $\dot{\theta}$ using the first,

$$\Pi = \frac{1}{2}m\left((1+\phi'^2)\dot{r}^2 + \frac{\mu^2}{r^2} - g\phi(r)\right)$$

Here, the constants Π and μ can be defined from the initial conditions. They represent, respectively, the whole energy of the system and the angular momentum; these quantities are conserved along the trajectory. These integrals alone allow for integration of the system, without computing the Euler equations. Solving for $\dot{r}$, we find

$$(\dot{r^2})^2 = 2\frac{\left(\frac{2\Pi}{m} + g\phi(r)\right)r^2 - \mu^2}{1+\phi'^2}$$

Consequently, we can find $r(t)$ and $\theta(t)$ (see Problem **??**.

A periodic trajectory corresponds to constant value $\dot{\theta}(t)$ and constant value of $r(t) = r_0$ which is defined by the initial energy, angular momentum, and the shape $\phi(r)$ of the funnel, and satifies the equation

$$\frac{\mu^2}{r_0^2} - g\phi(r_0) = \frac{2\Pi}{m}$$

This equation does not necessary has a solution. Physically speaking, a heavy particle can either tend to evolve around the funnel, or fall down it.

Example 4.3.6 (Three-dimensional pendulum) A heavy mass is attached to a hitch by a rod of unit length. Describe the motion of the mass. Since the mass moves along the spherical surface, we introduce a spherical coordinate system with the center at the hitch. The coordinates of mass are expressed through two spherical angles ϕ and θ which are the generalized coordinates. We compute

$$T = \dot{\phi}^2 + \dot{\theta}^2 \cos\phi$$

and

$$V = g\cos\phi$$

Two conservation laws follows

$$\dot{\theta}\cos\phi = \mu \tag{4.16}$$

(conservation of angular momentum) and

$$m(\dot{\phi}^2 + \dot{\theta}^2 \cos\phi) + g\cos\phi = \Pi \tag{4.17}$$

(conservation of energy)

The oscillations are described by these two first-order equations for ϕ and θ. The reader is encouraged to use Maple to model the motion.

Two special cases are immediately recognized. When $\mu = 0$, the pendulum oscillates in a plane, $\theta(t) = \theta_0$, and $\dot{\theta} = 0$. The Euler equation for ϕ becomes

$$m\ddot{\phi} + g\sin\phi = 0$$

This is the equation for a plane pendulum. The angle $\phi(t)$ is a periodic function of time, the period depends on the magnitude of the oscillations. For small θ, the equation becomes equation of linear oscillator.

When $\phi(t) = \phi_0 = \text{constant}$, the pendulum oscillates around a horizontal circle. In this case, the speed of the pendulum is constant (see (4.16)) and the generalized coordinate – the angle θ is

$$\theta = \frac{\mu}{\cos\phi_0} t + \theta_0$$

The motion is periodic with the period

$$T = \frac{2\pi\cos\phi_0}{\mu}$$

Next example illustrates that the Euler equations for generalized coordinates are similar for the simplest Newton equation $m\ddot{x} = f$ but m becomes a non-diagonal matrix.

Example 4.3.7 (Two-link pendulum) Consider the motion of two masses sequentially joined by two rigid rods. The first rod of length l_1 is attached to a hitch and can evolve around it and it has a mass m_1 on its other end. The second rod of the length l_2 is attached to this mass at one end and can evolve around it, and has a mass m_2 at its other end. Let us derive equation of motion of this system in the constant gravitational field.

The motion is expressed in terms of Cartesian coordinates of the masses x_1, y_1 and x_2, y_2. We place the origin in the point of the hitch: This is the natural stable point of the system. The distances between the hitch and the first mass, and between two masses are fixed,

$$l_1 = x_1^2 + y_1^2, \quad l_2 = (x_1 - x_2)^2 + (y_1 - y_2)^2,$$

which reduces the initial four degrees of freedom to two. The constraints are satisfied if we introduce two generalized coordinates: two angles θ_1 and θ_2 of the corresponding rods to the vertical, assuming

$$\begin{aligned} x_1 &= l_1 \cos\theta_1, \quad y_1 = l_1 \sin\theta_1 \\ x_2 &= l_1 \cos\theta_1 + l_2 \cos\theta_2, \quad y_2 = l_1 \sin\theta_1 + l_2 \sin\theta_2. \end{aligned}$$

The state of the system is defined by the angles θ_1 and θ_2.

The potential energy $V(\theta_1, \theta_2)$ is equal to the sum of vertical coordinates of the masses multiplied by the masses,

$$V(\theta_1, \theta_2) = m_1 y_1 + m_2 y_2 = m_1 l_1 \cos\theta_1 + m_2(l_1 \cos\theta_1 + l_2 \cos\theta_2).$$

The kinetic energy $T = T_1 + T_2$ is the sum of the kinetic energy of two masses:

$$T_1 = \frac{1}{2} m_1 (\dot{x}_1^2 + \dot{y}_1^2) = \frac{1}{2} m_1 \left((-l_1 \dot{\theta}_1 \sin\theta_1)^2 + (l_1 \dot{\theta}_1 \cos\theta_1)^2 \right) = \frac{1}{2} m_1 l_1^2 \dot{\theta}_1^2$$

and similarly

$$\begin{aligned} T_2 = \frac{1}{2} m_2 &\left((-l_1 \dot{\theta}_1 \sin\theta_1 - l_2 \dot{\theta}_2 \sin\theta_2)^2 + (l_1 \dot{\theta}_1 \cos\theta_1 + l_2 \dot{\theta}_2 \cos\theta_2)^2 \right) \\ &= \frac{1}{2} m_2 \left(l_1^2 \dot{\theta}_1^2 + l_2^2 \dot{\theta}_2^2 + 2 l_1 l_2 \cos(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2 \right) \end{aligned}$$

Combining these expression, we find

$$T = \frac{1}{2} \dot{\theta}^T R(\theta_1, \theta_2) \dot{\theta}$$

where θ is a vector of generalized coordinates$\theta = (\theta_1, \theta_2)^T$, and

$$R(\theta_1, \theta_2) = \begin{pmatrix} (m_2 + m_1)\, l_1^2 & m_2 l_1 l_2 \cos(\theta_1 - \theta_2) \\ m_2 l_1 l_2 \cos(\theta_1 - \theta_2) & m_2 l_2^2 \end{pmatrix}$$

is the inertia matrix for the generalized coordinates. The Lagrangian is composed as

$$L = T_1 + T_2 - V$$

Now we immediately derive the equations (4.14) for the motion:

$$\begin{aligned} s_1 &= \frac{d}{dt}\left(m_1 l_1^2 \dot{\theta_1} + m_2 l_2^2 \dot{\theta_1} + 2 m_2 l_1 l_2 \dot{\theta_1}\cos(\theta_1 - \theta_2)\right) \\ &+ 2 m_2 l_1 l_2 \sin(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 + (m_1 l_1 + m_2 l_2)\sin\theta_1 = 0 \\ s_2 &= \frac{d}{dt}\left(m_2 l_2^2 \dot{\theta_2} + l_1 l_2 \cos(\theta_1 - \theta_2)\dot{\theta_1}\right) - 2 l_1 l_2 \sin(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 \\ &- m_2 l_2 \sin\theta_2 = 0 \end{aligned}$$

and notice that the whole energy $T_1 + T_2 + V$ is constant at all time.

The linearized equations of motion can be derived in an additional assumption $|q_i| \ll 1, |\dot{q}_i| \ll 1, i = 1, 2.$ they are

$$\begin{aligned} sl_1 &= \left(m_1 l_1^2 + m_2 l_2^2\right)\ddot{\theta_1} + l_1 l_2 \ddot{\theta_2} + (m_1 l_1 + m_2 l_2)\theta_1 = 0 \\ sl_2 &= \left(m_2 l_2^2 \ddot{\theta_2} + m_2 l_1 l_2 \ddot{\theta_1}\right) - m_2 l_2 \theta_2 = 0 \end{aligned}$$

or, in the vector form

$$M\ddot{\theta} + C\theta = 0$$

where

$$M = \begin{pmatrix} m_1 l_1^2 + m_2 l_2^2 & m_2 l_1 l_2 \\ m_2 l_1 l_2 & m_2 l_2^2 \end{pmatrix}, \quad C = \begin{pmatrix} m_1 l_1 + m_2 l_2 & 0 \\ 0 & m_2 l_2 \end{pmatrix}$$

Notice that the matrix M that plays the role of the masses show the inertial elements and it not diagonal but symmetric. The matrix C shows the stiffness of the system. The solution is given by the vector formula

$$\theta(t) = A_1 \exp(iBt) + A_2 \exp(-iBt), \quad B = (M^{-1}C)^{\frac{1}{2}}$$

Chapter 5

Constrained problems

We pass to consideration of extremal problem with additional constraints imposed on the minimizer. These constraints may prescribe the values of integrals of some function of minimizer as the isoperimetric problem does, or they may pose the restriction on the minimizer on each point of an admissible trajectory, as the geodesics problem required.

5.1 Constrained minimum in vector problems

5.1.1 Lagrange Multipliers method

Reminding of the technique discussed in calculus, we first consider a finite-dimensional problem of constrained minimum. Namely, we want to find the condition of the minimum:

$$J = \min_x f(x), \quad x \in R^n, \quad f \in C_2(R^n) \tag{5.1}$$

assuming that m constraints are applied

$$g_i(x_1, \dots x_n) = 0 \quad i = 1, \dots m, \quad m \le n, \tag{5.2}$$

The vector form of the constraints is

$$\boldsymbol{g}(\boldsymbol{x}) = 0$$

where $\boldsymbol{g}$ is a m-dimensional vector-function of an n-dimensional vector $\boldsymbol{x}$.

To find the minimum, we add the constraints with the Lagrange multipliers $\boldsymbol{\mu} = (\mu_1, \dots \mu_p)$ and end up with the problem

$$J = \min_{\boldsymbol{x}} \left[f(x) + \sum_i^m \mu_i g_i(x) \right]$$

The stationary conditions become:

$$\frac{\partial f}{\partial x_k} + \sum_i^m \mu_i \frac{\partial g_i}{\partial x_k} = 0, \quad k = 1, \ldots, n$$

or, in the vector form

$$\frac{\partial f}{\partial \boldsymbol{x}} + W \cdot \boldsymbol{\mu} = 0 \tag{5.3}$$

where the $m \times n$ Jacobian matrix W is

$$W = \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}} \quad \text{or, by elements, } W_{nm} = \frac{\partial g_n}{\partial x_m}$$

The system (5.3) together with the constraints (5.2) forms a system of $n + p$ equations for $n + p$ unknowns: Components of the vectors $\boldsymbol{x}$ and $\boldsymbol{\mu}$.

Example Consider the problem

$$J = \min_x \sum_i A_i^2 x_i \quad \text{subject to } \sum_i \frac{1}{x_i - k} = \frac{1}{c}.$$

Using Lagrange multiplier λ we rewrite it in the form:

$$J_a = \min_x \sum_i A_i^2 x_i + \lambda \left(\sum_i \frac{1}{x_i - k} - \frac{1}{c} \right).$$

From the condition $\frac{\partial J_a}{\partial \boldsymbol{x}} = 0$ we obtain

$$A_i^2 - \frac{\lambda}{(x_i - k)^2} = 0, \quad \text{or } \frac{1}{x_i - k} = \frac{|A_i|}{\sqrt{\lambda}} \quad i = 1, \ldots, n.$$

We substitute these values into expression for the constraint and obtain an equation for λ

$$\frac{1}{c} = \sum_i \frac{1}{x_i - k} = \frac{1}{\sqrt{\lambda}} \sum_i |A_i|$$

Solving this equation, we find λ, the minimizer x_i

$$\sqrt{\lambda} = c \sum_i |A_i|, \quad x_i = k + \frac{\sqrt{\lambda}}{|A_i|},$$

and the value of the minimizing function J:

$$J = k \sum_i A_i^2 + c \left(\sum_i |A_i| \right)^2$$

Observe, the minimum is a sum of squares of L_2 and L_1 norms of the vector $A = [A_1, \ldots, A_n]$.

How does it work? (Min-max approach) Consider again the finite-dimensional minimization problem

$$J = \min_{x_1, \ldots x_n} F(x_1, \ldots x_n) \tag{5.4}$$

subject to one constraint

$$g(x_1, \ldots x_n) = 0 \tag{5.5}$$

and assume that there exist solutions to (5.5) in the neighborhood of the minimal point.

It is easy to see that the described constrained problem is equivalent to the unconstrained problem

$$J_* = \min_{x_1, \ldots x_n} \max_{\lambda} \left(F(x_1, \ldots x_n) + \lambda g(x_1, \ldots x_n) \right) \tag{5.6}$$

Indeed, the inner maximization gives

$$\max_{\lambda} \lambda g(x_1, \ldots x_n) = \begin{cases} \infty & \text{if } g \neq 0 \\ 0 & \text{if } g = 0 \end{cases}$$

because λ can be made arbitrary large or arbitrary small. This possibility forces us to choose such $\boldsymbol{x}$ that delivers equality in (5.5), otherwise the cost of the problem (5.6) would be infinite (recall that x "wants" to minimize J_*). By assumption, such x exists. At the other hand, the constrained problem (5.4)-(5.5) does not change its cost J if zero $g = 0$ is added to it. Thereby $J = J_*$ and the problem (5.4) and (5.5) is equivalent to (5.6).

If we interchange the sequence of the two extremal operations in (5.6), we would arrive at the *augmented* problem J_D

$$J_D(\boldsymbol{x}, \lambda) = \max_{\lambda} \min_{x_1, \ldots x_n} \left(F(x_1, \ldots x_n) + \lambda g(x_1, \ldots x_n) \right) \tag{5.7}$$

The interchange of max and min- operations preserves the problems cost if $F(x_1, \ldots x_n) + \lambda g(x_1, \ldots x_n)$ is a convex function of $x_1, \ldots x_n$; in this case $J = J_D$. In a general case, we arrive at an inequality $J \leq J_D$ (see the min-max theorem in Sectionintro)

The extended Lagrangian J_* depends on $n+1$ variables. The stationary point corresponds to a solution to a system

$$\frac{\partial L}{\partial x_k} = \frac{\partial F}{\partial x_k} + \lambda \frac{\partial g}{\partial x_k} = 0, \quad k = 1, \ldots n, \tag{5.8}$$

$$\frac{\partial L}{\partial \lambda} = g = 0 \tag{5.9}$$

The procedure is easily generalized for several constrains. In this case, we add each constraint with its own Lagrange multiplier to the minimizing function and arrive at expression (5.3)

5.1.2 Exclusion of Lagrange multipliers and duality

We can exclude the multipliers $\boldsymbol{\mu}$ from the system (5.3) assuming that the constraints are independent, that is rank$(W) = m$. We project n-dimensional vector ∇F onto a $n-m$-dimensional subspace allowed by the constraints, and require that this projection is zero. The procedure is as follows.

1. Multiply (5.3) by W^T:

$$W^T \frac{\partial f}{\partial \boldsymbol{x}} + W^T W \cdot \boldsymbol{\mu} = 0, \tag{5.10}$$

   Since the constraints are independent, $p \times p$ matrix $W^T W$ is nonsingular, $\det(W^T W) \neq 0$.

2. Find m-dimensional vector of multipliers $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = -(W^T W)^{-1} W^T \frac{\partial f}{\partial \boldsymbol{x}},$$

3. Substitute the obtained expression for $\boldsymbol{\mu}$ into (5.3) and obtain:

$$(I - W(W^T W)^{-1} W^T) \frac{\partial f}{\partial \boldsymbol{x}} = 0 \tag{5.11}$$

Matrix $W(W^T W)^{-1} W^T$ is called the projector to the subspace W. Notice that the rank of the matrix $W(W^T W)^{-1} W^T$ is equal to p; it has p eigenvalues equal to one and $n-p$ eigenvalues equal to zero. Therefore the rank of $I - W(W^T W)^{-1} W^T$ is equal to $n-p$, and the system (5.11) produces $n-p$ independent optimality conditions. The remaining p conditions are given by the constraints (5.2): $g_i = 0,\ i = 1, \ldots p$. Together these two groups of relations produce n equations for n unknowns $x_1, \ldots, x_n$.

Below, we consider several special cases.

Degeneration: No constraints When there is no constraints, $W = 0$, the problem trivially reduces to the unconstrained on, and the necessary condition (5.11) becomes $\frac{\partial f}{\partial x} = 0$ holds.

Degeneration: $\boldsymbol{n}$ constraints Suppose that we assign n independent constraints. They themselves define vector $\boldsymbol{x}$ and no additional freedom to choose it is left. Let us see what happens with the formula (5.11) in this case. The rank of the matrix $W(W^T W)^{-1} W^T$ is equal to n, (W^{-1} exists) therefore this matrix-projector is equal to I:

$$W(W^T W)^{-1} W^T = I$$

and the equation (5.11) becomes a trivial identity. No new condition is produced by (5.11) in this case, as it should be. The set of admissible values of $\boldsymbol{x}$ shrinks to the point and it is completely defined by the n equations $\boldsymbol{g}(\boldsymbol{x}) = 0$.

One constraint Another special case occurs if only one constraint is imposed; in this case $p = 1$, the Lagrange multiplier $\boldsymbol{\mu}$ becomes a scalar, and the conditions (5.3) have the form:

$$\frac{\partial f}{\partial x_i} + \mu \frac{\partial g}{\partial x_i} = 0 \quad i = 1, \dots n$$

Solving for μ and excluding it, we obtain $n - 1$ stationary conditions:

$$\frac{\partial f}{\partial x_1} \left(\frac{\partial g}{\partial x_1} \right)^{-1} = \dots = \frac{\partial f}{\partial x_n} \left(\frac{\partial g}{\partial x_n} \right)^{-1} \tag{5.12}$$

Let us find how does this condition follow from the system (5.11). This time, W is a $1 \times n$ matrix, or a vector,

$$W = \left[\frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_n} \right]$$

We have:

$$\operatorname{rank} W(W^T W)^{-1} W^T = 1, \quad \operatorname{rank}(I - W(W^T W)^{-1} W^T) = n - 1$$

Matrix $I - W(W^T W)^{-1} W^T$ has $n - 1$ eigenvalues equal to one and one zero eigenvalue that corresponds to the eigenvector W. At the other hand, optimality condition (5.11) states that the vector

$$\nabla f = \left[\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right]$$

lies in the null-space of the matrix $I - W(W^T W)^{-1} W^T$ that is vectors $\frac{\partial f}{\partial \boldsymbol{x}}$ and W are parallel. Equation (5.12) expresses parallelism of these two vectors.

Quadratic function Consider minimization of a quadratic function

$$F = \frac{1}{2} \boldsymbol{x}^T A \boldsymbol{x} + d^T x$$

subject to linear constraints

$$Bx = \beta$$

where $A > 0$ is a positive definite $n \times n$ matrix, B is a $n \times m$ matrix of constraints, d and β are the n- and m-dimensional vectors, respectively. Here, $W = B$. The optimality conditions consist of m constraints $Bx = \beta$ and $n - m$ linear equations

$$(I - B(B^T B)^{-1} B^T)(Ax + d) = 0$$

Duality Let us return to the constraint problem

$$J_D = \min_x \max_\mu (F(x) + \mu^T g(x))$$

with the stationarity conditions,

$$\nabla F + \mu^T W(x) = 0$$

Instead of excluding μ as is was done before, now we do the opposite: Exclude n-dimensional vector x from n stationarity conditions, solving them for x and thus expressing x through μ: $x = \phi(\mu)$. When this expression is substituted into original problem, the later becomes

$$J_D = \max_\mu \{F(\phi(\mu)) + \mu^T g(\phi(\mu))\};$$

it is called *dual* problem to the original minimization problem.

Dual form for quadratic problem Consider again minimization of a quadratic. Let us find the dual form for it. We solve the stationarity conditions $Ax+d+B^T\mu$ for x, obtain

$$x = -A^{-1}(d + B^T\mu)$$

and substitute it into the extended problem:

$$J_D = \max_{\mu \in R_m} \left\{ \frac{1}{2}(d^T + \mu^T B)A^{-1}(d + B^T\mu) - \mu^T BA^{-1}(d + B^T\mu) - \mu^T\beta \right\}$$

Simplifying, we obtain

$$J_D = \max_{\mu \in R_m} \left\{ -\frac{1}{2}\mu^T BA^{-1}B^T\mu - \mu^T\beta + \frac{1}{2}d^T A^{-1}d \right\}$$

The dual problem is also a quadratic form over the m dimensional vector of Lagrange multipliers μ; observe that the right-hand-side term β in the constraints in the original problem moves to the sift term in the dual problem. The shift d in the original problem generates an additive term $\frac{1}{2}d^T A^{-1}d$ in the dual one.

5.1.3 Finite-dimensional variational problem revisited

Consider the optimization problem for a finite-difference system of equations

$$J = \min_{y_1,\ldots,y_N} \sum_i^N f_i(y_i, z_i)$$

where $f_1, \ldots, f_N$ are given value of a function f, $y_1, \ldots, y_N$ is the N-dimensional vector of unknowns, and z_i $i = 2, \ldots, N$ are the finite differences of y_i:

$$z_i = \text{Diff}(y_i) \quad \text{where } \text{Diff}(y_i) = \frac{1}{\Delta}(y_i - y_{i-1}), \quad i = 1, \ldots, N \tag{5.13}$$

Assume that the boundary values y_1 and y_n are given and take (5.13) as constraints. Using Lagrange multiplies $\mu_1, \ldots, \mu_N$ we pass to the augmented function

$$J_a = \min_{y_1,\ldots,y_N;\ z_1,\ldots,z_N} \sum_i^N \left[f_i(y_i, z_i) + \mu_i \left(z_i - \frac{1}{\Delta}(y_i - y_{i-1}) \right) \right]$$

The necessary conditions are:

$$\frac{\partial J_a}{\partial y_i} = \frac{\partial f_i}{\partial y_i} + \frac{1}{\Delta}(-\mu_i + \mu_{i+1}) = 0 \quad 2 = 1, \ldots, N-1$$

and

$$\frac{\partial J_a}{\partial z_i} = \frac{\partial f_i}{\partial z_i} + \mu_i = 0 \quad i = 2, \ldots, N-1$$

Excluding μ_i from the last equation and substituting their values into the previous one, we obtain the conditions:

$$\frac{\partial J_a}{\partial y_i} = \frac{\partial f_i}{\partial y_i} + \frac{1}{\Delta} \left(\frac{\partial f_i}{\partial z_i} - \frac{\partial f_{i+1}}{\partial z_{i+1}} \right) = 0 \quad i = 2, \ldots, N-1$$

or, recalling the definition of the Diff -operator,

$$\text{Diff} \left(\frac{\partial f_{i+1}}{\partial z_{i+1}} \right) - \frac{\partial f_i}{\partial y_i} = 0 \quad z_i = \text{Diff}(y_i) \tag{5.14}$$

One can see that the obtained necessary conditions have the form of the difference equation of second-order.

On the other hand, Diff-operator is an approximation of a derivative and the equation (5.14) is a finite-difference approximation of the Euler equation.

5.1.4 Inequality constraints

Nonnegative Lagrange multipliers Consider the problem with a constraint in the form of inequality:

$$\min_{x_1,\ldots x_n} F(x_1, \ldots x_n) \quad \text{subject to } g(x_1, \ldots x_n) \le 0 \tag{5.15}$$

In order to apply the Lagrangian multipliers technique, we reformulate the constraint:

$$g(x_1, \ldots x_n) + v^2 = 0$$

where v is a new auxiliary variable.

The augmented Lagrangian becomes

$$L_*(x, v, \lambda) = f(x) + \lambda g(x) + \lambda v^2$$

and the optimality conditions with respect to v are

$$\frac{\partial L_*}{\partial v} = 2\lambda v = 0 \tag{5.16}$$

$$\frac{\partial^2 L_*}{\partial v^2} = 2\lambda \ge 0 \tag{5.17}$$

The second condition requires the nonnegativity of the Lagrange multiplier and the first one states that the multiplier is zero, $\lambda = 0$, if the constraint is satisfied by a strong inequality, $g(x_0) > 0$.

The stationary conditions with respect to x

$$\begin{array}{ll} \nabla f = 0 & \text{if } g \leq 0 \\ \nabla f + \lambda \nabla g = 0 & \text{if } g = 0 \end{array}$$

state that either the minimum correspond to an inactive constraint ($g > 0$) and coincide with the minimum in the corresponding unconstrained problem, or the constraint is active ($g = 0$) and the gradients of f and g are parallel and directed in opposite directions:

$$\frac{\nabla f(x_b) \cdot \nabla g(x_b)}{|\nabla f(x_b)|\,|\nabla g(x_b)|} = -1, \quad x_b : \; g(x_b) = 0$$

In other terms, the projection of $\nabla f(x_b)$ on the subspace orthogonal to $\nabla g(x_b)$ is zero, and the projection of $\nabla f(x)$ on the direction of $\nabla g(x_b)$ is nonpositive.

The necessary conditions can be expressed by a single formula using the notion of infinitesimal variation of x or a differential. Let x_0 be an optimal point, x_{trial} – an admissible (consistent with the constraint) point in an infinitesimal neighborhood of x_0, and $\delta x = x_{\text{trial}} - x_0$. Then the optimality condition becomes

$$\nabla f(x_0) \cdot \delta x \geq 0 \quad \forall \delta x \tag{5.18}$$

Indeed, in the interior point x_0 ($g(x_0) > 0$) the vector δx is arbitrary, and the condition (5.18) becomes $\nabla f(x_0) = 0$. In a boundary point x_0 ($g(x_0) = 0$), the admissible points are satisfy the inequality $\nabla g(x_0) \cdot \delta x \leq 0$, the condition (5.18) follows from (13.21).

It is easy to see that the described constrained problem is equivalent to the unconstrained problem

$$L_* = \min_{x_1, \ldots x_n} \max_{\lambda > 0} \left(F(x_1, \ldots x_n) + \lambda g(x_1, \ldots x_n) \right) \tag{5.19}$$

that differs from (5.7) by the requirement $\lambda > 0$.

Several constraints: Kuhn-Tucker conditions Several inequality constraints are treated similarly. Assume the constraints in the form

$$g_1(x) \leq 0, \ldots, g_m(x) \leq 0.$$

The stationarity condition can be expressed through nonnegative Lagrange multipliers

$$\nabla f + \sum_{i=1}^{m} \lambda_i \nabla g_i = 0, \tag{5.20}$$

where

$$\lambda_i \geq 0, \quad \lambda_i g_i = 0, \quad i = 1, \ldots, m. \tag{5.21}$$

The minimal point corresponds either to an inner point of the permissible set (all constraints are inactive, $g_i(x_0) < 0$), in which case all Lagrange multipliers λ_i are zero, or to a boundary point where $p \leq m$ constraints are active. Assume for definiteness that the first p constraints are active, that is

$$g_1(x_0) = 0, \quad \ldots, \quad g_p(x_0) = 0. \tag{5.22}$$

The conditions (5.21) show that the multiplier λ_i is zero if the ith constrain is inactive, $g_i(x) > 0$. Only active constraints enter the sum in (5.23), and it becomes

$$\nabla f + \sum_{i=1}^{p} \lambda_i \nabla g_i = 0, \quad \lambda_i > 0, \quad i = 1, \ldots, p. \tag{5.23}$$

The term $\sum_{i=1}^{p} \lambda_i \nabla g_i(x_0)$ is a cone with the vertex at x_0 stretched on the rays $\nabla g_i(x_0) > 0$, $i = 1, \ldots, p$. The condition (5.23) requires that the negative of $\nabla f(x_0)$ belongs to that cone.

Alternatively, the optimality condition can be expressed through the admissible vector δx,

$$\nabla f(x_0) \cdot \delta x \geq 0 \tag{5.24}$$

Assume again that the first p constraints are active, as in (**??**)

$$g_1(x_0) =, \ldots, = g_p(x_0) = 0.$$

In this case, the minimum is given by (5.24) and the admissible directions of δx satisfy the system of linear inequalities

$$\delta x \cdot \nabla g_i \geq 0, \quad i = 1, \ldots, p. \tag{5.25}$$

Assume that

These conditions are called Kuhn-Tucker conditions, see []

5.2 Isoperimetric problem

5.2.1 Stationarity conditions

Isoperimetric problem of the calculus of variations is

$$\min_u \int_a^b F(x, u, u')dx \quad \text{subject to} \quad \int_a^b G(x, u, u')dx = 0 \tag{5.26}$$

Applying the same procedure as in the finite-dimensional problem, we reformulate the problem using Lagrange multiplier λ:

$$\min_u \int_a^b \left[F(x, u, u') + \lambda G(x, u, u')\right] dx \tag{5.27}$$

To justify the approach, we may look on the finite-dimensional analog of the problem

$$\min_{u_i} \sum_{i=1}^{N} F_i(u_i, \text{Diff}(u_i)) \quad \text{subject to} \quad \sum_{i=1}^{N} G_i(u_i, \text{Diff}(u_i)) = 0$$

The Lagrange method is applicable to the last problem which becomes

$$\min_{u_i} \sum_{i=1}^{N} \left[F_i(u_i, \text{Diff}(u_i)) + \lambda G_i(u_i, \text{Diff}(u_i))\right].$$

Passing to the limit when $N \to \infty$ we arrive at (5.27).

The procedure of solution is as follows: First, we solve Euler equation for the problem(5.27)

$$\frac{d}{dx}\frac{\partial}{\partial u'}(F + \lambda G) - \frac{\partial}{\partial u}(F + \lambda G) = 0.$$

Keeping λ undefined and arrive at minimizer $u(x, \lambda)$ which depends on parameter λ. The equation

$$\int_a^b G(x, u(x, \lambda), u'(x, \lambda))dx = 0$$

defines this parameter.

Remark 5.2.1 The method assumes that the constraint is consistent with the variation: The variation must be performed upon a class of functions u that satisfy the constraint. Parameter λ has the meaning of the cost for violation of the constraint.

Of course, it is assumed that the constraint can be satisfied for all varied functions that are close to the optimal one. For example, the method is not applicable to the constraint

$$\int_a^b u^2 dx \leq 0$$

because this constraint allows for only one function $u = 0$ and will be violated at any varied trajectory.

5.2.2 Dido problem revisited

Let us apply the variational technique to Dido Problem discussed in Chapter **??**. It is required to maximize the area A between the OX axes and a positive curve $u(x)$

$$A = \int_a^b u dx \quad u(x) \geq 0 \forall x \in [a, b]$$

assuming that the length L of the curve is given

$$L = \int_a^b \sqrt{1 + u'^2} dx$$

and that the beginning and the end of the curve belong to OX-axes: $u(a) = 0$ and $u(b) = 0$. Without lose of generality we assume that $a = 0$ and we have to find b.

The constrained problem has the form

$$J = A + \lambda L = \int_0^b \left(u + \lambda \sqrt{1 + u'^2} \right) dx$$

where λ is the Lagrange multiplier.

The Euler equation for the extended Lagrangian is

$$1 - \lambda \frac{d}{dx} \left(\frac{u'}{\sqrt{1 + u'^2}} \right)$$

Let us fix λ and find u as a function of x and λ. Integrating, we obtain

$$\lambda \frac{u'}{\sqrt{1 + u'^2}} = x - C_1$$

where C_1 is a constant of integration. Solving for $u' = \frac{du}{dx}$, we rewrite the last equation as

$$du = \pm \frac{(x - C_1) dx}{\sqrt{\lambda^2 + (x - C_1)^2}},$$

integrate it:

$$u = \mp \sqrt{\lambda^2 + (x - C_1)^2} + C_2$$

and rewrite the result as

$$(x - C_1)^2 + (u - C_2)^2 = \lambda^2 \tag{5.28}$$

The extremal is a part of the circle. The constants C_1, C_2 and λ can be found from boundary conditions and the constraints.

To find the length b of the trajectory, we use the transversality condition (**??**):

$$u' \frac{\partial F}{\partial u'} - F = - \frac{\lambda}{\sqrt{1 + u'^2}} - u = 0$$

which gives $|u'(b)| = \infty$ – the optimal trajectory approaches OX-axis perpendicular to it. By symmetry, $|u'(a)| = \infty$, and the optimal trajectory is the semicircle of the radius λ, symmetric with respect to OX-axis. We find $\lambda = \frac{L}{\pi}$, $C_1 = a + \frac{L}{2\pi}$, and $C_2 = 0$.

5.2.3 Catenoid

The classical problem of the shape of a heavy chain (catenoid, from Latin *catena*) was considered by Euler ?? using a *variational principle.* It is postulated, that the equilibrium minimizes minimizes the potential energy W of the chain

$$W = \int_0^1 g\rho u\, ds = g\rho \int_0^1 u\sqrt{1+(u')^2}dx$$

defined as the limit of the sum of vertical coordinates of the parts of the chain. Here, ρ is the density of the mass of the chain, ds is the element of its length, x and u are the horizontal and vertical coordinates, respectively. The length of the chain

$$L = \int_0^1 \sqrt{1+(u')^2}dx$$

and the coordinates of the ends are fixed. Normalizing, we put $g\rho = 1$. Formally, the problem becomes

$$I = \min_{u(x)} \ (W(u) + \lambda L(u)), \quad W(u) + \lambda L(u) = \int_0^1 (u+\lambda)\sqrt{1+(u')^2}dx$$

The Lagrangian is independent of x and therefore permits the first integral

$$(u+\lambda)\left(\frac{(u')^2}{\sqrt{1+(u')^2}} - \sqrt{1+(u')^2}\right) = C$$

that is simplified to

$$\frac{u+\lambda}{\sqrt{1+(u')^2}} = C.$$

We solve for u'

$$\frac{du}{dx} = \sqrt{\left(\frac{u+\lambda}{C}\right)^2 - 1}$$

integrate

$$x = \ln\left(\lambda + u + \sqrt{\left(\frac{u+\lambda}{C}\right)^2 - 1}\right) - \ln C + x_0$$

and find the extremal $u(x)$

$$u = -C\cosh\left(\frac{x-x_0}{C}\right) + \lambda$$

The equation – the catenoid – defines the shape of a chain; it also gave the name to the hyperbolic cosine.

5.2.4 General form of a variational functional

Lagrange method allows for reformulation of an extremal problem in a general form as a simplest variational problem. The minimizing functional can be the product, ratio, superposition of other differentiable function of integrals of the minimizer and its derivative. Consider the problem

$$J = \min_u \Phi(I_1, \dots, I_n) \tag{5.29}$$

where

$$I_k(u) = \int_a^b F_k(x, u, u')dx \quad k = 1, \dots n \tag{5.30}$$

and Φ is a continuously differentiable function. Using Lagrange multipliers λ_1, λ_n, we transform the problem (5.29) to the form

$$J = \min_u \min_{I_1,\dots,I_n} \max_{\lambda_1,\dots\lambda_n} \left\{ \Phi + \sum_{k=1}^n \lambda_k \left(I_k - \int_a^b F_k(x, u, u')dx \right) \right\}. \tag{5.31}$$

The stationarity conditions for (5.31) consist of n algebraic equations

$$\frac{\partial \Phi}{\partial I_k} + \lambda_i = 0 \tag{5.32}$$

and the differential equation – the Euler equation

$$S(\Psi, u) = 0$$

$$\left(\text{recall that } S(\Psi, u) = \frac{d}{dx}\frac{\partial \Psi}{\partial u'} - \frac{\partial \Psi}{\partial u} \right)$$

for the function

$$\Psi(u) = \sum_{k=1}^n \lambda_k F_k(x, u, u')$$

Together with the definitions (5.30) of I_k, this system enables us to determine the real parameters I_k and λ_k and the function $u(x)$. The Lagrange multipliers can be excluded from the previous expression using (5.32), then the remaining stationary condition becomes an integro-differential equation

$$S(\bar{\Psi}, u) = 0, \quad \bar{\Psi}(I_k, u) = \sum_{k=1}^n \frac{\partial \Phi}{\partial I_k} F_k(x, u, u') \tag{5.33}$$

Next examples illustrate the approach.

The product of integrals

Consider the problem

$$\min_u J(u), \quad J(u) = \left(\int_a^b \phi(x,u,u')dx\right)\left(\int_a^b \psi(x,u,u')dx\right).$$

We rewrite the minimizing quantity as

$$J(u) = I_1(u)I_2(u), \quad I_1(u) = \int_a^b \phi(x,u,u')dx, \quad I_2(u) = \int_a^b \psi(x,u,u')dx,$$

apply stationary condition (5.33), and obtain the condition

$$I_1\delta I_2 + I_2\delta I_1 = I_2(u)S(\phi(u),u) + I_1(u)S(\psi(u),u) = 0.$$

or

$$\left(\int_a^b \phi(x,u,u')dx\right)^{-1} S(\phi(u),u) + \left(\int_a^b \psi(x,u,u')dx\right)^{-1} S(\psi(u),u) = 0$$

The equation is nonlocal: Solution u at each point depends on its first and second derivatives and integrals of $\phi(x,u,u')$ and $\phi(x,u,u')$ over the whole interval $[a,b]$.

Example 5.2.1 Solve the problem

$$\min_u \left(\int_0^1 (u')^2dx\right)\left(\int_0^1 (u+1)dx\right) \quad u(0)=0, \quad u(1)=a$$

We compute the Euler equation

$$u'' - R = 0, \quad u(0)=0, \quad u(1)=a.$$

where

$$R = \frac{I_1}{I_2} = 0, \quad I_1 = \int_0^1 (u')^2dx, \quad I_2 = \int_0^1 (u-1)dx$$

The integration gives

$$u(x) = \frac{1}{2}Rx^2 + \left(a - \frac{1}{2}R\right)x,$$

We obtain the solution that depends on R – the ratio of the integrals of two function of this solution. To find R, we substitute the expression for $u = u(R)$ into right-hand sides of I_1 and I_2,

$$I_1 = \frac{R^2}{12} + a^2, \quad I_2 = -\frac{R}{12} + \frac{1}{2}a + 1$$

compute the ratio, $\frac{I_1}{I_2} = R$ and obtain the equation for R,

$$R = \frac{R^2 + 12a^2}{R + 6a + 12}$$

Solving it, we find $R = \frac{1}{2}(3a + 6 \pm \sqrt{36 + 36a - 15a^2})$.

At this point, we do not know whether the solution correspond to minimum or maximum. This question is investigated later in Chapter 6.

The ratio of integrals

Consider the problem

$$\min_u J(u), \quad J(u) = \frac{\int_a^b \phi(x, u, u')dx}{\int_a^b \psi(x, u, u')dx}.$$

We rewrite it as

$$J = \frac{I_1}{I_2}, \quad I_1(u) = \int_a^b \phi(x, u, u')dx, \quad I_2(u) = \int_a^b \psi(x, u, u')dx, \tag{5.34}$$

apply stationary condition (5.33), and obtain the condition

$$\frac{1}{I_2(u)} S(\phi(u), u) - \frac{I_1(u)}{I_2^2(u)} S(\psi(u), u) = 0.$$

Using definition (5.34) of the goal functional, we bring the previous expression to the form

$$S(\phi(x, u, u') - J\psi(x, u, u'), u) = 0$$

Observe that the stationarity condition depends on the cost J of the problem. The examples will be given in the next section.

Superposition of integrals

Consider the problem

$$\min_u \int_a^b R\left(x, u, u', \int_a^b \phi(x, u, u')dx\right) dx$$

We introduce a new variable I

$$I = \int_a^b \phi(x, u, u')dx$$

and reformulate the problem as

$$\min_u \int_a^b \left[R(x, u, u', I) + \lambda \left(\phi(x, u, u') - \frac{I}{b-a}\right)\right] dx$$

where λ is the Lagrange multiplier. The stationarity conditions are:

$$S((R+\lambda\phi),u)=0, \quad \frac{\partial R}{\partial I}-\frac{1}{b-a}=0.$$

and the above definition of I.

The general procedure is similar: We always can rewrite a minimization problem in the standard form adding new variables (as the parameter c in the previous examples) and corresponding Lagrange multipliers.

Inequality in the isoperimetric condition Often, the isoperimetric constraint is given in the form of an inequality

$$\min_u \int_a^b F(x,u,u')dx \quad \text{subject to} \quad \int_a^b G(x,u,u')dx \geq 0 \tag{5.35}$$

In this case, the additional condition $\lambda \geq 0$ is added to the Euler-Lagrange equations (13.21) according to the (13.21).

Remark 5.2.2 Sometimes, the replacement of an equality constraint with the corresponding inequality can help to determine the sign of the Lagrange multiplier. For example, consider the Dido problem, and replace the condition that the perimeter is fixed with the condition that the perimeter is smaller than or equal to a constant. Obviously, the maximal area corresponds to the maximal allowed perimeter and the constraint is always active. On the other hand, the problem with the inequality constraint requires positivity of the Lagrange multiplier; so we conclude that the multiplier is positive in both the modified and original problem.

5.2.5 Homogeneous functionals and Eigenvalue Problem

The next two problems are *homogeneous*: The functionals do not vary if the solution is multiplied by any number. Therefore, the solution is defined up to a constant multiplier.

The eigenvalue problem correspond to the functional

$$\min_u \frac{\int_0^1 (u')^2 dx}{\int_0^1 u^2 dx} \quad x(0)=x(1)=0$$

it can be compared with the problem:

$$\min_u \frac{\int_0^1 (u')^2 dx}{\left(\int_0^1 u dx\right)^2} \quad x(0)=x(1)=0$$

Do these problem have nonzero solutions?

5.3 Constraints in boundary conditions

Constraints on the boundary, fixed interval Consider a variational problem (in standard notations) for a vector minimizer $\boldsymbol{u}$. If there are no constrains imposed on the end of the trajectory, the solution to the problem satisfies n natural boundary conditions

$$\delta \boldsymbol{u}(b) \cdot \left.\frac{\partial F}{\partial \boldsymbol{u}'}\right|_{x=b} = 0$$

(For definiteness, we consider here conditions on the right end, the others are clearly identical).

The vector minimizer of a variational problem may have some additional constraints posed at the end point of the optimal trajectory. Denote the boundary value of $u_i(b)$ by v_i The constraints are

$$\phi_i(v_1, \dots v_n) = 0 \quad i = 1, \dots, k; \quad k \le n$$

or in vector form,

$$\Phi(x, \boldsymbol{v}) = 0,$$

where Φ is the corresponding vector function. The minimizer satisfies these conditions and $n - k$ supplementary natural conditions that arrive from the minimization requirement. Here we derive these supplementary boundary conditions for the minimizer.

Let us add the constrains with a vector Lagrange multiplier $\boldsymbol{\lambda} = (\lambda_1, \dots . \lambda_k)$ to the problem. The variation of $\boldsymbol{v} = \boldsymbol{u}(b)$ gives the conditions

$$\delta \boldsymbol{v} \cdot \left[\left.\frac{\partial F}{\partial \boldsymbol{u}'}\right|_{x=b, \boldsymbol{u}=\boldsymbol{v}} + \frac{\partial \Phi}{\partial \boldsymbol{v}} \boldsymbol{\lambda} \right] = 0$$

The vector in the square brackets must be zero because of arbitrariness of $\nu = \delta \boldsymbol{u}(b)$.

Next, we may exclude $\boldsymbol{\lambda}$ from the last equation (see the previous section 5.1.2):

$$\boldsymbol{\lambda} = - \left[\left(\frac{\partial \Phi}{\partial \boldsymbol{u}} \right)^T \frac{\partial \Phi}{\partial \boldsymbol{u}} \right]^{-1} \left.\frac{\partial F}{\partial \boldsymbol{u}'}\right|_{x=b, \boldsymbol{u}=\boldsymbol{v}} \tag{5.36}$$

and obtain the conditions

$$\left(I - \frac{\partial \Phi}{\partial \boldsymbol{u}}^T \left[\left(\frac{\partial \Phi}{\partial \boldsymbol{u}} \right)^T \frac{\partial \Phi}{\partial \boldsymbol{u}} \right]^{-1} \frac{\partial \Phi}{\partial \boldsymbol{u}} \right) \left.\frac{\partial F}{\partial \boldsymbol{u}'}\right|_{x=b, \boldsymbol{u}=\boldsymbol{v}} = 0 \tag{5.37}$$

The rank of the matrix in the parenthesis is equal to $n - k$. Together with k constrains, these conditions are the natural conditions for the variational problem.

Example

$$\min_{u_1,u_2} \int_a^b (u_1'^2 + u_2'^2 + u_3')dx, \quad u_1(b) + u_2(b) = 1, \ u_1(b) - u_3(b) = 1,$$

We compute

$$\frac{\partial F}{\partial \boldsymbol{u}'} = \begin{pmatrix} 2u_1 \\ 2u_2 \\ 1 \end{pmatrix}, \quad \frac{\partial \Phi}{\partial \boldsymbol{u}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & -1 \end{pmatrix},$$

(please continue..)

Free boundary with constraints Consider a general case when the constraints $\Phi(x,u) = 0$ are posed on the solution at the end point. Variation of these constrains results in the condition:

$$\delta\Phi(x,u)|_{x=b} = \frac{\partial \Phi}{\partial u}\delta u + \left(\frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial u}u'\right)\delta x$$

Adding the constraints to the problem with Lagrange multiplier $\boldsymbol{\lambda}$, performing variation, and collecting terms proportional to δx, we obtain the condition at the unknown end point $x = b$

$$F(x,u,u') - \frac{\partial F}{\partial u'}u' + \boldsymbol{\lambda}^T \left(\frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial u}u'\right) = 0$$

where $\boldsymbol{\lambda}$ is defined in (5.36). Together with $n-k$ conditions (5.37) and k constraints, they provide $n+1$ equations for the unknowns $u_1(b), \ldots, u_n(b), b$.

5.4 Pointwise Constraints

5.4.1 Stationarity conditions

Consider a variational problem for a vector-valued minimizer $u = u_1, \ldots u_n$.

$$\min_u \int_a^b F(x, \boldsymbol{u}, \boldsymbol{u}')dx$$

Assume that the minimizer obeys certain constraint(algebraic or differential) in each point of any admissible trajectory,

$$G(x, \boldsymbol{u}, \boldsymbol{u}') = 0, \forall x \in (a,b) \tag{5.38}$$

The number of constraints is less than the number of minimizers. This way, we arrive at the constrained variational problem

$$\min_u \int_a^b F(x, \boldsymbol{u}, \boldsymbol{u}')dx \quad \text{subject to } G(x, \boldsymbol{u}, \boldsymbol{u}') = 0, \forall x \in (a,b) \tag{5.39}$$

As in the isoperimetric problem, we use the Lagrange multipliers method to account for the constrain. This time, the constraint must be enforced in every point of the trajectory, therefore the Lagrange multiplier becomes a function of x. To prove the method, it is enough to pass to finite-dimensional problem; after discretization, the constraint is replaced by the array of equations

$$G(x,u,u')=0 \ \Rightarrow \ G_i(u_i, \mathrm{Diff}\,(u_i))=0, \quad i=1,\dots N.$$

Each of this constrains, multiplied by its own Lagrange multiplier $\mu_1,\dots\mu_N$, must be added to the functional. The array of these multipliers converge to a function $\mu(x)$ when $N\to\infty$. The variational problem becomes

$$\min_u \int_a^b [F(x,u,u')+\mu(x)G(x,u,u')]\,dx \tag{5.40}$$

The necessary conditions are expressed in the form of differential constraints (5.38) and Euler equation:

$$G(x,u,u')=0 \tag{5.41}$$

$$\left(\frac{d}{dx}\frac{\partial}{\partial u'}+\frac{\partial}{\partial u}\right)(F+\mu G)=0. \tag{5.42}$$

They define functions $u(x)$ and $\mu(x)$.

Algebraic constraints Notice that if the constraints are algebraic, $G=G(x,u)$, then (5.42) does not depend on μ' and is an algebraic relation for μ.

Consider the case of one constraint $G(x,u)=0$ The Euler equations are

$$\frac{\partial F}{\partial u_k}-\frac{d}{d\,x}\frac{\partial F}{\partial u_k'}+\mu\frac{\partial G}{\partial u_k}=0, \quad k=1\dots,n.$$

We may exclude $\mu=\mu(x)$ from the system and obtain $n-1$ equations

$$\left(\frac{\partial F}{\partial u_1}-\frac{d}{d\,x}\frac{\partial F}{\partial u_1'}\right)\left(\frac{\partial G}{\partial u_1}\right)^{-1}=\left(\frac{\partial F}{\partial u_k}-\frac{d}{d\,x}\frac{\partial F}{\partial u_k'}\right)\left(\frac{\partial G}{\partial u_k}\right)^{-1}, \quad k=2,\dots,n$$

for $u_1,\dots u_n$; this system is supplemented with the constraint $G(x,u)=0$.

The general case is considered similarly. Euler equation forms a linear system for vector-function μ; it can be excluded from the system, as it will be shown in following examples.

Example 5.4.1 (Euler equation revisited) As a first example, we derive Euler equation in a different manner: The minimization problem

$$\min_u \int_a^b F(x,u,v)dx \quad \text{subject to } v=u' \tag{5.43}$$

is obviously equivalent to the canonic variational problem.

Using Lagrange multiplier, we rewrite the problem as

$$\min_u \int_a^b \left(F(x,u,v) + \mu(u' - v)\right) dx$$

Variation with respect to u, v gives, respectively,

$$\mu' - \frac{\partial F}{\partial u} = 0, \quad \mu + \frac{\partial F}{\partial v} = 0$$

(the term $\mu u'$ is transformed by integration by parts). We exclude μ by differentiation of the second equation and subtraction of the first one:

$$\left(\frac{\partial F}{\partial v}\right)' - \frac{\partial F}{\partial u} = 0$$

Accounting for the constraint $v = u'$ we arrive at Euler equation.

Geodesics as constrained problem We return to the problem of geodesics – the shortest path on a surface between two points at this surface. Here we will develop a general approach to the problem without assumptions that the local coordinates and the metric is introduced on the surface. We simply assume that the surface is parameterized as

$$\psi(x_1, x_2, x_3) = 0 \tag{5.44}$$

where x_1, x_2, x_3 are Cartesian coordinates. The distance D along a path $x(t), y(t), z(t)$ is

$$D = \int_0^1 \sqrt{x_1'(t)^2 + x_2'(t)^2 + x_3'(t)^2} dt$$

The extended Lagrangian is

$$F = \sqrt{x_1'(t)^2 + x_2'(t)^2 + x_3'(t)^2} + \mu(t)\psi(x_1, x_2, x_3) = 0$$

where $\mu(t0$ is the Lagrange multiplier. Euler equations are

$$\frac{d}{dt}\frac{x_i'}{\sqrt{x_i'(t)^2 + x_2'(t)^2 + x_3'(t)^2}} - \mu\frac{\partial \psi}{\partial x_i} = 0, \quad i = 1, 2, 3.$$

Excluding μ, we obtain two equalities

$$\left(\frac{\partial \psi}{\partial x_i}\right)^{-1} \frac{d}{dt}\left(\frac{x_i'}{\sqrt{x_i'(t)^2 + x_2'(t)^2 + x_3'(t)^2}}\right) = \mu(t) \quad i = 1, 2, 3$$

which together with equation (5.44) for the surface determine the optimal trajectory: the geodesic.

5.4.2 Constraints in the form of differential equations

The same idea of constrained variational problem can be used to take into account the differential equations of the motion as constraints

$$g(u, u') = 0 \forall x \in [0, 1]. \tag{5.45}$$

This idea is fully exploited in the control theory, (see below, Section 10.1). The formal scheme is as in the previous case, but this time the derivatives of the Lagrange multipliers participate in the Euler equation:

$$\frac{dd}{dx}\left(\frac{\partial F}{\partial u'} + \lambda \frac{\partial g}{\partial u'}\right) - \frac{\partial F}{\partial u} - \frac{\partial g}{\partial u'}$$

that should be solved together with (5.45) to determine u and λ. Here we illustrate it by an example.

Sailing boat Consider the problem: How to use the minimal resources to sail to a proper destination. First, let us do the modelling. The equations of the boat in the water are

$$m\ddot{x} + \gamma \dot{x} = f(t)$$

where x is the coordinate of the boat, m is its mass, γ is the dissipation, and $f(t)$ is the time-dependent driving force that depends on the used amount of fuel

$$|f| = r^q.$$

It is required to bring the boat to the moorage $x(T) = P$ from the moorage $x(0) = 0$ in the given time T; the speed in the beginning and in the end must be zero, $\dot{x}(0) = \dot{x}(T) = 0$.

The total amount R of the fuel

$$R = \int_0^T r(t) dt \tag{5.46}$$

must be minimized:

Remark 5.4.1 In the modelling, it was assumed that the boat is moving straight from the start to the destination and the forward and backward acceleration require the same amount of fuel.

We formulate the variational problem for the unknown fuel consumption rate $r(t)$ and the boat's speed $v(t) = \dot{r}$ subject to differentia constraint

$$m\dot{v} + \gamma v = r^q \tag{5.47}$$

boundary conditions

$$v(0) = v(T) = 0$$

and the integral constraint

$$\int_0^T v(t) = L. \tag{5.48}$$

Accounting for the constrains (5.47) and (5.48) by Lagrange multiplier $\lambda = \lambda(x)$ and $\mu =$ constant, we obtain the variational problem

$$\min_{x(x),r(x)} \int_a^b F(r, v, \lambda, \mu)dx, \quad v(0) = 0$$

with the Lagrangian

$$F = r + \lambda(m\dot{v} + \gamma v - r^q) + \mu v$$

The Euler equations are respectively (from the variation with respect to v and r)

$$\begin{array}{ll} \delta v: & -m\dot{\lambda} + \gamma\lambda + \mu = 0 \\ \delta r: & 1 + qr^{q-1}\lambda = 0 \end{array}$$

Solving this system, we find

$$\lambda = -\frac{\mu}{\gamma} + C\exp\left(\frac{\gamma t}{m}\right), \quad , r(t) = \frac{1}{q}\lambda^{\frac{1}{q-1}} = \frac{1}{q}\left[-\frac{\mu}{\gamma} + C\exp\left(\frac{\gamma t}{m}\right),\right]^{\frac{1}{q-1}}$$

where μ and C are still undefined constants. Those are found evaluating $v(t)$

$$v(t) = -\exp\left(-\frac{\gamma t}{m}\right)\int_0^t \frac{r(t)^q}{m}\exp\left(\frac{\gamma t}{m}\right)dt$$

and applying the integral constraint (5.48) and boundary condition $v(T) = 0$.

5.4.3 Notion of variational inequalities

The variational problem with pointwise constraints in the form of inequalities, called variational inequalities, were investigated only recently, see [**?**]. These problems are formulated as a variational problem plus an inequality.

$$\min_{u(x)\geq\phi(x)} \int_a^b F(x, u, u')dx \tag{5.49}$$

The increment of the objective functional $I(u + \delta u) - I(u)$ is nonnegative at the optimal trajectory

$$I(u + \delta u) - I(u) = -\int_a^b S(F, u)\delta u\, dx \geq 0.$$

Here, (F, u) is the Euler equation (13.21). Let us analyze this formula.

When the constraint is satisfied as strict inequality, $u > \phi(x)$, an infinitesimal variation variation δu is not constrained and the minimizer u satisfies the Euler equation $S(F, u) = 0$ to keep the increment nonnegative. Otherwise, the extremal is coincide with the constraint, $u = \phi(x)$, variation δu of the trajectory but must be positive $\delta u \geq 0$ because all admissible trajectories $u + \delta u$ are above the constraint $\phi(x)$, $u(x) + \delta u(x) \geq \phi(x)$ for all x. Correspondingly, the variation $I(u + \delta u) - I(u)$ is nonnegative if the inequality holds

$$(F, u)|_{u(x)=\phi(x)} \leq 0.$$

To sum up, we formulate the obtained optimality conditions. The optimal trajectory satisfies one of the two supplementary conditions:
Either

$$S(F, u) = 0 \quad \text{ and } u(x) \geq \phi(x)$$

or

$$u(x) = \phi(x) \quad \text{ and } S(F, u) \leq 0$$

The equalities define the minimizer in each regime, and the inequalities check the optimality of the regime.

String (membrane) over an obstacle

Consider again the problem of catenoid, assuming in addition that the chain is hanged over a plane surface and is cannot go beneath it. The variational problem is

$$\min_{u(x) \geq 0, u(a)=A, u(b)=B} \int_a^b () dx$$

and its solution is

$$\begin{array}{l} u(x) = a \quad u'' \geq 0 \\ u''(x) = q \;\; u > a \end{array}$$

Convex envelope

Consider the problem about the shortest path around an obstacle discussed in Chapter **??** in geometric terms. Now we formulate the problem as a variational inequality. We find a curve $u(x) \geq 0$ that has the shortest length L

$$L = \int_a^b \sqrt{1 + u'^2} dx,$$

passes through the points $(a, 0)$ and $(b, 0)$, lies over the obstacle $\phi(x)$

$$u(x) \geq \phi(x), \quad \forall x \in [a, b]$$

Remark 5.4.2 We assume that the equation of the obstacle $\phi(x)$ is defined for all $x \in [a, b]$. If it is not defined for some x, we put $\phi(x) = 0$ for these x.

The Euler equation $S(F,u)=0$ corresponds to the operator

$$S(F,u) = \frac{d}{dx}\frac{u'}{\sqrt{1+u'^2}} = \frac{u''}{\sqrt{1+u'^2}^3}$$

its sign coincide with the sign of u'',

$$S(F,u) = A^2 u'', \quad \text{where } A = \frac{1}{(1+u'^2)^{\frac{3}{4}}} > 0$$

The extremal is found from the conditions (**??**) which take the form:

$$\begin{array}{llll} \text{Either} & u(x)=\phi(x) & \text{and} & u'' \leq 0, \\ \text{or} & u''=0 & \text{and} & u(x) > \phi(x) \end{array}$$

Multidimensional version The problem of the convex envelope of a function of a vector argument can be formulated as the variational inequality as well. The conditions of convexity of a differentiable function are

$$\begin{array}{ll} u(x) = f(x) \quad H(u) \geq 0 & \\ \det H(u) = 0, \ H(u) \geq 0, & u(x) < f(x) \end{array}$$

This problem will be discussed in Chapter **??**

5.5 Summary

1. Euler equations and Lagrange method in variational problems can be viewed as limits of stationary conditions of a finite-dimensional minimization problem.

2. Lagrange method allows to solve isoperimetric or constrained extremal problem of a rather general form, reducing it to the canonic variational problem. The solution is first defined as a function of the unknown multipliers, which are later determined from the constraints. Alternatively, the multipliers can be algebraically excluded from the optimality condition.

3. The total number of boundary conditions in a variational problem always matches the order of differential equations. The boundary conditions are either given or can be obtained from the minimization requirement.

4. The length of the interval of integration, if unknown, also can be obtained from the minimization requirement.

We will observe these features also in the optimization of multiple integrals: The variational problems supply of missing components of the problem formulation. We will see that they also can make the solution "better" that is more stable and even can help define the solution to the problem where no solution exist.

Chapter 6

Distinguishing minimum from maximum or saddle

Stationary conditions point to a possibly optimal trajectory but they do not answer the question of the sense of extremum. A stationary solution can correspond to minimum, local minimum, maximum, local maximum, of a saddle point of the functional. In this chapter, we establish methods aiming to distinguish local minimum from local maximum or saddle. In addition to being a solution to the Euler equation, the true minimizer satisfies necessary conditions in the form of inequalities. Here, we introduce two variational tests, Weierstrass and Jacobi conditions, that supplement each other examining various variations of the stationary trajectory.

6.1 Local variations

6.1.1 Legendre and Weierstrass Tests

The Weierstrass test detects stability of a solution to a variational problem against strong local perturbations. It compares trajectories that coincide everywhere except a small interval where their derivatives significantly differ.

Suppose that u_0 is the minimizer of the variational problem (3.1) that satisfies the Euler equation (3.7). Additionally, u_0 should satisfy another test that uses a type of variation δu different from (3.3). The variation used in the Weierstrass test is an infinitesimal triangle supported on the interval $[x_0, x_0+\varepsilon]$ in a neighborhood of a point $x_0 \in (0,\ 1)$ (see **??**):

$$\Delta u(x) = \begin{cases} 0 & \text{if } \ x \notin [x_0, x_0+\varepsilon], \\ v_1(x-x_0) & \text{if } \ x \in [x_0, x_0+\alpha\varepsilon], \\ v_2(x-x_0) - \alpha\varepsilon(v_1 - v_2) & \text{if } \ x \in [x_0+\alpha\varepsilon, x_0+\varepsilon] \end{cases}$$

where the parameters $\alpha,\ v_1,\ v_2$ are related

$$\alpha v_1 + (1-\alpha) v_2 = 0. \tag{6.1}$$

which provides the continuity of $u_0 + \Delta u$ at the point $x_0 + \varepsilon$, because it yields to the equality $\Delta u(x_0 + \varepsilon - 0) = 0$.

The considered variation (the Weierstrass variation) is localized and has an infinitesimal absolute value (if $\varepsilon \to 0$), but its derivative $(\Delta u)'$ is finite, unlike the variation in (3.3) (see **??**):

$$(\Delta u)' = \begin{cases} 0 & \text{if } \ x \notin [x_0, x_0 + \varepsilon], \\ v_1 & \text{if } \ x \in [x_0, x_0 + \alpha\varepsilon], \\ v_2 & \text{if } \ x \in [x_0 + \alpha\varepsilon, x_0 + \varepsilon]. \end{cases} \tag{6.2}$$

Computing δI from (3.2) and rounding up to ε, we find that the inequality holds

$$\begin{aligned} \delta I = \ & \varepsilon[\alpha F(x_0, u_0, u_0' + v_1) + (1-\alpha) F(x_0, u_0, u_0' + v_2) \\ & -F(x_0, u_0, u_0')] + o(\varepsilon) \geq 0 \end{aligned} \tag{6.3}$$

for a minimizer u_0. Notice that we approximately replace $u_0 + \delta u_0$ with u_0 keeping only terms of the order of $O(1)$ in the varied integrand, but we have to count for different value of the derivative.

The last expression yields to the Weierstrass test and the necessary Weierstrass condition. Any minimizer $u(x)$ of (3.1) satisfies the inequality

$$\alpha F(x_0, u_0, u_0' + v_1) + (1-\alpha) F(x_0, u_0, u_0' + v_2) - F(x_0, u_0, u_0') \geq 0. \tag{6.4}$$

Comparing this with the definition of convexity (**??**), we observe that the Weierstrass condition requires convexity of the Lagrangian $F(x, y, z)$ with respect to its third argument $z = u'$. The first two arguments x, $y = u$ here are the coordinates x, $u(x)$ of the testing minimizer $u(x)$. Recall that the tested minimizer $u(x)$ is a solution to the Euler equation.

Theorem 6.1.1 (Weierstrass test) A differentiable minimizer $u(x)$ of the simplest variational problem that solves Euler equation yields to convexity of the integrand $F(x, u, v)$ with respect of its third argument $v = u'$ when $x, u(x), u'(x)$ is an arbitrary point of the stationary trajectory.

Example 6.1.1 Consider the Lagrangian

$$F(u, u') = [(u')^2 - u^2]^2$$

It is convex as a function of u' if $|u'| \geq |u|$. Consequently, the solution u of Euler equation

$$\frac{d}{dx}[(u')^3 - u^2 u'] + u(u')^2 - u^3 = 0, \quad u(0) = a_0, \ u(1) = a_1$$

or

$$(3(u')^2 - u^2)u'' - u((u')^2 + u^2) = 0 \quad u(0) = a_0, \ u(1) = a_1$$

corresponds to a local minimum of the functional if, in addition, the inequality $|u'(x)| \geq |u(x)|$ is satisfied in all points $x \in (0, 1)$.

Remark 6.1.1 Convexity of the Lagrangian does not guarantee the existence of a solution to variational problem. It states only that a differentiable minimizer (if it exists) is stable against fine-scale perturbations. However, the minimum may not exist at all or be unstable to other variations.

If the solution of a variational problem fails the Weierstrass test, then its cost can be decreased by adding infinitesimal centered wiggles to the solution. The wiggles are the Weierstrass trial functions, which decrease the cost. In this case, we call the variational problem ill-posed, and we say that the solution is unstable against fine-scale perturbations.

Example 6.1.2 Notice that Weierstrass condition is always satisfied in the geometric optics. The Lagrangian depends on the derivative as $L = \frac{\sqrt{1+y'^2}}{v(y)}$ and its second derivative

$$\frac{\partial^2 L}{\partial y'^{\,2}} = \frac{1}{v(y)(1+y'^2)^{\frac{3}{2}}}$$

is always nonnegative if $v > 0$. It is physically obvious that the fastest path is stable to short-term perturbations.

Vector-Valued Minimizer The Euler equation and the Weierstrass condition can be naturally generalized to the problem with the vector-valued minimizer.

The Weierstrass test requires convexity of $F(x, \boldsymbol{y}, \boldsymbol{z})$ with respect to the last vector argument. Here again $\boldsymbol{y} = \boldsymbol{u}_0(x)$ represents a minimizer. If the Lagrangian is twice differentiable function of z, the convexity condition becomes

$$He(F, z) \geq 0 \tag{6.5}$$

(see Section 2.1) where $He(F, z)$ is the Hessian

$$He(F, z) = \begin{pmatrix} \frac{\partial^2 F}{\partial z_1 \partial z_1} & \cdots & \frac{\partial^2 F}{\partial z_1 \partial z_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 F}{\partial z_1 \partial z_n} & \cdots & \frac{\partial^2 F}{\partial z_n \partial z_n} \end{pmatrix}$$

and inequality in (6.5) means that the matrix is nonnegative definite (all eigenvalues are nonnegative).

Example 6.1.3 Notice that Weierstrass condition is always satisfied in the Lagrangian mechanics. The Lagrangian depends on the derivatives of the generalized coordinates through the kinetic energy $T = \frac{1}{2}\dot{q}R(q)\dot{q}$ and its Hessian equals generalized inertia R which is always positive definite. Physically speaking, inertia does not allow for infinitesimal oscillations because they always increase the kinetic energy while potential energy is insensitive to them.

Figure 6.1: The construction of Weierstrass $\mathcal{E}$-function. The graph of a convex function and its tangent plane.

Weierstrass $\mathcal{E}$-function Weierstrass suggested a convenient test for convexity of Lagrangian, the so-called $\mathcal{E}$-function equal to the difference between the value of Lagrangian $L(x,u,\hat{z})$ in a trial point $u, z = z'$ and the tangent hyperplane $L(x,u,u') - (\hat{z}-u')^T \frac{\partial L(x,u,u')}{\partial u'}$ to the optimal trajectory at the point u, u':

$$\mathcal{E}(L(x,u,u',\hat{z}) = L(x,u,\hat{z}) - L(x,u,u') - (\hat{z}-u')^T \frac{\partial L(x,u,u')}{\partial u'}$$

Function $\mathcal{E}(L(x,u,u',\hat{z})$ vanishes together with the derivative $\frac{\partial \mathcal{E}(L)}{\partial \hat{z}}$ when $\hat{z} = u'$:

$$\mathcal{E}(L(x,u,u',\hat{z})|_{\hat{z}=u'} = 0, \quad \frac{\partial}{\partial \hat{z}}\mathcal{E}(L(x,u,u',\hat{z})|_{\hat{z}=u'} = 0.$$

According to the basic definition of convexity, the graph of a convex function is greater than or equal to a tangent hyperplane. Thereafter, the Weierstrass condition of minimum of the objective functional can be written as the condition of positivity of the Weierstrass $\mathcal{E}$-function for the Lagrangian,

$$\mathcal{E}(L(x,u,u',\hat{z}) \geq 0 \quad \forall \hat{z}, \ \forall x, u(x)$$

where $u(x)$ tested trajectory.

Example 6.1.4 Check the optimality of Lagrangian

$$L = u'^4 - \phi(u,x)u'^2 + \psi(u,x)$$

where ϕ and ψ are some functions of u and x using Weierstrass $\mathcal{E}$-function.

The Weierstrass $\mathcal{E}$-function for this Lagrangian is

$$\mathcal{E}(L(x,u,u',\hat{z}) = \left[\hat{z}^4 - \phi(u,x)\hat{z}^2 + \psi(u,x)\right]$$
$$- \left[u'^4 - \phi(u,x)u'^2 + \psi(u,x)\right] - (\hat{z}-u')(4u'^3 - 2\phi(u,x)u).$$

or

$$\mathcal{E}(L(x,u,u',\hat{z}) = (\hat{z}-u')^2\left(\hat{z}^2 + 2\hat{z}u' - \phi + 3u'^2\right).$$

As expected, $\mathcal{E}(L(x,u,u',\hat{z})$ is independent of an additive term ψ and contains a quadratic coefficient $(\hat{z}-u')^2$. It is positive for any trial function $\hat{z}$ if the quadratic

$$\pi(\hat{z}) = -\hat{z}^2 - 2\hat{z}u' + \phi - 3u'^2$$

does not have real roots, or if

$$\phi(u,x) - 2u^2 \leq 0$$

If this condition is violated at a point of an optimal trajectory $u(x)$, the trajectory is nonoptimal.

6.1.2 Null-Lagrangians and convexity

Find the Lagrangian cannot be uniquely reconstructed from its Euler equation. Similarly to antiderivative, it is defined up to some term called null-Lagrangian.

Definition 6.1.1 The Lagrangians $\phi(x, \boldsymbol{u}, \boldsymbol{u}')$ for which the operator $S(\phi, u)$ of the Euler equation (3.26) identically vanishes

$$S(\phi, u) = 0 \quad \forall u$$

are called *Null-Lagrangians*.

Null-Lagrangians in variational problems with one independent variable are linear functions of $\boldsymbol{u}'$. Indeed, the Euler equation is a second-order differential equation with respect to $\boldsymbol{u}$:

$$\frac{d}{dx}\left(\frac{\partial}{\partial \boldsymbol{u}'}\phi\right) - \frac{\partial}{\partial \boldsymbol{u}}\phi = \frac{\partial^2 \phi}{\partial(\boldsymbol{u}')^2} \cdot \boldsymbol{u}'' + \frac{\partial^2 \phi}{\partial \boldsymbol{u}' \partial \boldsymbol{u}} \cdot \boldsymbol{u}' + \frac{\partial^2 \phi}{\partial \boldsymbol{u} \partial x} - \frac{\partial \phi}{\partial \boldsymbol{u}} \equiv 0. \quad (6.6)$$

The coefficient of $\boldsymbol{u}''$ is equal to $\frac{\partial^2 \phi}{\partial(\boldsymbol{u}')^2}$. If the Euler equation holds identically, this coefficient is zero, and therefore $\frac{\partial \phi}{\partial \boldsymbol{u}'}$ does not depend on $\boldsymbol{u}'$. Hence, ϕ linearly depends on $\boldsymbol{u}'$:

$$\begin{aligned} \phi(x, \boldsymbol{u}, \boldsymbol{u}') &= \boldsymbol{u}' \cdot A(\boldsymbol{u}, x) + B(\boldsymbol{u}, x); \\ A &= \tfrac{\partial^2 \phi}{\partial \boldsymbol{u}' \partial \boldsymbol{u}}, \quad B = \tfrac{\partial^2 \phi}{\partial \boldsymbol{u} \partial x} - \tfrac{\partial \phi}{\partial \boldsymbol{u}}. \end{aligned} \quad (6.7)$$

If, inn addition, the following equality holds

$$\frac{\partial A}{\partial x} = \frac{\partial B}{\partial \boldsymbol{u}}, \quad (6.8)$$

then the Euler equation vanishes identically. In this case, ϕ is a null-Lagrangian.

We notice that the Null-Lagrangian (6.7) is simply a full differential of a function $\Phi(x, u)$:

$$\phi(x, \boldsymbol{u}, \boldsymbol{u}') = \frac{d}{dx}\Phi(x, u) = \frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial u}u';$$

equations (6.8) are the integrability conditions (equality of mixed derivatives) for Φ. The vanishing of the Euler equation corresponds to the Fundamental theorem of calculus: The equality

$$\int_a^b \frac{d\Phi(x, u)}{dx} dx = \Phi(b, u(b)) - \Phi(a, u(a)).$$

that does not depend on $u(x)$ only on its end-points values.

Example 6.1.5 Function $\phi = u\,u'$ is the null-Lagrangian. Indeed,we check

$$\frac{d}{dx}\left(\frac{\partial}{\partial u'}\phi\right) - \frac{\partial}{\partial u}\phi = u' - u' \equiv 0.$$

Remark 6.1.2 We will show in Section **??** that nonlinear null-Lagrangians in multivariable problems exist that express the integrability conditions.

Null-Lagrangians and Convexity The convexity requirements of the Lagrangian F that follow from the Weierstrass test are in agreement with the concept of null-Lagrangians (see, for example [?]).

Consider a variational problem with the Lagrangian F,

$$\min_{\boldsymbol{u}} \int_0^1 F(x, \boldsymbol{u}, \boldsymbol{u}')dx.$$

Adding a null-Lagrangian ϕ to the given Lagrangian F does not affect the Euler equation of the problem. The family of problems

$$\min_{\boldsymbol{u}} \int_0^1 \left(F(x, \boldsymbol{u}, \boldsymbol{u}') + t\phi(x, \boldsymbol{u}, \boldsymbol{u}')\right) dx,$$

where t is an arbitrary number, corresponds to the same Euler equation. Therefore, each solution to the Euler equation corresponds to a family of Lagrangians $F(x, \boldsymbol{u}, \boldsymbol{z}) + t\phi(x, \boldsymbol{u}, \boldsymbol{z})$, where t is an arbitrary real number. This says, in particular, that a Lagrangian cannot be uniquely defined by the solution to the Euler equation.

The stability of the minimizer against the Weierstrass variations should be a property of the Lagrangian that is independent of the value of the parameter t. It should be a common property of the family of equivalent Lagrangians. On the other hand, if $F(x, \boldsymbol{u}, \boldsymbol{z})$ is convex with respect to $\boldsymbol{z}$, then $F(x, \boldsymbol{u}, \boldsymbol{z}) + t\phi(x, \boldsymbol{u}, \boldsymbol{z})$ is also convex. Indeed, $\phi(x, \boldsymbol{u}, \boldsymbol{z})$ is linear as a function of $\boldsymbol{z}$, and adding the term $t\phi(x, \boldsymbol{u}, \boldsymbol{z})$ does not affect the convexity of the sum. In other words, convexity is a characteristic property of the family. Accordingly, it serves as a test for the stability of an optimal solution.

6.2 Weak and strong local minima

6.2.1 Norms in functional space

Calculus of variation studies increment of a functional at close-by curves. The answer to the question whether or not two curves are close to each other, depends on definition of closeness. This question is studied in theory of topological spaces. Unlike the distance between two points in finite-dimensional Euclidian space, the same two curves can be considered to be infinitesimally close or far parted depending of the meaning of "distance." The variational tests examine the stability of the stationary solutions to small perturbations; different tests differently define the smallness of perturbation.

In calculus of variations, there are three mostly used criteria to measure the closeness of two differentiable functions $f_1(x)$ and $f_2(x)$: The norm $\mathcal{N}_1$ of difference $\delta f(x) = f_1(x) - f_2(x)$ in the values of functions

$$\mathcal{N}_1(\delta f) = \max_{x \in (0,1)} |\delta f(x)|$$

the norm $\mathcal{N}_2$ of difference of their derivatives,

$$\mathcal{N}_2(\delta f) = \max_{x \in (0,1)} |\delta f'(x)|$$

and the length $\mathcal{N}_3$ of the interval on which these functions are different

$$\mathcal{N}_3(\delta f) = \Delta \quad \text{if} \quad \delta f(x) = 0 \ \forall x \notin [x, x + \Delta]$$

None of variational tests guarantees the global optimality of the tested trajectory, only local minimum; at the other hand, these tests are simple enough to be applied to practically interesting problems. The *local minimum* satisfies the inequality

$$I(u) \leq I(u + \delta u) \quad \forall \delta u : \ \mathcal{N}(\delta u(x)) < \varepsilon$$

where ε is infinitesimally small and $\mathcal{N}$ is a norm. The definition of what is local minimum depends on the above definitions of the norm $\mathcal{N}$.

If the perturbation is small in the following sense

$$\mathcal{N}_{\text{Legendre}}(\delta u) = \mathcal{N}_1(\delta u) + \mathcal{N}_2(\delta u) + \mathcal{N}_3(\delta u) < \varepsilon$$

the Legendre text is satisfied. The test assumes that the compared functions and their derivatives are close everywhere, and they are identical outside of an infinitesimal interval.

The Weierstrass text assumes that the compared functions are close everywhere, and they are identical outside of an infinitesimal interval, but their derivatives are not close in that infinitesimal interval of variation:

$$\mathcal{N}_{\text{Weierstrass}}(\delta u) = \mathcal{N}_1(\delta u) + \mathcal{N}_3(\delta u) < \varepsilon, \quad \mathcal{N}_2(\delta u) \text{ is arbitrary.}$$

If the objective functional satisfy the Weierstrass test we say that the extremal $u(x)$ realizes a *strong local minimum*. The Weierstrass test is stronger than the Legendre test.

The Jacobi test (see below, Section 6.3) assumes that

$$\mathcal{N}_{\text{Jacobi}}(\delta u) = \mathcal{N}_1(\delta u) + \mathcal{N}_2(\delta u) < \varepsilon, \quad \mathcal{N}_3(\delta u) \text{ is arbitrary}$$

that is the compared functions and their derivatives are close everywhere, but the variation is not localized. The Jacobi test is stronger than the Legendre test. If Jacobi test is satisfied we say that the extremal $u(x)$ realizes a *weak local minimum* (not to be confused with minimum of weakly convergent sequence or with minimum for localized variations). Neither Weierstrass and Jacobi tests is stronger than the other: They test the stationary trajectory from different angles.

6.2.2 Sufficient condition for the weak local minimum

We assume that a trajectory $u(x)$ satisfies the stationary conditions and Legendre condition. We investigate the increment caused by a nonlocal variation δu of an infinitesimal magnitude:

$$\mathcal{N}_{\text{Jacobi}}(\delta u) = \mathcal{N}_1(\delta u) + \mathcal{N}_2(\delta u) < \varepsilon, \quad \mathcal{N}_3(\delta u) \text{ is arbitrary.}$$

To compute the increment, we expand the Lagrangian into Taylor series keeping terms up to $O(\epsilon^2)$. Recall that the linear of ϵ terms are zero because the Euler equation $S(u,u')=0$ for $u(x)$ holds. We have

$$\delta I = \int_0^r S(u,u')\delta u\, dx + \int_0^r \delta^2 F dx + o(\epsilon^2) \tag{6.9}$$

where

$$\delta^2 F = \frac{\partial^2 F}{\partial u^2}(\delta u)^2 + 2\frac{\partial^2 F}{\partial u \partial u'}(\delta u)(\delta u') + \frac{\partial^2 F}{\partial (u')^2}(\delta u')^2 \tag{6.10}$$

No variation of this kind can improve the stationary solution if the quadratic form

$$Q(u,u') = \begin{pmatrix} \frac{\partial^2 F}{\partial u^2} & \frac{\partial^2 F}{\partial u \partial u'} \\ \frac{\partial^2 F}{\partial u \partial u'} & \frac{\partial^2 F}{\partial (u')^2} \end{pmatrix}$$

is positively defined,

$$Q(u,u') > 0 \quad \text{on the stationary trajectory } u(x) \tag{6.11}$$

This condition is called the sufficient condition for the weak minimum because it neglects the relation between δu and $\delta u'$ and treats them as independent trial functions. If the sufficient condition is satisfied, no trajectory that is smooth and sufficiently close to the stationary trajectory can increase the objective functional of the problem compared with the objective at that tested stationary trajectory.

Notice that the term $\frac{\partial^2 F}{\partial u'^2}$ is nonnegative because of the Legendre condition (**??**).

Example 6.2.1 Show that the sufficient condition is satisfied for the Lagrangians

$$F = \frac{1}{2}u^2 + \frac{1}{2}(u')^2 \text{ and } F_2 = \frac{1}{|u|}(u')^2$$

Next example shows that violation of the sufficient conditions can yield to nonexistence of the solution.

Example 6.2.2 (Stationary solution is not a minimizer) Consider the variational problem:

$$I = \min_u \int_0^r \left(\frac{1}{2}(u')^2 - \frac{c}{2}u^2\right) dx \quad u(0)=0; \quad u(r)=A$$

where c is a constant. The first variation δI is zero,

$$\delta I = \int_0^r \left(u'' + c^2 u\right) \delta u dx = 0$$

if $u(x)$ satisfies the Euler equation

$$u'' + c^2 u = 0, \quad u(0)=0, u(r)=A. \tag{6.12}$$

The stationary solution $u(x)$ is

$$u(x) = \left(\frac{A}{\sin(cr)}\right)\sin(cx)$$

The Weierstrass test is satisfied, because the dependence of the Lagrangian on the derivative u' is convex, $\frac{\partial L}{\partial^2 u'^2} = c^2$.

The second variation equals

$$\delta^2 I = \int_0^r \left(\frac{1}{2}(\delta u')^2 - \frac{c^2}{2}(\delta u)^2\right) dx$$

Since the ends of the trajectory are fixed, the variation δu satisfies homogeneous conditions $\delta u(0) = \delta u(r) = 0$. Let us choose the variation as follow:

$$\delta u = \begin{cases} \epsilon x(a-x), & 0 \le x \le a \\ 0 & x > a \end{cases}$$

where the interval of variation $[0, a]$ is not greater that $[0, r]$, $a \le r$. Computing the second variation, we obtain

$$\delta^2 I(a) = \frac{\epsilon^2}{60} a^3 (c^2 a^2 - 10), \quad a \le r$$

If second variation $\delta^2 I(a)$ is negative, $\delta^2 I(a) < 0$ the stationary solution does not correspond to the minimum of I. The second variation of the chosen type depends on a and $\delta^2 I$ is maximal when $a = r$. This maximum is negative when

$$r > r_{\text{crit}} = \frac{\sqrt{10}}{c}$$

We conclude that the stationary solution does not correspond to the minimum of I if the length of the trajectory is larger than r_{crit}. If the length is smaller than r_{crit}, the situation is inconclusive because we could choose another type of variation different from considered here and disprove the optimality of the stationary solution.

6.3 Jacobi variation

The Jacobi condition examines the optimality of "long" trajectories. It complements the Weierstrass test that investigates stability of a Lagrangian to strong localized variations. Jacobi condition tries to disprove optimality of a stationary trajectory by testing the dependence of Lagrangian on the minimizer itself not of its derivative. This condition is stronger than the sufficient condition for the weak minimum. We assume that a trajectory $u(x)$ satisfies the stationary conditions and Weierstrass condition but does not satisfy the sufficient conditions for weak minimum, $Q(u, u')$ is not positively defined.

To derive Jacobi condition, we apply again an infinitesimal nonlocal variation: $\delta u = O(\epsilon) \ll 1$ and $\delta u' = O(\epsilon) \ll 1$ and examine the expression (6.10) for the second variation. Notice that we denote the upper limit of integration in (6.10) by r; we are testing the stability of the trajectory depending on its length. When a nonlocal "shallow" variation is applied, the increment increases because of assumed positivity of $\frac{\partial^2 F}{\partial (u')^2}$ and decreases because of assumed nonpositivity of the matrix Q. Depending on the length r of the interval of integration and of chosen form of the variation δu, one of these effects prevails. If the second effect is stronger, the extremal fails the test and is nonoptimal.

Let us choose the best shape δu of the variation. The expression (6.10) itself is a variational problem for δu which we rename as v; the Lagrangian is quadratic of v and v' and the coefficients are functions of x determined by the stationary trajectory $u(x)$:

$$\delta I = \int_0^r \left[A v^2 + 2B\, v\, v' + C (v')^2 \right] dx, \quad v(0) = v(r) = 0 \tag{6.13}$$

where

$$A = \frac{\partial^2 F}{\partial u^2}, \quad B = \frac{\partial^2 F}{\partial u \partial u'}, \quad C = \frac{\partial^2 F}{\partial (u')^2}$$

The problem (6.13) correspond to the Euler equation that is a solution to Storm-Liusville problem:

$$\frac{d}{dx}(Cv' + Bv) - Av = 0, \quad v(0) = v(r_{\text{conj}}) = 0 \quad \text{if } r < r_{\text{conj}} \tag{6.14}$$

with boundary conditions $v(0) = v(r) = 0$. The point r_{conj} is called a conjugate point to the end of the interval. The problem is homogeneous: If $v(x)$ is a solution and c is a real number, $cv(x)$ is also a solution.

Jacobi condition is satisfied if the interval does not contain conjugate points, that is there is no nontrivial solutions to (6.14) on any subinterval of $[0, r_{\text{conj}}]$, that is if there are no nontrivial solutions of (6.14) with boundary conditions $v(r) = v(r_{\text{conj}}) = 0$ where $0 \le r_{\text{conj}} \le r$.

If this condition is violated, than there exist a family of trajectories

$$u(x) \begin{cases} u_0 + v & \text{if } \ x \in [0, r_{\text{conj}}] \\ u_0 & \text{if } \ x \in [r_{\text{conj}}, r] \end{cases}$$

that deliver the same value of the cost. Indeed, v is defined up to a multiplier: If v is a solution, αv is a solution too. These trajectories have discontinuous derivative at the points r_1 and r_2 which leads to a contradiction to the Weierstrass-Erdman condition that does not allow a broken extremal at these points.

Examples

Example 6.3.1 (Nonexistence of the minimizer: Blow up) Consider again the problem in example 6.2.2

$$I = \min_u \int_0^r \left(\frac{1}{2}(u')^2 - \frac{c^2}{2} u^2 \right) dx \quad u(0) = 0; \quad u(r) = A$$

The stationary trajectory and the second variation are give by formulas (13.21) and (13.21), respectively.

Instead of arbitrary choosing the second variation, we now choose it as a solution to the homogeneous problem (6.14) for $v = \delta u$

$$v'' + c^2 v = 0, \quad u(0) = 0, u(r_{\text{conj}}) = 0, \quad r_{\text{conj}} \leq r \tag{6.15}$$

This problem has a nontrivial solution $v = \epsilon \sin(cx)$ if the length of the interval is large enough to satisfy homogeneous condition of the right end, $cr_{\text{conj}} = \pi$ or

$$r \geq r(_{\text{conj}}) = \frac{\pi}{c}$$

The second variation $\delta^2 I$is negative when r is large,

$$\delta^2 I \leq \frac{1}{r}\epsilon^2 \left(\frac{\pi^2}{r^2} - c^2\right) < 0 \quad \text{if } r > \frac{\pi}{c}$$

which shows that the a stationary solution is not a minimizer.

To clarify the phenomenon, let us compute the stationary solution from the Euler equation (6.12). We have

$$u(x) = \left(\frac{A}{\sin(cr)}\right)\sin(cx) \quad \text{and } I(u) = \frac{A^2}{\sin^2(cr)}\left(\frac{\pi^2}{r^2} - c^2\right)$$

When r increases approaching the value $\frac{c}{\pi}$, the magnitude of the stationary solution indefinitely grows, and the cost indefinitely decreases:

$$\lim_{r \to \frac{c}{\pi} - 0} I(u) = -\infty$$

Obviously, the solution of the Euler equation that corresponds to finite $I(u)$ when $r > \frac{\pi}{c}$ is not a minimizer.

Remark 6.3.1 Comparing this result with the result in Example (6.3.1), we see that the optimal choice of variation improved the result at only 0.65%.

6.3.1 Does Nature minimize action?

The next example deals with a system of multiple degrees of freedom.

Consider the variational problem with the Lagrangian

$$L = \sum_{i=1}^{n} \frac{1}{2} m {u_i'}^2 - \frac{1}{2} C(u_i - u_{i-1})^2, \quad u(0) = u_0$$

We will see later in Chapter **??** that this Lagrangian describes the *action* of a chain of particles with masses m connected by springs with constant C. The second variation

$$\delta^2 L = \sum_{i=1}^{n} \frac{1}{2} m \dot{v_i}^2 - \frac{1}{2} C(v_i - v_{i-1})^2, \quad v_0 = 0, \quad v_n = 0$$

corresponds to the Euler equation – the eigenvalue problem

$$m\ddot{V} = \frac{C}{m} AV$$

where $V = v_1(t), \dots, v_n(t)$ and

$$A = \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ 0 & 1 & -2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -2 \end{pmatrix}.$$

The problem has a solution – vector $v(t)$

$$v(t) = \sum \alpha_k v_k \sin \omega_k t \quad v(0) = v(T_{\text{conj}}) = 0, \quad T_{\text{conj}} \le T$$

where v_k are the eigenvectors, α are coefficients found from initial conditions, and ω_k are the square roots of eigenvalues of the matrix A. Solving the characteristic equation for eigenvalues $\det(A - \omega^2 I) = 0$ we find that these eigenvalues are

$$\omega_k = 2\sqrt{\frac{C}{m}} \sin^2\left(\sqrt{\frac{C}{m}}\frac{\pi k}{n}\right), \quad k = 1, \dots n$$

The Jacobi condition is violated if $v(t)$ is consistent with the homogeneous initial and final conditions that is if the time interval is short enough. Namely, The condition is violated when the duration T is larger than

$$T \ge \frac{\pi}{\max(\omega_k)} \approx 2\pi\sqrt{\frac{m}{C}}$$

.

The continuous limit of the chain with the masses is achieved when the number N of notes indefinitely growth and their mass decreases correspondingly as $\frac{m}{N}$, and the stiffness of one link growth as CN as it become N times shorter. Correspondingly,

$$\sqrt{\frac{C(N)}{m(N)}} = N\sqrt{\frac{C(0)}{m(0)}}$$

and the maximal eigenvalue tends to infinity as $N \to \infty$. This implies that the action J of the continuous system is not minimized at any time interval T.

What is minimized in classical mechanics?

Chapter 7

Irregular solutions and Relaxation

Every problem of the calculus of variations has a solution, provided that the word "solution" is suitably understood.

David Hilbert

7.1 Exotic and classical solutions

The classical approach to variational problems assumes that the optimal trajectory is a differentiable curve – a solution to the Euler equation that, in addition, satisfies the Weierstrass and Jacobi tests. In this chapter, we consider the variational problems which solutions do not satisfy necessary conditions of optimality. Either the Euler equation does not have solution, or Jacobi or Weierstrass tests are not satisfied; in any case, the extremal cannot be found from stationarity conditions. We have seen such solution in the problem of minimal surface (Goldschmidt solution, Section 4.1.3).

A minimization problem always can be solved in a way because it allows for a minimizing sequence: the functions $u^s(t)$ with the property $I(u^\epsilon) \geq I(u^{s+1})$. The functionals $I(u^\epsilon)$ form a monotonic sequence of real number that converges to a real or improper limit. In this sense, every variational problem can be solved, but the limiting solution $\lim_{s\to\infty} u^S$ may be irregular; in other terms, it may not exist in an assumed set of functions. Especially, derivation of Euler equation uses an assumption that the minimum is a differentiable function. This assumption leads to complications because the set of differentiable functions is open and the limits of sequences of differentiable functions are not necessary differentiable functions themselves. For example, the limit of the sequence of infinitely differentiable function

$$\phi_n(x) = \frac{n}{2\pi} \exp\left(\frac{x^2}{2n}\right)$$

is not even a function but a distribution - the δ function. The limit $H(x)$ of infinitely differentiable functions

$$H(x) = \int_{-\infty}^{x} \phi_n(x)dx = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 1 \end{cases}$$

is a discontinuous Heaviside function. The limit

$$\lim_{n\to\infty} \sin(nx)$$

does not exist for any $x \neq 0$. The sequence $\frac{1}{n}\sin(nx)$ converges to zero, but its derivative does not converge to a limit but infinitely often oscillates in the interval $[-1, 1]$, and derivative of a converging to zero sequence $\frac{1}{\sqrt{n}}\sin(nx)$ does not converge to a limit and is unbounded. These or similar sequences can represent minimizing sequences in a variational problems. Here we give a brief introduction to the methods to deal with such "exotic" solutions"

How to deal with irregular problems The possible nonexistence of minimizer poses several challenging questions. First, criteria are needed to establish which problems have a classical solution and which have not. These criteria analyze the type of Lagrangians and results in existence theorems.

The are two alternative ideas in handling problems with nondifferentiable minimizers. The admissible class of minimizers can be enlarged and closed it in a way so that the "exotic" limits of minimizers would be included in the admissible set. This procedure, the *relaxation* underlined in the Hilbert's quotation motivated the introduction of distributions and the corresponding functional spaces, as well as development of relaxation methods. Below, we consider several ill-posed problems that require rethinking of the concept of a solution.

Alternatively, the minimization problem can be constrained so that the exotic solutions are penalized and the problem will avoid them; this approach called *regularization* forces the problem select a classical solution at the expense of increasing the objective functional.

Existence conditions We do not prove here existence theorems because the arguments use theorems from functional analysis. Instead, we outline the ideas of such theorems and refer to more advance and rigorous books for the proofs. We formulate here a list of conditions guarantying the smooth classical solution to a variational problem.

1. The Lagrangian superlinearly growths with respect to u':

$$\lim_{|u'|\to\infty} \frac{F(x, u, u')}{|u'|} = \infty \quad \forall x, u(x) \tag{7.1}$$

   This condition forbids any finite jumps of the optimal trajectory $u(x)$; any such jump leads to an infinite penalty in the problem's cost.

2. The cost of the problem growths indefinitely when $|u| \to \infty$. This condition forbids a blow-up of the solution.

3. The Lagrangian is convex with respect to u':

$$F(x, u, u') \quad \text{is convex function of } u' \quad \forall x, u(x)$$

at the optimal trajectory u. This condition forbids infinite oscillations because they would increase the cost of the problem.

Idea of the proof:

1. First two conditions guarantee that the limit of any minimizing sequence is bounded and continuous. The cost of the problem unlimitedly growths when either the function or its derivative tent to infinity at a set of nonzero measure.

2. It is possible extract a weakly convergent subsequence $u^S \rightharpoonup u^0$ from a bounded minimizing sequence. Roughly, this means that the subsequence $u^\epsilon(x)$ in a sense approximates a limiting function u^0, but may wiggle around it infinitely often.

3. The convexity of Lagrangian eliminates the possibility of wiggling, because the cost of the problem with convex Lagrangian is smaller at a smooth function than on at any close-by wiggling function.

The conditions of the theorem can be alternated. Example?

7.2 Unbounded solutions. Regularization

7.2.1 Examples of discontinuous solutions

We start with three simple examples of variational problems with discontinuous solution. We apply regularization to it, an approach to deal with ill-posed variational problems as (7.2). According to this approach, we slightly change the Lagrangian and arrive at a regular (differentiable) solution. Then we may consider the sequence of perturbed solutions when the perturbation parameter tends to zero.

A problem with discontinuous extremal

Consider the minimization problem

$$I_0 = \min_{u(x)} I(u), \quad I(u) = \int_{-1}^{1} x^2 u'^2 dx, \quad u(-1) = -1, \quad u(1) = 1, \qquad (7.2)$$

We observe that $I(u) \geq 0 \ \forall u$, and therefore $I_0 \geq 0$. The Lagrangian is convex function of u', and the third condition is satisfied. However, the second condition is violated in $x = 0$:

$$\lim_{|u'| \to \infty} \left. \frac{x^2 u'^2}{|u'|} \right|_{x=0} = 0$$

The functional is of sublinear growth at only one point $x = 0$.

Let us show that the solution is discontinuous. Assume the contrary, that the solution satisfies the Euler equation $(x^2 u')' = 0$ everywhere. The equation admits the integral

$$\frac{\partial L}{\partial u'} = 2x^2 u' = C.$$

If $C \neq 0$, the value of $I(u)$ is infinity, because then $u' = \frac{C}{2x^2}$, the Lagrangian becomes

$$x^2 u'^2 = \frac{C^2}{x^2} \quad \text{if } C \neq 0.$$

and the integral of Lagrangian diverges. A finite value of the objective corresponds to $C = 0$ which implies that $u_0'(x) = 0$ if $x \neq 0$. Accounting for the boundary conditions, we find

$$u_0(x) = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases}$$

and $u_0(0)$ is not defined.

We arrived at the unexpected result that violate the assumptions used when the Euler equation is derived: $u_0(x)$ is discontinuous at $x = 0$ and u_0' exists only in the sense of distributions:

$$u_0(x) = -1 + 2H(x), \quad u_0'(x) = 2\delta(x)$$

The question is how should this result be interpreted. In the classical sense, the solution of this problem does not exist. However, the discontinuous minimizer makes sense even if it does not belong to the set of differentiable functions.

Stabilizers We may slightly perturb the problem so that it has a classical solution that is close to the discontinuous solution of the original problem. Regularization can be performed by adding to the Lagrangian a stabilizer, a strictly convex function $\epsilon\rho(u')$ of superlinear growth.

Consider the perturbed problem for the Example 7.2:

$$I_\epsilon = \min_{u(x)} I_\epsilon(u), \quad I_\epsilon(u) = \int_{-1}^{1} \left(x^2 u'^2 + \epsilon^2 u'^2\right) dx, \quad u(-1) = -1, \quad u(1) = 1, \tag{7.3}$$

Here, the perturbation $\epsilon^2 u'$ is added to the original Lagrangian $\epsilon^2 u'$; the perturbed Lagrangian is of superlinear growth everywhere.

The first integral of the Euler equation for the perturbed problem becomes

$$(x^2 + \epsilon^2) u' = C, \quad \text{or } du = C \frac{dx}{x^2 + \epsilon^2}$$

Integrating and accounting for the boundary conditions, we obtain

$$u_\epsilon(x) = C \arctan \frac{x}{\epsilon}, \quad C = \left(\arctan \frac{1}{\epsilon}\right)^{-1}$$

When $\epsilon \to 0$, the solution $u_\epsilon(x)$ converges to $u_0(x)$ although the convergence is not uniform at $x = 0$.

Solution with δ-sequence

Consider the variational problem with an inequality constraint

$$\max_{u(x)} \int_0^\pi u' \sin(x) dx, \quad u(0) = 0,\ u(\pi) = 1,\ u'(x) \geq 0\ \forall x.$$

The minimizer either corresponds to the limit of derivative

$$u'(x) = 0, \quad x \in [\alpha_i, \beta_i],$$

on subintervals $[\alpha_i, \beta_i]$ of $[0, \pi]$, or it satisfies the stationary condition in $[\beta_i, \alpha_{i+1}]$ between the intervals of constancy. The derivative cannot be zero everywhere, because this would correspond to a constant $u(x)$ and would violate the boundary conditions.

However, the minimizer cannot correspond to the solution of Euler equation at any interval. Indeed, the Lagrangian depends only on x and u'. The first integral $\frac{\partial L}{\partial u'} = C$ of the Euler equation yields to an absurd result

$$\sin(x) = \text{ constant } \quad \forall x \in [\beta_i, \alpha_{i+1}]$$

The Euler equation does not gives the minimizer. Something is wrong!

This problem can be immediately solved by the inequality

$$\int_0^\pi f(x) g(x) dx \leq \left(\max_{x \in [0,\pi]} g(x) \right) \int_0^\pi |f(x)| dx.$$

that is valid for all functions f and g if the involved integrals exist. We set $g(x) = \sin(x)$ and $f(x) = |f(x)| = u'$ (because u' is nonnegative), account for the constraints

$$\int_0^\pi |f(x)| dx = u(\pi) - u(0) = 1 \quad \text{ and } \max_{x \in [0,\pi]} \sin(x) = 1,$$

and obtain the upper bound

$$I(u) = \int_0^\pi u' \sin(x) dx \leq 1 \quad \forall u.$$

This bound is achievable by the limit the minimizing sequence that tends to a Heaviside function $\phi_n(x) \to H(x - \pi/2)$. Notice that the derivative of such sequence tends to the δ-function, $u'(x) = \delta(x - \pi/2)$. We check that the bound is realizable,

$$\int_0^\pi \delta\left(x - \frac{\pi}{2}\right) \sin(x) dx = \sin\left(\frac{\pi}{2}\right) = 1.$$

but the minimizer is discontinuous.

This problem also can be regularized. Here, we show another way to regularization, by imposing an additional pointwise inequality $u'(x) \leq M\ \forall x$. Because the intermediate values of u' are never optimal, it would alternate the limiting values:

$$u'_M(x) = \begin{cases} 0 & \text{if } x \notin \left[\frac{\pi}{2} - \frac{1}{2M}, \frac{\pi}{2} + \frac{1}{2M}\right], \\ M & \text{if } x \in \left[\frac{\pi}{2} - \frac{1}{2M}, \frac{\pi}{2} + \frac{1}{2M}\right], \end{cases}$$

The objective functional is equal to

$$I(u_M) = M \int_{\frac{\pi}{2}-\frac{1}{2M}}^{\frac{\pi}{2}+\frac{1}{2M}} \sin(x)dx = 2M \sin\left(\frac{1}{2M}\right)$$

When M tends to infinity, I_M goas to its limit

$$\lim_{M\to\infty} I_M = 1,$$

the length $\frac{1}{M}$ of the interval where $u' = M$ goes to zero so that $u'_M(t)$ weakly converges to the δ-function for u', $u'_M(t) \rightharpoonup \delta\left(x - \frac{\pi}{2}\right)$.

The constrained variational problems are subject to control theory; they are are discussed later in Chapter 10.1.

Discontinuities in problems of geometrical optics

We have already seen in Section 4.1.3 that the minimal surface problem

$$I_0 = \min_{u(x)} I(u), \quad I(u) = \int_o^L u\sqrt{1+(u')^2}dx, \quad u(-1) = 1, \quad u(1) = 1, \quad (7.4)$$

can lead to a discontinuous solution (Goldschmidt solution)

$$u = -H(x+1) + H(x-1)$$

if L is larger than a threshold. The minimal surface is an example of problem of geometric optics. We should ask: What is the reason of discontinuous solutions in such problems? Notice that the Lagrangian in minimal surface problem (and in any other problem of geometric optics) is of linear growth:

$$F = u\sqrt{1+(u')^2}, \quad \lim_{|u'|\to\infty} \frac{F(x,u,u')}{|u'|} = u$$

which hints of a possible appearance of the discontinuous solution.

Let us investigate the discontinuous solutions of Lagrangians of linear growth. Suppose that a minimizing sequence u^ϵ of differentiable functions tends to a discontinuous at the point x_0 function, as follows

$$\begin{aligned} u^\epsilon(x) &= \phi^\epsilon(x) + \psi^\epsilon(x) \\ \phi^\epsilon(x) &\rightharpoonup \phi(x) \\ \phi^\epsilon(x) &\rightharpoonup \beta H(x - x_0), \quad \beta \neq 0 \end{aligned}$$

where ϕ and ψ are differentiable functions. More specifically, assume that ψ is piece-wise linear,

$$\psi^\epsilon(x) = \begin{cases} 0 & \text{if } x < x_0 - \epsilon \\ \frac{\beta}{\epsilon}(x - x_0 + \epsilon) & \text{if } x_0 - \epsilon \le x \le x_0 \\ \beta & \text{if } x > x_0. \end{cases}$$

where $\epsilon = \frac{1}{s}$.

Notice that ψ' is zero outside of the interval $[x_0 - \epsilon, x_0]$ where it is equal to a large constant,

$$\psi' = \begin{cases} 0 & \text{if } x \notin [x_0 - \epsilon, x_0] \\ \frac{1}{\epsilon} & \text{if } x \in [x_0 - \epsilon, x_0] \end{cases}$$

and compute

$$\sqrt{1 + [(u^\epsilon)']^2} = \begin{cases} \sqrt{1 + [(\phi^\epsilon)']^2} & \text{if } x \notin [x_0 - \epsilon, x_0] \\ \sqrt{1 + \left(\frac{\beta^2}{\epsilon^2} + (\phi^\epsilon)'\right)^2} = \frac{\beta}{\epsilon} + o\left(\frac{1}{\epsilon}\right) & \text{if } x \in [x_0 - \epsilon, x_0] \end{cases}$$

We observe that $\sqrt{1 + [(u^\epsilon)']^2}$ is independent of ψ outside of the interval $[x_0 - \epsilon, x_0]$ where it is equal to a large constant that depends only on the magnitude of the jump.

The objective integral stays finite in spite of the indefinite growth of the derivative in $[x_0 - \epsilon, x_0]$ because of smallness of this interval,

$$\begin{aligned} I(u^\epsilon) \;=\; & \int_0^{x_0 - \epsilon} \phi^\epsilon \sqrt{1 + (\phi^{\epsilon\prime})^2} dx \\ & + \; M + \int_{x_0}^{L} (\beta + \phi^\epsilon) \sqrt{1 + (\phi^{\epsilon\prime})^2} dx + o\left(\frac{1}{\epsilon}\right) \end{aligned}$$

where M represent the contribution of the interval $[x_0 - \epsilon, x_0]$. Computing M, we replace the term $\sqrt{1 + (u^{\epsilon\prime})^2}$ by its estimate,

$$M = \int_{x_0 - \epsilon}^{x_0} u^\epsilon \sqrt{1 + (u^{\epsilon\prime})^2} dx \approx \frac{\beta}{\epsilon} \int_{x_0 - \epsilon}^{x_0} \phi(x) dx + \frac{\beta}{\epsilon} \int_{x_0 - \epsilon}^{x_0} \psi^\epsilon(x)) dx.$$

Substituting the expression for ψ^ϵ and using the continuity of ϕ, we obtain

$$M = \beta\phi(x_0) + \frac{1}{2}\beta^2 \quad \text{if } \epsilon \to 0$$

The contribution M due to the discontinuity of the minimizer is finite when the magnitude $|\beta|$ of the jump is finite. Therefore, discontinuous solutions are tolerated in the geometric optics: They do not lead to infinitely large values of the objective functionals. The Goldschmidt solution corresponds to zero smooth component $u(x) = 0$, $x = (a, b)$ and two jumps M_1 and M_2 of the magnitudes $u(a)$ and $u(b)$, respectively. The smooth component gives zero contribution, and the contributions of the jumps are

$$I = \frac{1}{2}\left(u^2(a) + u^2(b)\right)$$

Why are the problems nonregular Notice that both examples, (7.4) and (7.2) do not satisfy assumption (**??**). Notice also, that in the problem (7.2), assumption (**??**) is violated at only one point, $x = 0$.

To the contrary, problems of Lagrange mechanics do satisfy this assumption because kinetic energy depends of the speed $\dot{q}$ quadratically.

7.2.2 Regularization

Regularization as smooth approximation The smoothing out feature of regularization is easy demonstrated on the following example of a quadratic approximation of a function by a smoother one.

Approximate a function $f(x)$ where $x \in \mathcal{R}$, by the function $u(x)$, adding a quadratic stabilizer; this problem takes the form

$$\min_u \int_{-\infty}^{\infty} [\epsilon^2 (u')^2 + (u - f)^2] dx$$

The Euler equation

$$\epsilon^2 u'' - u = -f \tag{7.5}$$

can be easily solved using the Green function

$$G(x, y) = \frac{1}{2\epsilon} \exp\left(-\frac{|x - y|}{\epsilon} \right)$$

of the operator in the left-hand side of (7.5). We have

$$u(x) = \frac{1}{2\epsilon} \int_{-\infty}^{\infty} \exp\left(-\frac{|x - y|}{\epsilon} \right) f(y) dy$$

that is the expression of the averaged f. The smaller is ϵ the closer is the average to f.

Quadratic stabilizers Besides the stabilizer $\varepsilon u'^2$, other stabilizers can be considered: The added term εu^2 penalizes for large values of the minimizer, $\varepsilon (u'')^2$ penalizes for the curvature of the minimizer and is insensitive to linearly growing solutions. The stabilizers can be inhomogeneous like $\varepsilon (u - u_{\text{target}})^2$; they force the solution stay close to a target value. The choice of a specific stabilizer depends on the physical arguments (see Tikhonov).

For example, solve the problem with the Lagrangian

$$F = \epsilon^4 (u'')^2 + (u - f(x)^2$$

Show that $u = f(x)$ if $f(x)$ is any polynomial of the order not higher than three. Find an integral representation for $u(f)$ if the function $f(x)$ is defined at the interval $|x| \leq 1$ and at the axis $x \in R$.

7.2.3 Regularization of a finite-dimensional problem

As the most of variational methods, the regularization has a finite-dimensional analog. It is applicable to the minimization problem of a convex but not strongly convex function which may have infinitely many solutions. The idea of regularization is to slightly perturb the function by small but a strictly convex term; the perturbed problem has a unique solution to matter how small the perturbation is. The numerical advantage of the regularization is the convergence of minimizing sequences.

Let us illustrate ideas of regularization by studying a finite dimensional problem. Consider a linear system

$$Ax = b \tag{7.6}$$

where A is a square $n \times b$ matrix and b is a known n-vector.

We know from linear algebra that the Fredholm Alternative holds:

- If $\det A \neq 0$, the problem has a unique solution:

$$x = A^{-1}b \quad \text{if } \det A \neq 0 \tag{7.7}$$

- If $\det A = 0$ and $Ab \neq 0$, the problem has no solutions.
- If $\det A = 0$ and $Ab = 0$, the problem has infinitely many solutions.

In practice, we also deal with an additional difficulty: The determinant $\det A$ may be a "very small" number and one cannot be sure whether its value is a result of rounding of digits or it has a "physical meaning." In any case, the errors of using the formula (7.7) can be arbitrary large and the norm of the solution is not bounded.

To address this difficulties, it is helpful to restate linear problem (7.6) as an extremal problem:

$$\min_{x \in R^n} (Ax - b)^2 \tag{7.8}$$

This problem does have at least one solution, no matter what the matrix A is. This solution coincides with the solution of the original problem (7.6) when this problem has a unique solution; in this case the cost of the minimization problem (7.8) is zero. Otherwise, the minimization problem provides "the best approximation" of the non-existing solution.

If the problem (7.6) has infinitely many solutions, so does problem (7.8). Corresponding minimizing sequences $\{x^s\}$ can be unbounded, $\|x^s\| \to \infty$ when $s \to \infty$.

In this case, we may select a solution with minimal norm. We use the *regularization*, passing to the perturbed problem

$$\min_{x \in R^n} (Ax - b)^2 + \epsilon x^2$$

The solution of the last problem exists and is unique. Indeed, we have by differentiation

$$(A^T A + \epsilon I)x - A^T b = 0$$

and

$$x = (A^T A + \epsilon I)^{-1} A^T b$$

We mention that

1. The inverse exists since the matrix $A^T A$ is nonnegative defined, and ϵ is positively defined. The eigenvalues of the matrix $(A^T A + \epsilon I)^{-1}$ are not smaller than ϵ^{-1}

2. Suppose that we are dealing with a well-posed problem (7.6), that is the matrix A is not degenerate. If $\epsilon \ll 1$, the solution approximately is $x = A^{-1}b - \epsilon(A^2 A^T)^{-1} b$ When $\epsilon \to 0$, the solution becomes the solution (7.7) of the unperturbed problem, $x \to A^{-1}b$.

3. If the problem (7.6) is ill-posed, the norm of the solution of the perturbed problem is still bounded:
$$||x|| \leq \frac{1}{\epsilon} ||b||$$

Remark 7.2.1 Instead of the regularizing term ϵx^2, we may use any positively define quadratic $\epsilon(x^T P x + p^T x)$ where matrix P is positively defined, $P > 0$, or other strongly convex function of x.

7.2.4 Growth conditions and discontinuous extremals

Growth conditions of Lagrangians Assume that the Lagrangian is bounded as follows:

$$C_L(x,y)|z|^L \leq \lim_{|z|\to\infty} F(x,y,z) \leq C_U(x,y)|z|^U \quad \forall x,y$$

where C_L and C_U are lower and upper bound and L and U are positive numbers. We prove that

Theorem 7.2.1 (Growth conditions) Depending on the growth conditions, the problems can be classified as follows.

- If $L > 1$ the Lagrangian is of superlinear growth. The derivative of a minimizer is bounded almost everywhere. The minimizer is continuous. Euler equation could correspond to a minimizer.

- If $U < 1$ the Lagrangian is of sublinear growth. Notice that it is also nonconvex, and the Weierstrass condition is violated. The derivative of a minimizer can be unbounded almost everywhere. Optimal trajectory is a saw-teeth curve with dense set of intervals of arbitrary fast growth. The Euler equation never corresponds to a minimum. A variational problem with Lagrangian of sublinear growth
$$I = \inf_{u:u(0)=a_0,u(1)=a_1} \int_0^1 F_{\text{sub}}(x,u,u')dx,$$

has the objective functional equal to

$$I = \int_0^1 \phi(x)\, dx \quad \phi(x) = \min_z \ \min_y F(x, y, z)$$

independently of boundary values a_1 and a_2.

- If $L \leq 1$ and $U \geq 1$ the Lagrangian may have either continuous or discontinuous solutions. Euler equation could correspond to a minimizer.

Discontinuous minimizers We investigate here if the minimizer $u(x)$ can be discontinuous at a point x_0 inside the interval of integration $[0, 1]$:

$$u(x) = \begin{cases} \phi_1(x), \ x < x_0 \\ \phi_2(x), \ x > x_0 \end{cases}$$

and

$$\phi_1(x_0 - 0) \neq \phi_2(x_0 + 0).$$

We start with a lemma:

Theorem 7.2.2 Consider a differentiable function $F(x, y, z)$ and a sequence of functions $u_s(x) = u_0(x) + j_s(x)$ such that u_0 and its derivative are bounded everywhere at $[0, 1]$,

$$|u_0(x)| \leq C, \quad |u_0'(x)| \leq C \quad \forall x \in [0, 1],$$

$j_s(x)$ is constant everywhere except in an interval $[x_0, x_0 + \frac{1}{s}] \subset [0, 1]$,

$$j_s(x) = \begin{cases} \alpha & \text{if } x \in [0, x_0] \\ \beta & \text{if } x \in \left[x_0 + \frac{1}{s}, 1\right] \end{cases}$$

where it growths unlimitedly fast, and a proper or improper limit exists

$$\lim_{z \to \infty} \frac{F(x, y, z)}{z} = A \tag{7.9}$$

where $-\infty \leq A \leq \infty$.

Then

$$\lim_{s \to \infty} \int_0^1 F(x, u_s, u_s')dx = \int_0^{x_0} F(x, u_s + \alpha, u_s')dx + \int_{x_0}^1 F(x, u_s + \beta, u_s')dx + A(\beta - \alpha) \tag{7.10}$$

Proof: The first term in the right-hand side of (21.15) is obtained by direct substitution of the value of j_s; the second term is the limit of the integral

$$\int_{x_0 + \frac{1}{s}}^1 F(x, u_s, u_s')dx$$

when $s \to \infty$. The third term appears by virtue of (7.9) and because $j_s' \gg u_0'$:

$$\lim_{s\to\infty} \int_{x_0}^{x_0+\frac{1}{s}} F(x, u_0 + j_s, u_0' + j_s)dx =$$
$$\lim_{s\to\infty} \int_{x_0}^{x_0+\frac{1}{s}} \frac{F(x, u_0 + j_s, u_0' + j_s')}{u_0' + j_s'}(u_0' + j_s')dx =$$
$$A \lim_{s\to\infty} \int_{x_0}^{x_0+\frac{1}{s}} (j_s')dx = A(\beta - \alpha)$$

We assume that the functions $\phi_1(x)$ and $\phi_2(x)$ minimize the variational functional on the intervals $[0, x_0]$ and $[x_0, 1]$, respectively, respectively,

$$I_1(\phi_1) = \min_u J_1(u), \quad I_2(\phi_2) = \min_u J_2(u)$$

If the Lagrangian is of superlinear growth, $L > 1$, then discontinuities never occurs in minimizing sequences: The penalty for them is infinitely high. If the Lagrangian is of sublinear growth, $U < 1$, then the finite discontinuities in minimizing sequences are not penalized at all. If the Lagrangian is of linear growth, $L = U = 1$, then I_3 is finite. The discontinuities may occurs in minimizing sequences: The penalty for them is finite.

7.3 Infinitely oscillatory solutions: Relaxation

7.3.1 Nonconvex Variational Problems. An example

When a Lagrangian $F(x, \boldsymbol{y}, \boldsymbol{z})$ of the problem

$$\inf_u J(u), \quad J(u) = \inf_u \int_0^1 F(x, u, u')dx, \quad u(0) = a_0, \ u(1) = a_1 \tag{7.11}$$

is nonconvex with respect to $\boldsymbol{z}$, the Weierstrass test fails. A minimizing sequence cannot tend to a differentiable curve in the limit; otherwise it would satisfy the Euler equation and the Weierstrass test,

Definition 7.3.1 We call the *forbidden region* $\mathcal{Z}_f$ the set of $\boldsymbol{z}$ for which $F(x, \boldsymbol{y}, \boldsymbol{z})$ is not convex with respect to $\boldsymbol{z}$,

The derivative $\boldsymbol{u}'$ of a minimizer $\boldsymbol{u}$ of (7.11) should never belong to the region Z_f:

$$\boldsymbol{u}' \notin Z_f. \tag{7.12}$$

Instead, the minimizer "jumps over" the forbidden set, and does it infinitely often. Because of this jumps, the minimizer stays outside of the forbidden interval but its average can take any value within or outside the forbidden region.

We will demonstrate that a minimizing sequence tends to a "generalized curve." It consists of infinitesimal zigzags. The limiting curve has a dense set of points of discontinuity of the derivative. A detailed explanation of this phenomenon can be found, for example, in Young:1942:GSCa, Young:1942:GSCb, Gamkrelidze:1962:SOS,Young:1969:LCV, Warga:1972:OCD, Gamkrelidze:1985:SMO. Here we give a brief description of it, mainly by working on several examples.

To deal with a nonconvex problem, we "relax" it. *Relaxation* means that we replace the problem with another one that has the same cost but whose solution is stable against fine-scale perturbations; particularly, it cannot be improved by the Weierstrass variation. The relaxed problem has the following two basic properties:

- The relaxed problem has a classical solution.
- The infimum of the functional (the cost of the problem) in the initial problem coincides with the cost of the relaxed problem.

Here we will demonstrate two approaches to relaxation based on necessary and sufficient conditions. Each of them yields to the same construction but uses different arguments to achieve it. In the next chapters we will see similar procedures applied to variational problems with multiple integrals; sometimes they also yield the same construction, but generally they result in different relaxations.

A non-convex problem Consider a simple variational problem that yields to an exotic solution [**?**]:

$$\inf_u I(u) = \inf_u \int_0^1 G(u,u')dx, \quad u(0) = u(1) = 0 \tag{7.13}$$

where

$$G(u,v) = u^2 + \begin{cases} (v-1)^2, & \text{if } \ v \geq \frac{1}{2} \\ \frac{1}{2} - v^2 & \text{if } \ -\frac{1}{2} \leq v \leq \frac{1}{2} \\ (v+1)^2 & \text{if } \ v \leq -\frac{1}{2} \end{cases} . \tag{7.14}$$

The graph of the function $G(.,v)$ is presented in **??**B; it is a nonconvex twice differentiable function of v of superlinear growth.

The Lagrangian G penalizes the trajectory u for having the speed $|u'|$ different from ± 1 and penalizes the deflection of the trajectory u from zero. These contradictory requirements cannot be resolved in the class of classical trajectories.

Indeed, a differentiable minimizer satisfies the Euler equation (**??**) that takes the form

$$\begin{array}{ll} u'' - u = 0 \ \text{ if } & |u'| \geq \frac{1}{2} \\ u'' + u = 0 \ \text{ if } & |u'| \leq \frac{1}{2}. \end{array} \tag{7.15}$$

The Weierstrass test additionally requires convexity of $G(u,v)$ with respect to v; the Lagrangian $G(u,v)$ is nonconvex in the interval $v \in (-1, \ 1)$ (see **??**).

The Weierstrass test requires that the extremal (7.20) is supplemented by the inequality (recall that $v = u'$)

$$u' \notin (-1,\ 1) \quad \text{at the optimal trajectory.} \tag{7.16}$$

and it is not clear how to satisfy it. Indeed, the Euler equation does not leave a freedom to change the trajectory to avoid the *forbidden interval*. Notice also, that the second regime in (7.20) is never optimal because it is realized inside of the forbidden interval.

The Weierstrass-Erdman condition that requires continuity of $\frac{\partial L}{\partial u'}$ permits switching between the first ($u' > 1/2$) and third ($u' < -1/2$) regimes in (**??**) when

$$2\left(u'_{(1)} - 1\right) = 2\left(u'_{(3)} + 1\right)$$

or when

$$u'_{(1)} = 1, \quad u'_{(3)} = -1$$

which means the switching from one end of the forbidden interval to another

Remark 7.3.1 Observe, that the easier verifiable Legendre condition $\frac{\partial^2 F}{\partial (u')^2} \geq 0$ gives a twice smaller forbidden region $|u'| \leq \frac{1}{2}$ and is not in the agreement with Weierstrass-Erdman condition. One should always use stronger conditions!

Minimizing sequence The minimizing sequence for problem (7.13) can be immediately constructed. Indeed, the infimum of (7.13) obviously is nonnegative, $\inf_u I(u) \geq 0$. Therefore, a sequence u^s with the property

$$\lim_{s \to \infty} I(u^s) = 0 \tag{7.17}$$

is a minimizing sequence.

Consider a set of functions $\tilde{u}^s(x)$ with the derivatives equal to ± 1 at each point,

$$\tilde{u}'(x) = \pm 1 \quad \forall x.$$

These functions belong to the boundary of the *forbidden interval* of the nonconvexity of $G(., v)$; they make the second term in the Lagrangian (7.14) vanish, $G(u, v) = u^2$, and the problem becomes

$$I(\tilde{u}^s, (\tilde{u}^s)') = \int_0^1 (\tilde{u}^s)^2 dx. \tag{7.18}$$

The sequence $\tilde{u}^s$ oscillates near zero if the derivative $(\tilde{u}^s)'$ changes its sign on intervals of equal length. The cost $I(\tilde{u}^s)$ depends on the density of switching points and tends to zero when the number of these points increases (see **??**). Therefore, the minimizing sequence consists of the saw-tooth functions $\tilde{u}^s$; the heights of the teeth tend to zero and their number tends to infinity as $s \to \infty$.

Note that the minimizing sequence $\{\tilde{u}^s\}$ does not converge to any classical function. This minimizer $\tilde{u}^s(x)$ satisfies the contradictory requirements, namely, the derivative must keep the absolute value equal to one, but the function itself must be arbitrarily close to zero:

$$|(\tilde{u}^s)'| = 1 \quad \forall x \in [0,1], \qquad \max_{x\in[0,1]} \tilde{u}^s \to 0 \quad \text{as } s \to \infty. \tag{7.19}$$

The limiting curve u_0 has zero norm in $C_0[0,\ 1]$ but a finite norm in $C_1[0,\ 1]$.

Remark 7.3.2 Below, we consider this problem with arbitrary boundary values; the solution corresponds partly to the classical extremal (7.20), (7.16), and partly to the saw-tooth curve; in the last case u' belongs on the boundary of the forbidden interval $|u'| = 1$.

Regularization and relaxation This considered nonconvex problem is another example of an ill-posed variational problem. For these problems, the classical variational technique based on the Euler equation fails to work. Here, The limiting curve is not a discontinuous curve as in the previous example, but a limit of infinitely fast oscillating functions, similar to $\lim_{\omega\to\infty} \sin(\omega x)$.

We may apply regularization to discourage the solution to oscillate infinitely often. For example, we may penalize for the discontinuity of the u' adding the stabilizing term $\epsilon(u'')^2$ to the Lagrangian. Doing this, we pass to the problem

$$\min_u \int_0^1 (\epsilon^2(u'')^2 + G(u,u'))dx$$

that corresponds to Euler equation:

$$\begin{aligned} \epsilon^2 u^{IV} - u'' + u = 0 \quad &\text{if} \quad |u'| \geq \tfrac{1}{2} \\ \epsilon^2 u^{IV} + u'' + u = 0 \quad &\text{if} \quad |u'| \leq \tfrac{1}{2}. \end{aligned} \tag{7.20}$$

The Weierstrass condition this time requires the convexity of the dependence of Lagrangian on u''; this condition is satisfies.

One can see that the solution of equation (7.20) is oscillatory; the period of oscillation is of the order of $\epsilon \ll 1$: The solution still tends to an infinitely often oscillating distribution. When ϵ is positive but small, the solution has finite but large number of wiggles. The computation of such solutions is difficult and often unnecessary: It strongly depends on an artificial parameter ϵ, which is difficult to justify physically. Although formally the solution of regularized problem exists, here the regularization does not accomplish much: The problem is still computationally difficult and the difficulty grows when $\epsilon \to 0$. Other methods are needed to deal with such problems.

Below we describe the relaxation of a nonconvex variational problem. The idea of relaxation is in a sense opposite to regularization. Instead of penalization for fast oscillations, we admit them as a legitime minimizers enlarging set of minimizers. The main problem is to find an adequate description of infinitely often switching controls in terms of smooth functions. It turns out that the limits of oscillating minimizers allows for a parametrization and can be effectively described by a several smooth functions: the values of alternating limits for u' and the average time that minimizer spends on each limit.

7.3.2 Minimal Extension

We introduce the idea of relaxation of a variational problem. Consider the class of Lagrangians $\mathcal{N}F(x,y,z)$ that are smaller than $F(x,y,z)$ and satisfy the Weierstrass test $\mathcal{W}(\mathcal{N}F(x,y,z)) \geq 0$:

$$\begin{cases} \mathcal{N}F(x,y,z) - F(x,y,z) \leq 0, \\ \mathcal{W}(\mathcal{N}F(x,y,z)) \geq 0 \end{cases} \quad \forall\ x,y,z. \tag{7.21}$$

Let us take the maximum on $\mathcal{N}F(x,y,z)$ and call it $\mathcal{S}F$. Clearly, $\mathcal{S}F$ corresponds to turning one of these inequalities into an equality:

$$\begin{array}{llll} \mathcal{S}F(x,y,z) = F(x,y,z), & \mathcal{W}(\mathcal{S}F(x,y,z)) \geq 0 & \text{if} & z \notin \mathcal{Z}_{\text{f}}, \\ \mathcal{S}F(x,y,z) \leq F(x,y,z), & \mathcal{W}(\mathcal{S}F(x,y,z)) = 0 & \text{if} & z \in \mathcal{Z}_{\text{f}}. \end{array} \tag{7.22}$$

This variational inequality describes the extension of the Lagrangian of an unstable variational problem. Notice that

1. The first equality holds in the region of convexity of F and the extension coincides with F in that region.
2. In the region where F is not convex, the Weierstrass test of the extended Lagrangian is satisfied as an equality; this equality serves to determine the extension.

These conditions imply that $\mathcal{S}F$ is convex everywhere. Also, $\mathcal{S}F$ is the maximum over all convex functions that do not exceed F. Again, $\mathcal{S}F$ is equal to the convex envelope of F:

$$\mathcal{S}F(x,y,z) = \mathcal{C}_z F(x,y,z). \tag{7.23}$$

The cost of the problem remains the same because the convex envelope corresponds to a minimizing sequence of the original problem.

Remark 7.3.3 Note that the geometrical property of convexity never explicitly appears here. We simply satisfy the Weierstrass necessary condition everywhere. Hence, this relaxation procedure can be extended to more complicated multidimensional problems for which the Weierstrass condition and convexity do not coincide.

Recall that the derivative of the minimizer never takes values in the region $\mathcal{Z}_{\text{f}}$ of nonconvexity of F. Therefore, a solution to a nonconvex problem stays the same if its Lagrangian $F(x,\boldsymbol{y},\boldsymbol{z})$ is replaced by any Lagrangian $\mathcal{N}F(x,\boldsymbol{y},\boldsymbol{z})$ that satisfies the restrictions

$$\begin{array}{l} \mathcal{N}F(x,\boldsymbol{y},\boldsymbol{z}) = F(x,\boldsymbol{y},\boldsymbol{z}) \quad \forall\ z \notin \mathcal{Z}_{\text{f}}, \\ \mathcal{N}F(x,\boldsymbol{y},\boldsymbol{z}) > \mathcal{C}F(x,\boldsymbol{y},\boldsymbol{z})\ \forall\ z \in \mathcal{Z}_{\text{f}}. \end{array} \tag{7.24}$$

Indeed, the two Lagrangians $F(x,\boldsymbol{y},\boldsymbol{z})$ and $\mathcal{N}F(x,\boldsymbol{y},\boldsymbol{z})$ coincide in the region of convexity of F. Therefore, the solutions to the variational problem also coincide in this region. Neither Lagrangian satisfies the Weierstrass test in the forbidden region of nonconvexity. Therefore, no minimizer can distinguish between these two problems: It never takes values in Z_{f}. The behavior of the Lagrangian in the forbidden region is simply of no importance. In this interval, the Lagrangian cannot be computed back from the minimizer.

Minimizing Sequences Let us prove that the considered extension preserves the value of the objective functional. Consider the extremal problem (7.11) of superlinear growth and the corresponding stationary solution $u(x)$ that may not satisfy the Weierstrass test. Let us perturb the trajectory u by a differentiable function $\omega(x)$ with the properties:

$$\max_x |\omega(x)| \leq \varepsilon, \quad \omega(x_k) = 0 \ k = 1 \dots N$$

where the points x_k uniformly cover the interval (a, b). The perturbed trajectory wiggles around the stationary one, crossing it at N uniformly distributed points; the derivative of the perturbation is not bounded.

The integral $J(u, \omega)$on the perturbed trajectory

$$J(u, \omega) = \int_0^1 F(x, u + \omega, u' + \omega')dx$$

is estimated as

$$J(u, \omega) = \int_0^1 F(x, u, u' + \omega')dx + o(\varepsilon).$$

because of the smallness of ω. The derivative $\omega'(x) = v(x)$ is a new minimizer constrained by N conditions

$$\int_{\frac{k}{N}}^{\frac{k+1}{N}} v(x)dx = 0, k = 0, \dots N - 1;$$

correspondingly, the variational problem can be rewritten as

$$J(u, \omega) = \sum_{k=1}^{N-1} \int_{\frac{k}{N}}^{\frac{k+1}{N}} F(x, u, u' + \omega')dx + o\left(\frac{1}{N}\right).$$

Perform minimization of a term of the above sum with respect of v:

$$I_k(u) = \min_{v(x)} \int_{\frac{k}{N}}^{\frac{k+1}{N}} F(x, u, u' + v)dx \quad \text{subject to} \quad \int_{\frac{k}{N}}^{\frac{k+1}{N}} v(x)dx = 0$$

This is exactly the problem (13.21) of the convex envelope with respect to v.

By referring to the Carathéodory theorem (2.13) we conclude that the minimizer $v(x)$ is a piece-wise constant function in $(\frac{k}{N}, \frac{k+1}{N})$ that takes at most $n + 1$ values $v_1, \dots v_{n+1}$ at the intervals of the length $m_1 L, \dots m_{n+1} L$, where $L = \frac{k}{N}$ is the length of the interval of integration. These values are subject to the constraints (see (**??**))

$$m_i(x) \geq 0, \quad \sum_{i=1}^{n} m_i = 1, \quad \sum_{i=1}^{p} m_i \boldsymbol{v}_i = 0. \qquad (7.25)$$

Average derivative	Pointwise derivatives	Optimal concentrations	Convex envelope $\mathcal{C}G(u,v)$
$v < -1$	$v_1^0 = v_2^0 = v$	$m_1^0 = 1,\ m_2^0 = 0$	$u^2 + (v-1)^2$
$\lvert v\rvert < 1$	$v_1^0 = 1, v_2^0 = -1$	$m_1^0 = m_2^0 = \frac{1}{2}$	u^2
$v > 1$	$v_1^0 = v_2^0 = v$	$m_1^0 = 0,\ m_2^0 = 1$	$u^2 + (v+1)^2$

Table 7.1: Characteristics of an optimal solution in Example **??**.

This minimum coincides with the convex envelope of the original Lagrangian with respect to its last argument (see (2.13)):

$$I_k = \min_{m_i, \boldsymbol{v}_i \in (7.25)} \frac{1}{N} \left(\sum_{i=1}^{p} m_i F(x, \boldsymbol{u}, u' + \boldsymbol{v}_i) \right) \tag{7.26}$$

Summing I_k and passing to the limit $N \to \infty$, we obtain the relaxed variational problem:

$$I = \min_{\boldsymbol{u}} \int_0^1 \mathcal{C}_{\boldsymbol{u}'} F\left(x, \boldsymbol{u}(x), \boldsymbol{u}'(x)\right) dx. \tag{7.27}$$

Note that $n+1$ constraints (7.25) leave the freedom to choose $2n+2$ inner parameters m_i and $\boldsymbol{v}_i$ to minimize the function $\sum_{i=1}^{p} m_i F(u, \boldsymbol{v}_i)$ and thus to minimize the cost of the variational problem (see (7.26)). If the Lagrangian is convex, $\boldsymbol{v}_i = 0$ and the problem keeps its form: The wiggle trajectories do not minimize convex problems.

The cost of the reformulated (relaxed) problem (7.27) corresponds to the cost of the problem (7.11) on the minimizing sequence (**??**). Therefore, the cost of the relaxed problem is equal to the cost of the original problem (7.11). The extension of the Lagrangian that preserves the cost of the problem is called the *minimal extension*. The minimal extension enlarges the set of classical minimizers by including generalized curves in it.

7.3.3 Examples

Relaxation of nonconvex problem in Example ?? We revisit Example **??**. Let us solve this problem by building the convex envelope of the Lagrangian $G(u,v)$:

$$\begin{aligned} \mathcal{C}_v G(u,v) &= \min_{m_1, m_2} \min_{v_1, v_2} \left\{ u^2 + m_1(v_1-1)^2 + m_2(v_2+1)^2 \right\}, \\ v &= m_1 v_1 + m_2 v_2, \quad m_1 + m_2 = 1, \quad m_i \geq 0. \end{aligned} \tag{7.28}$$

The form of the minimum depends on the value of $v = u'$. The convex envelope $\mathcal{C}G(u,v)$ coincides with either $G(u,v)$ if $v \notin [0,\ 1]$ or $\mathcal{C}G(u,v) = u^2$ if $v \in [0,\ 1]$; see Example 2.1.6. Optimal values $v_1^0,\ v_2^0,\ m_1^0\ m_2^0$ of the minimizers and the convex envelope $\mathcal{C}G$ are shown in Table 7.1. The relaxed form of the problem with zero boundary conditions

$$\min_u \int_0^1 \mathcal{C}G(u,\ u')dx, \quad u(0) = u(1) = 0, \tag{7.29}$$

has an obvious solution,

$$u(x) = u'(x) = 0, \tag{7.30}$$

that yields the minimal (zero) value of the functional. It corresponds to the constant optimal value $m_{\rm opt}$ of $m(x)$:

$$m_{\rm opt}(x) = \frac{1}{2}\ \forall x \in [0,1]$$

.

The relaxed Lagrangian is minimized over four functions u, m_1, v_1, v_2 bounded by one equality, $u' = m_1 v_1 + (1 - m_1)v_2$ and the inequalities $0 \le m \le 1$, while the original Lagrangian is minimized over one function u. In contrast to the initial problem, the relaxed one has a differentiable solution in terms of these four controls.

Inhomogeneous boundary conditions Let us slightly modify this example. Assume that boundary conditions are

$$u(0) = V\ (0 < V < 1), \quad u(1) = 0$$

In this case, an optimal trajectory of the relaxed problem consists of two parts,

$$u' < -1 \quad \text{if } x \in [0, x_0), \quad u = u' = 0 \quad \text{if } x \in [x_0, 1]$$

At the first part of the trajectory, the Euler equation $u'' - u = 0$ holds; the extremal is

$$u = \begin{cases} Ae^x + Be^{-x} & \text{if } x \in [0, x_0) \\ 0 & \text{if } x \in [x_0, 1] \end{cases}$$

Since the contribution of the second part of the trajectory is zero, the problem becomes

$$I = \min_{u, x_0} \int_O^{x_0} \mathcal{C}_v G(u, u')dx$$

To find unknown parameters A, B and x_0 we use the conditions

$$u(0) = V, \quad u(x_0) = 0, \quad u' = -1$$

The last condition expresses the optimality of x_0, it is obtained from the condition (see (**??**))

$$\mathcal{C}_v G(u, u')|_{x=x_0} = 0.$$

We compute

$$A + B = V, \quad Ae^{x_0} + Be^{-x_0} = 0, \quad Ae^x - Be^{-x} = 1$$

which leads to

$$u(x) = \begin{cases} \sinh(x - x_0) & \text{if } x < x_0, \\ 0 & \text{if } x > x_0, \end{cases}$$
$$x_0 = \sinh^{-1}(V)$$

The optimal trajectory of the relaxed problem decreases from V to zero and then stays equal zero. The optimal trajectory of the actual problem decays to zero and then become infinite oscillatory with zero average.

Relaxation of a two-wells Lagrangian We turn to a more general example of the relaxation of an ill-posed nonconvex variational problem. This example highlights more properties of relaxation. Consider the minimization problem

$$\min_{u(x)} \int_0^z F_p(x, u, u')dx, \quad u(0) = 0, \; u'(z) = 0 \tag{7.31}$$

with a Lagrangian

$$F_p = (u - \alpha x^2)^2 + F_n(u'), \tag{7.32}$$

where

$$F_n(v) = \min\{a\, v^2, \; b\, v^2 + 1\}, \quad 0 < a < b, \; \alpha > 0.$$

Note that the second term F_n of the Lagrangian F_p is a nonconvex function of u'.

The first term $(u - \alpha x^2)^2$ of the Lagrangian forces the minimizer u and its derivative u' to increase with x, until u' at some point reaches the interval of nonconvexity of $F_n(u')$. The derivative u' must vary outside of the forbidden interval of nonconvexity of the function F_n at all times.. Formally, this problem is ill-posed because the Lagrangian is not convex with respect to u' (**??**); therefore, it needs relaxation. Convexification of the Lagrangian (top) and the minimizer (bottom); points a and b are equal to v_1 and v_2, respectively. Convexification of the Lagrangian and the minimizer f2.4 0.4

To find the convex envelope $\mathcal{C}F$ we must transform $F_n(u')$ (in this example, the first term of F_p (see (7.32)) is independent of u' and it does not change after the convexification). The convex envelope $\mathcal{C}F_p$ is equal to

$$\mathcal{C}F_p = (u - \alpha x^2)^2 + \mathcal{C}F_n(u'). \tag{7.33}$$

The convex envelope $\mathcal{C}F_n(u')$ is computed in Example 2.1.7 (where we use the notation $v = u'$). The relaxed problem has the form

$$\min_u \int \mathcal{C}F_p(x, u, u')dx, \tag{7.34}$$

where

$$\mathcal{C}F_p(x, u, u') = \begin{cases} (u - \alpha x^2)^2 + a(u')^2 & \text{if } |u'| \le v_1, \\ (u - \alpha x^2)^2 + 2u'\sqrt{\frac{ab}{a-b}} - \frac{b}{a-b} & \text{if } v_1 \le |u'| \le v_2, \\ (u - \alpha x^2)^2 + b(u')^2 + 1 & \text{if } |u'| \ge v_2. \end{cases}$$

Note that the variables u, v in the relaxed problem are the averages of the original variables; they coincide with those variables everywhere when $\mathcal{C}F = F$. The Euler equation of the relaxed problem is

$$\begin{cases} au'' - (u - \alpha x^2) = 0 & \text{if } \ |u'| \le v_1, \\ (u - \alpha x^2) = 0 & \text{if } \ v_1 \le |u'| \le v_2, \\ bu'' - (u - \alpha x^2) = 0 & \text{if } \ |u'| \ge v_2. \end{cases} \tag{7.35}$$

The Euler equation is integrated with the boundary conditions shown in (7.31). Notice that the Euler equation degenerates into an algebraic equation in the interval of convexification. The solution u and the variable $\frac{\partial}{\partial u'}\mathcal{C}F$ of the relaxed problem are both continuous everywhere.

Integrating the Euler equations, we sequentially meet the three regimes when both the minimizer and its derivative monotonically increase with x (see **??**). If the length z of the interval of integration is chosen sufficiently large, one can be sure that the optimal solution contains all three regimes; otherwise, the solution may degenerate into a two-zone solution if $u'(x) \le v_2 \ \forall x$ or into a one-zone solution if $u'(x) \le v_1 \ \forall x$ (in the last case the relaxation is not needed; the solution is a classical one).

Let us describe minimizing sequences that form the solution to the relaxed problem. Recall that the actual optimal solution is a generalized curve in the region of nonconvexity; this curve consists of infinitely often alternating parts with the derivatives v_1 and v_2 and the relative fractions $m(x)$ and $(1 - m(x))$:

$$v = \langle u'(x) \rangle = m(x)v_1 + (1 - m(x))v_2, \quad u' \in [v_1, v_2], \tag{7.36}$$

where $\langle \ \rangle$ denotes the average, u is the solution to the original problem, and $\langle u \rangle$ is the solution to the homogenized (relaxed) problem.

The Euler equation degenerates in the second region into an algebraic one $\langle u \rangle = \alpha x^2$ because of the linear dependence of the Lagrangian on $\langle u \rangle'$ in this region. The first term of the Euler equation,

$$\frac{d}{dx}\frac{\partial F}{\partial \langle u \rangle'} \equiv 0 \quad \text{ if } \ v_1 \le |\langle u \rangle'| \le v_2, \tag{7.37}$$

vanishes at the optimal solution.

The variable m of the generalized curve is nonzero in the second regime. This variable can be found by differentiation of the optimal solution:

$$(\langle u \rangle - \alpha x^2)' = 0 \quad \Longrightarrow \quad \langle u \rangle' = 2\alpha x. \tag{7.38}$$

This equality, together with (7.36), implies that

$$m = \begin{cases} 0 & \text{if } \ |u'| \le v_1, \\ \frac{2\alpha}{v_1 - v_2}x - \frac{v_2}{v_1 - v_2} & \text{if } \ v_1 \le |u'| \le v_2, \\ 1 & \text{if } \ |u'| \ge v_2. \end{cases} \tag{7.39}$$

Variable m linearly increases within the second region (see **??**). Note that the derivative u' of the minimizing generalized curve at each point x lies on the boundaries v_1 or v_2 of the forbidden interval of nonconvexity of F; the average derivative varies only due to varying of the fraction $m(x)$ (see **??**).

7.4 Lagrangians of sublinear growth

Discontinuous extremals Some applications, such as an equilibrium in organic or breakable materials, deal with Lagrangians of sublinear growth. If the Lagrangian $F_{\text{sub}}(x,u,u')$ growths slower that $|u'|$,

$$\lim_{|z|\to\infty} \frac{F_{\text{sub}}(x,y,z)}{|z|} = 0 \quad \forall x,y$$

then the discontinuous trajectories are expected because the functional is insensitive to finite jumps of the trajectory.

The Lagrangian is obviously a nonconvex function of u', The convex envelope of a bounded from below function $F_{\text{sub}}(x,y,z)$ of a sublinear with respect to z growth is independent of z.

$$\mathcal{C}F_{\text{sub}}(x,y,z) = \min_z F_{\text{sub}}(x,y,z) = F_{\text{conv}}(x,y)$$

In the problems of sublinear growth, the minimum $f(x)$ of the Lagrangian correspond to pointwise condition

$$f(x) = \min_u \min_v F(x,u,v)$$

instead of Euler equation. The second and the third argument become independent of each other. The condition $v' = u$ is satisfied (as an average) by fast growth of derivatives on the set of dense set of interval of arbitrary small the summary measure. Because of sublinear growth of the Lagrangian, the contribution of this growth to the objective functional is infinitesimal.

Namely, at each infinitesimal interval of the trajectory $x_0, x_0+\varepsilon$ the minimizer is a broken curve with the derivative

$$u'(x) = \begin{cases} v_0 & \text{if } x \in [x_0, x_0+\gamma\varepsilon] \\ v_0 & \text{if } x \in [x_0+\gamma\varepsilon, x_0+\varepsilon] \end{cases}$$

where $v_0 = \arg\min_z F(x,y,z)$, $1-\gamma \ll 1$, and v_1 is found from the equation

$$u'(x) \approx \frac{u(x+\varepsilon)-u(x)}{\varepsilon} = \frac{v_1\gamma\varepsilon + v_2(1-\gamma)\varepsilon}{\varepsilon}$$

to approximate the derivative u'. When $\gamma \to 1$, the contribution of the second interval becomes infinitesimal even if $v_2 \to \infty$.

The solution $u(x)$ can jump near the boundary point, therefore the main boundary conditions are irrelevant. The optimal trajectory will always satisfy natural boundary conditions that correspond to the minimum of the functional, and jump at the boundary points to meet the main conditions.

Example 7.4.1

$$F = \log^2(u+u') \quad u(0) = u(1) = 10$$

The minimizing sequence converges to a function from the family

$$u(x) = A\exp(-x) + 1 \quad x \in (0,1)$$

(A is any real number) and is discontinuous on the boundaries.

Problem with everywhere unbounded derivative This example shows an instructive minimizing sequence in a problem of sublinear growth. Consider the problem with the Lagrangian

$$J(u) = \int_0^1 F(x, u, u')dx, \quad F = (ax - u)^2 + \sqrt{|u'|}$$

This is an approximation problem: we approximate a linear function $f(x) = ax$ on the interval $[0, 1]$ by a function $u(x)$ using function $\sqrt{|u'|}$ as a penalty. We show that the minimizer is a distribution that perfectly approximate $f(x)$, is constant almost everywhere, and is nondifferentiable everywhere.

We mention two facts first: (i) The cost of the problem is nonnegative,

$$J(u) \geq 0 \quad \forall u,$$

and (ii) when the approximating function simply follows $f(x)$, $u_{trial} = ax$, the cost J of the problem is $J = \sqrt{a} > 0$ because of the penalty term.

Minimizing sequence Let us construct a minimizing sequence $u^k(x)$ with the property:

$$J(u^k) \to 0 \quad \text{if } s \to \infty$$

Partition the interval $[0, 1]$ into N equal subintervals and request that approximation $u(x)$ be equal to $f(x) = ax$ at the ends $x_k = \frac{k}{N}$ of the subintervals, and that the approximation is similar in all subintervals of partition,

$$\begin{aligned} u(x) &= u_0\left(x - \frac{k}{N}\right) + a\frac{k}{N} \quad \text{if } x \in \left[\frac{k}{N}, \frac{k+1}{N}\right], \\ u_0(0) &= 0, \quad u_0\left(\frac{1}{N}\right) = \frac{a}{N} \end{aligned}$$

Because of self-similarity, he cost J of the problem becomes

$$J = N \int_0^{\frac{1}{N}} \left((ax - u_0)^2 + \sqrt{|u_0'|}\right) dx \tag{7.40}$$

The minimizer $u_0(x)$ in a small interval $x \in \left[0, \frac{1}{N}\right]$ is constructed as follows

$$u_0(x) = \begin{cases} 0 & \text{if } x \in [0, \epsilon] \\ a\frac{1+\delta}{\delta}(x - \epsilon) & \text{if } x \in [\epsilon, \epsilon(1 + \delta)] \end{cases}$$

Here, ϵ and δ are two small positive parameters, linked by the condition $\epsilon(1 + \delta) = \frac{1}{N}$. The minimizer stays constant in the interval $x \in [0, \epsilon]$ and then linearly growths on the supplementary interval $x \in [\epsilon, \epsilon(1 + \delta)]$. We also check that

$$u_0\left(\frac{1}{N}\right) = u_0(\epsilon + \delta\epsilon) = \frac{a}{N}$$

Derivative $u_0'(x)$ equals

$$u_0'(x) = \begin{cases} 0 & \text{if } x \in [0, \epsilon] \\ a\frac{1+\delta}{\delta} & \text{if } x \in [\epsilon, \epsilon(1+\delta)] \end{cases}$$

Computing the functional (7.40) of the suggested function u_0,

$$J = N \left(\int_0^{\epsilon} ((ax)^2 dx + \int_{\epsilon}^{\epsilon+\delta} \left[\left(ax - a\frac{1+\delta}{\delta}(x-\epsilon) \right)^2 + \sqrt{a\frac{1+\delta}{\delta}} \right] dx \right)$$

we obtain, after obvious simplifications,

$$J = N \left(\frac{a^2 \epsilon^3}{3}(1+\delta) + \epsilon\sqrt{a(1+\delta)\delta} \right)$$

Excluding $\epsilon = \frac{1}{N(1+\delta)}$ we finally compute

$$J = \frac{a^2}{3N^2(1+\delta)^2} + \sqrt{\frac{a\delta}{1+\delta}}$$

Increasing N, $N \to \infty$ and decreasing δ, $\delta \to 0$ we can bring the cost functional arbitrary close to zero.

The minimizing sequence consists of the functions that are constant almost everywhere and contain a dense set of intervals of rapid growth. It tends to a nowhere differentiable function of the type of Cantor's "devils steps." The derivative is unbounded on a dense in $[0, 1]$ set. Because of slow growth of F,

$$\lim_{|u'| \to \infty} \frac{F(x, u, u')}{|u'|} \to 0$$

the functional is not sensitive to large values of u', if the growth occurs at the interval of infinitesimal measure. The last term of the Lagrangian does not contribute at all to the cost.

Regularization and relaxation To make the solution regular, we may go in two different directions. The first way is to forbid the wiggles by adding a penalization term $\epsilon(u' - a)^2$ to the Lagrangian which is transformed to:

$$F_{\epsilon} = (u - ax)^2 + \sqrt{|u'|} + \epsilon(u' - a)^2$$

The solution would become smooth, but the cost of the problem would significantly increase because the term $\sqrt{|u'|}$ contributes to it and the cost $J\epsilon = J(F_{\epsilon})$ would depend on ϵ and will rapidly grow to be close to $\sqrt{a}$. Until the cost grows to this value, the solution remain nonsmooth.

Alternatively, we may "relax" the problem, replacing it with another one that preserves its cost and has a classical solution that approximates our nonregular minimizing sequence. To perform the relaxation, we simply ignore the term $\sqrt{|u'|}$ and pass to the Lagrangian

$$F_{\text{relax}} = (u - ax)^2$$

which corresponds the same cost as the original problem and a classical solution $u_{\text{class}} = ax$ that in a sense approximate the true minimizer, but not its derivative; it is not differentiable at all.

7.5 Nonuniqueness and improper cost

Unbounded cost functional An often source of ill-posedness (the nonexistence of the minimizer) is the convergence to minimizing functional to $-\infty$ or the maximizing functional to $+\infty$. To illustrate this point, consider the opposite of the brachistochrone problem: Maximize the travel time between two points. Obviously, this time can be made arbitrary large by different means: For example, consider the trajectory that has a very small slop in the beginning and then rapidly goes down. The travel time in the first part of the trajectory can be made arbitrary large (Do the calculations!). Another possibility is to consider a very long trajectory that goes down and then up; the larger is the loop the more time is needed to path it. In both cases, the maximizing functional goes to infinity. The sequences of maximizing trajectories either tend to a discontinuous curve or is unbounded and diverges. The sequences do not convergence to a finite differentiable curve.

Generally, the problem with an improper cost does not correspond to a classical solution: a finite differentiable curve on a finite interval. Such problems have minimizing sequences that approach either non-smooth or unbounded curve or do not approach anything at all. One may either accept this "exotic solution," or assume additional constraints and reformulate the problem. In applications, the improper cost often means that something essential is missing in the formulation of the problem.

Nonuniqueness Another source of irregular solutions is nonuniqueness. If the problem has families of many extremal trajectories, the alternating of them can occur in infinitely many ways. The problem could possess either classical or nonclassical solution. To detect such problem, we investigate the Weierstrass-Erdman conditions which show the possibilities of broken extremals.

An example of nonuniqueness, nonconvex Lagrangian As a first example, consider the problem

$$I(v) = \min_u \int_0^1 \left(1 - (u')^2\right)^2 dx, \quad u(0) = 0, \ u(1) = v \tag{7.41}$$

The Euler equation admits the first integral, because the Lagrangian depends only on u',

$$\left(1 - (u')^2\right)(1 - 2u') = C;$$

the optimal slope is constant everywhere and is equal to V.

When $-1 \leq v \leq 1$, the constant C is zero and the value of I is zero as well. The solution is not unique. Indeed, in this case one can joint the initial and the final points by the curve with the slope equal to either one or negative one in all points. The Weierstrass-Erdman condition

$$\left[\left(1-(u')^2\right)(1-2u')\right]_-^+ = 0$$

is satisfied if $u' = \pm 1$ to the left and to the right of the point of break. There are infinitely many extremals with arbitrary number of breaks that all join the end points and minimize the functional making it equal to zero. Notice that Lagrangian is not convex function of u'.

Similarly to the finite-dimensional case, regularization of variational problems with nonunique solutions can be done by adding a penalty $\epsilon(u')^2$, or $\epsilon(u'')^2$ to the minimizer. Penalty would force the minimizer to prefer some trajectories. Particularly, the penalty term may force the solution to become infinitely oscillatory at a part of trajectory.

Another example of nonuniqueness, convex Lagrangian Work on the problem

$$I(v) = \min_u \int_0^1 (1-u')^2 \sin^2(mu)dx, \quad u(0) = 0,\ u(1) = v \tag{7.42}$$

As in the previous problem, here there are two kinds of "free passes" (the trajectories that correspond to zero Lagrangian that is always nonnegative): horizontal ($u = \pi k/m$, $u' = 0$) and inclined ($u = c+x$, $u' = 1$). The Weierstrass-Erdman condition

$$[\sin(mu)^2(1-u')]_-^+ = 0$$

allows to switch these trajectories in infinitely many ways.

Unlike the previous case, the number of possible switches is finite; it is controlled by parameter m. The optimal trajectory is monotonic; it becomes unique if $v \geq 1$ or $v \leq 0$, and if $|m| < \frac{1}{\pi}$.

7.6 Conclusion and Problems

We have observed the following:

- A one-dimensional variational problem has the fine-scale oscillatory minimizer if its Lagrangian $F(x, u, u')$ is a nonconvex function of its third argument.
- Homogenization leads to the relaxed form of the problem that has a classical solution and preserves the cost of the original problem.

- The relaxed problem is obtained by replacing the Lagrangian of the initial problem by its convex envelope. It can be computed as the second conjugate to F.

- The dependence of the Lagrangian on its third argument in the region of nonconvexity does not effect the relaxed problem.

To relax a variational problem we have used two ideas. First, we replaced the Lagrangian with its convex envelope and obtained a stable variational problem of the problem. Second, we proved that the cost of variational problem with the transformed Lagrangian is equal to the cost of the problem with the original Lagrangian if its argument $\boldsymbol{u}$ is a zigzag-like curve.

Problems

1. Formulate the Weierstrass test for the extremal problem

$$\min_u \int_0^1 F(x, u, u', u'')$$

that depends on the second derivative u''.

2. Find the relaxed formulation of the problem

$$\min_{u_1, u_2} \int_0^1 \left(u_1^2 + u_2^2 + F(u_1', u_2')\right),$$
$$u_1(0) = u_2(0) = 0, \quad u_1(1) = a, \quad u_2(1) = b,$$

where $F(v_1, v_2)$ is defined by (2.18). Formulate the Euler equations for the relaxed problems and find minimizing sequences.

3. Find the relaxed formulation of the problem

$$\min_u \int_0^1 \left(u^2 + \min\left\{|u' - 1|,\ |u' + 1| + 0.5\right\}\right),$$
$$u(0) = 0, \quad u(1) = a.$$

Formulate the Euler equation for the relaxed problems and find minimizing sequences.

4. Find the conjugate and second conjugate to the function

$$F(x) = \min\left\{x^2,\ 1 + ax^2\right\}, \quad 0 < a < 1.$$

Show that the second conjugate coincides with the convex envelope $\mathcal{C}F$ of F.

5. Let $x(t) > 0$, $y(t)$ be two scalar variables and $f(x, y) = x\,y^2$. Demonstrate that
$$f(\langle x\rangle, \langle y\rangle) \geq \langle y\rangle^2 \left\langle \frac{1}{x} \right\rangle^{-1}.$$
When is the equality sign achieved in this relation?

Hint: Examine the convexity of a function of two scalar arguments,
$$g(y, z) = \frac{y^2}{z}, \quad z > 0.$$

Chapter 8

Hamiltonian, Invariants, and Duality

In this chapter, we return to study of the Euler equations transforming them to different forms. The variational problem is viewed here as the convenient and compact form to generate these equations. We will also focus on invariant properties of solutions that can be obtained from variational formulations. We will see that the stationary conditions in classical mechanics usually do not lead to a true minimizer but are adequate to describe the motion and equilibria of a mechanical system.

8.1 Hamiltonian

8.1.1 Canonic form

The structure of Euler equations can be simplified and unified if we consider $2N$ first-order differential equations instead of N second-order ones

$$\frac{d}{dx}\frac{\partial L}{\partial u_i'} - \frac{\partial L}{\partial u_i} = 0.$$

A first-order system can be obtained from the above equations if we introduce a new variable p,

$$p(x) = \frac{\partial L(x,u,u')}{\partial u'} \tag{8.1}$$

In mechanics, p is called the impulse. The Euler equation takes the form

$$p' = \frac{\partial L(x,u,u')}{\partial u} = n(x,u,u'), \tag{8.2}$$

where n is function of x, u, u'.

The system (8.1), (8.2) becomes symmetric with respect to p and u if we algebraically solve (8.1) for u' as follows:

$$u' = \phi(x, u, p), \tag{8.3}$$

and substitute this expression (8.2):

$$p' = n(x, u, \phi(x, u, p)) = \psi(x, u, p) \tag{8.4}$$

where ψ is a function of the variables u and p but not of their derivatives.

In summary, system (8.1), (8.2) is transferred to the canonic from (or Cauchy form)

$$\begin{array}{l} u' = \phi(x, u, p) \\ p' = \psi(x, u, p) \end{array} : \tag{8.5}$$

It is resolved for the derivatives u' and p' and is symmetric with respect to variables u and p. The properties of the solution are entirely determined by the algebraic vector functions ϕ, ψ in the right-hand side, which do not contain derivatives.

Remark 8.1.1 The equation (8.1) can be solved for u' and (8.3) can be obtained if the Lagrangian is convex function of u' of a superlinear growth. As we know, (see Chapter (7)), we expec this condition to be satisfied if the problem has a classical minimizer.

Impulses The equations of Lagrangian mechanics correspond to the stationarity of the action

$$L = \frac{1}{2}\dot{q}^T R(q)\dot{q} - V(q)$$

Variables $p = \frac{\partial L}{\partial \dot{q}}$ are called *impulses* are are equal to $p = R(q)\dot{q}$. The canonic system becomes

$$\dot{q} = R^{-1}p, \quad \dot{p} = \frac{\partial L}{\partial q} = p^T R^{-1} \frac{\partial R}{\partial q} R^{-1} p - \frac{\partial V}{\partial q}$$

The last equation is obtained by excluding $\dot{q}$ from the $\frac{\partial L}{\partial q}$.

Example 8.1.1 (Quadratic Lagrangian) Assume that

$$L = \frac{1}{2}(a(x)u'^2 + b(x)u^2).$$

We introduce p as in (8.1)

$$p = \frac{\partial L(x, u, u')}{\partial u'} = au'$$

and obtain the canonic system

$$u' = \frac{1}{a(x)}p, \quad p' = b(x)u.$$

Notice that the coefficient $a(x)$ is moved into denominator.

Next example deals with the pendulum problem in Example 13.21

Example 8.1.2 (Two-link pendulum)

8.1.2 Hamiltonian

Although the system (8.5) is a convenient first-order system to deal with, we may rewrite it in a more symmetric form introducing a special function called *Hamiltonian.*

The Hamiltonian is defined by the formula:

$$H(x,u,p) = pu'(x,u,p) - L(x,u,u'(u,p)) = p\phi(x,u,p) - L(x,u,\phi(x,u,p))$$

where u is a stationary trajectory – the solution of Euler equation. Let us compute the partial derivatives:

$$\frac{\partial H}{\partial u} = p\frac{\partial \phi}{\partial u} - \frac{\partial L}{\partial u} - \frac{\partial L}{\partial \phi}\frac{\partial \phi}{\partial u}$$

But, by the definition of p, $p = \frac{\partial L}{\partial u'} = \frac{\partial L}{\partial \phi}$ the first and third term in the right-hand side cancel. By virtue of the Euler equation, the remaining term $\frac{\partial L}{\partial u}$ is equal to p' and we obtain

$$p' = -\frac{\partial H}{\partial u} \tag{8.6}$$

Next, compute $\frac{\partial H}{\partial p}$. We have

$$\frac{\partial H}{\partial p} = p\frac{\partial \phi}{\partial p} + \phi - \frac{\partial L}{\partial \phi}\frac{\partial \phi}{\partial p}$$

By definition of p, the first and the third term in the right-hand side cancel, and by definition of ϕ ($\phi = u'$) we have

$$u' = \frac{\partial H}{\partial p} \tag{8.7}$$

The system (8.6), (8.7) is called the canonic system, it is remarkable symmetric.

In Lagrangian mechanics, the Hamiltonian H is equal to the sum of kinetic and potential energy, $H = T + V$ where $\dot{q} = R(q)p$ is excluded,

$$H(q,p) = \frac{1}{2}p^T R^{-1} p + \frac{\partial}{\partial q}(p^T R^{-1} p + V)$$

Here, we use the we already established property $\frac{\partial T}{\partial \cdot q} = 2T$ of kinetic energy – a homogeneous second degree function of $\dot{q}$.

Example 8.1.3 Compute the Hamiltonian and canonic equations for the system in the previous example.

We have

$$L = \frac{1}{2}(a(x)u'^2 + b(x)u^2) = \frac{1}{2}\left(\frac{1}{a(x)}p^2 + b(x)u^2\right)$$

then

$$H = p\left(\frac{p}{a}\right) - L = \frac{1}{2}\left(\frac{1}{a(x)}p^2 - b(x)u^2\right)$$

and

$$\frac{\partial H}{\partial u} = -b(x)u = -p', \quad \frac{\partial H}{\partial p} = \frac{1}{a(x)}p = u'$$

which coincides with the previous example.

Natural boundary conditions and end-point conditions Using Hamiltonian, we conveniently reformulate the variations of the boundary condition and the length of the interval of integration. Natural boundary conditions (13.21) become $\frac{\partial L}{\partial u'} = p = 0$, and the condition (13.21) for the interval of unknown length becomes $L - u'\frac{\partial L}{\partial u'} = H = 0$. The first variation takes the form

$$\delta I = p_i \delta u_i + H\delta x|_0^T + \int_0^T \left(p' - \frac{\partial H}{\partial u}\right)\delta u \, dx \tag{8.8}$$

The Weierstrass-Erdman conditions at the moving boundary are rewritten as

$$[p \cdot \delta u + H\delta x]_-^+ = 0 \tag{8.9}$$

If both variations δu and δx are free, we obtain the condition for a broken extremal $[p]_-^+ = 0$ and $[H]_-^+ = 0$.

8.1.3 Geometric optics

The results of study of geometic optics (Section13.21) can be conveniently presented using Hamiltonian. It is convenient to introduce the slowness $w(x,y) = \frac{1}{v(x,y)}$ - reciprocal to the speed v. Then the Lagrangian for the geometric optic problem is

$$L(x, y, y') = w\sqrt{1 + (y')^2} \quad y' > 0.$$

Canonic system To find a canonic system, we use the outlined procedure: Define a variable p dual to $y(x)$ by the relation $p = \frac{\partial L}{\partial y'}$

$$p = \frac{wy'}{\sqrt{1 + (y')^2}}.$$

Solving for y', we obtain first canonic equation:

$$y' = \frac{p}{\sqrt{w^2 - p^2}} = \phi(x, y, p), \tag{8.10}$$

Excluding y' from the expression for L,

$$L(x, y, \phi) = L_*(w(x, y), p) = \frac{w^2}{\sqrt{w^2 - p^2}}.$$

and recalling the representation for the solution y of the Euler equation

$$p' = \frac{\partial L}{\partial y} = \frac{\partial L_*}{\partial w} \frac{dw}{dy}$$

we obtain the second canonic equation:

$$p' = -\frac{w}{\sqrt{w^2 - p^2}} \frac{dw}{dy} \tag{8.11}$$

Hamiltonian Hamiltonian $H = p\phi - L_*(x, y, p)$ can be simplified to the form

$$H = -\sqrt{w^2 - p^2}$$

It satisfies the remarkably symmetric relation

$$H^2 + p^2 = w^2$$

that contains the whole information about the geometric optic problem. The elegancy of this relation should be compared with messy straightforward calculations that we previously did. The geometric sense of the last formula becomes clear if we denote as α the angle of declination of the optimal trajectory to OX axis; then $y' = \tan\alpha$, and (see (13.21))

$$p = \psi(x, y) \sin\alpha, \quad H = -\psi(x, y) \cos\alpha.$$

Refraction: Snell's law Assume the media has piece-wise constant properties, speed $v = 1/\psi$ is piece-wise constant $v = v_1$ in Ω_1 and $v = v_2$ in Ω_2; denote the curve where the speed changes its value by $y = z(x)$. Let us derive the refraction law. The variations of the extremal $y(x)$ on the boundary $z(x)$ can be expressed through the angle θ to the normal to this curve

$$\delta x = \sin\theta, \quad \delta y = \cos\theta$$

Substitute the obtain expressions into the Weierstrass-Erdman condition (8.9) and obtain the refraction law

$$[\psi(\sin\alpha\cos\theta - \cos\alpha\sin\theta)]_-^+ = [\psi\sin(\alpha - \theta)]_-^+ = 0$$

Finally, , recall that $\psi = \frac{1}{v}$ and rewrite it in the conventional form (Snell's law)

$$\frac{v_1}{v_2} = \frac{\sin\gamma_1}{\sin\gamma_2}$$

where $\gamma_1 = \alpha_1 - \theta$ and $\gamma_2 = \alpha_2 - \theta$ are the angles between the normal to the surface of division and the incoming ?? and the refracted rays respectively.

Weierstrass-Erdman condition Although the classical derivation of the Euler equation required the existence of second derivative of u, the system (8.5) is different: The functions ϕ and ψ do not need to be even continuous functions of x. However, the variables p and u are to be differentiable to satisfy (8.5).

The Weierstrass-Erdman condition, see Section 3.2.1, expresses the continuity of

$$p = \frac{\partial F}{\partial u'}$$

along the optimal trajectory. In each point, the jump of p is zero,

$$\left[\frac{\partial F}{\partial u'}\right]_-^+ = 0, \quad \text{along the optimal trajectory } u(x)$$

Example 8.1.4 (Quadratic Lagrangian, continuity) In the previous Example 8.1.3 we may assume that the coefficient $a = a(x)$ is discontinuous and it switches from a_- to a_+ at the point x_0. Applying the Weierstrass-Erdman condition, we find that

$$a_- u'_- = a_+ u'_+$$

at the point x_0. This shows that the extremal breaks at this point; it is called the *broken extremal*.

8.1.4 Fast oscillating coefficients. Homogenization

The mentioned continuity of the canonic variables u and p allows for easy handling of system with fast oscillating coefficients.

Consider again the Lagrangian

$$F = \frac{1}{2}(a(x)u'^2 + b(x)u^2).$$

and assume that $a(x) > 0$ and $b(x)$ are rapidly oscillating functions of x. Accordingly, the solution $u(x)$ is also an oscillating function. Dealing with such problems, it is desirable to find a variational formulation of the averaged Lagrangian. This approach is called *homogenization*. The averaged variables are denoted by a subindex $_\epsilon$. They are defined as follows:

$$z_\epsilon(x) = \frac{1}{2\epsilon}\int_{x-\epsilon}^{x+\epsilon} z(\xi)d\xi$$

Let us average the equations (8.5) over an interval of x that is small comparing with $b - a$ but large comparing with a scale of oscillations.

Averaged Lagrangian is

$$[F]_\epsilon = \frac{1}{2}\left([a(x)u'^2]_\epsilon + [b(x)u^2]_\epsilon\right).$$

This form, however, is not convenient since it is not clear how to compute the average derivative $[u']_\epsilon$. The derivative is not a smooth or even continuous function and the term $[a(x)u'^2]_\epsilon$ is a product of two oscillatory terms. Therefore, we pass to the canonic variables p and u that are both differentiable, and their derivatives are bounded. Therefore, we may use the continuity of u and p and consider them as constants on the interval of averaging. If $\epsilon \ll 1$, we may assume that all differentiable variables are close to their average, in particular,

$$u_\epsilon(x) = u(x) + O(\epsilon), \quad p_\epsilon(x) = p(x) + O(\epsilon).$$

We compute, as before:

$$p = \frac{u'}{a}, \quad L = \frac{1}{2a}p^2 + \frac{b}{2}u^2$$

In terms of canonic variables, the average Lagrangian becomes

$$[F]_\epsilon = \frac{1}{2}\left(\left[\frac{1}{a(x)}\right]_\epsilon p^2 + [b(x)]_\epsilon u^2\right).$$

Here we use the continuity of u and p to compute averages:

$$\left[\frac{p^2}{a(x)}\right]_\epsilon = \left[\frac{1}{a(x)}\right]_\epsilon p^2 + O(\epsilon), \quad [b(x)u^2]_\epsilon = [b(x)]_\epsilon u^2 + O(\epsilon)$$

Returning to the original notations, we find that

$$[u']_\epsilon = \left[\frac{1}{a(x)}\right]_\epsilon^{-1} p_\epsilon$$

and we obtain the homogenized Lagrangian

$$L(u_\epsilon, u'_\epsilon) = \frac{1}{2}\left[\frac{1}{a(x)}\right]_\epsilon^{-1} (u')^2 + [b(x)]_\epsilon u^2$$

We arrive at interesting results: the oscillating coefficients a and b are replaced by their *harmonic* and *arithmetic* means, respectively, in the homogenized system.

Let us find the equation for the extremal. The averaged (homogenenized) Hamiltonian is

$$[H]_\epsilon = \frac{1}{2}\left(\left[\frac{1}{a(x)}\right]_\epsilon^{-1} p^2 - [b(x)]_\epsilon u^2\right).$$

The canonic system for the averaged canonic variables u_ϵ, p_ϵ becomes

$$u'_\epsilon = \left[\frac{1}{a(x)}\right]_\epsilon^{-1} p_\epsilon \quad p'_\epsilon = [b(x)]_\epsilon u_\epsilon \tag{8.12}$$

Example 8.1.5 Let us specify the oscillating coefficients $a(x)$ and $b(x)$ as follows

$$a(x) = \alpha_1 + \alpha_2 \sin^2\left(\frac{x}{\epsilon}\right), \quad b(x) = \beta_1 + \beta_2 \sin^2\left(\frac{x}{\epsilon}\right)$$

where $\alpha_1 > 0,\ a_2 > 0$. The homogenized coefficients are computed (with the help of Maple) as:

$$a_\epsilon = \left(\frac{1}{T}\int_0^T \frac{1}{\alpha_1 + \alpha_2 \sin^2\left(\frac{x}{\epsilon}\right)} dx\right)^{-1}, \quad \lim_{\epsilon\to 0} a_\epsilon = \alpha_1\sqrt{1+\frac{\alpha_2}{\alpha_1}};$$

$$b_\epsilon = \frac{1}{T}\int_0^T \left[\beta_1 + \beta_2 \sin^2\left(\frac{x}{\epsilon}\right)\right] dx, \quad \lim_{\epsilon\to 0} b_\epsilon = \beta_1.$$

We observe that the average coefficients nonlinearly depend on the magnitude α_2 of oscillations of the $a(x)$, but not on the magnitude β_2. The homogenized problem corresponds to Hamiltonian

$$H = \frac{1}{2}\left(\alpha_1\sqrt{1+\frac{\alpha_2}{\alpha_1}}\right)p^2 - \frac{1}{2}\beta_1 u^2.$$

Derive equation of the stationary trajectory.

8.2 Symmetries and invariants

We discuss here methods of systematically determination of invariants of solutions to variational problems. We want to find quantities that stay constant along the trajectory – solution to the Euler equation.

8.2.1 Poisson brackets

There is an algebraic construction for checking whether a function $G(x,u,p)$ is constant at a stationary trajectory based on the so-called Poisson brackets. Compute the whole differential of G:

$$\frac{dG}{dx} = \frac{\partial G}{\partial x} + \frac{\partial G}{\partial u}\cdot u' + \frac{\partial G}{\partial p}\cdot p'$$

The derivatives u' and p' at the stationary trajectory can be expressed through Hamiltonian (see (**??**)), $u' = \frac{\partial H}{\partial p}, p' = -\frac{\partial H}{\partial u}$; consequently, we rewrite the expression for the derivative of G,

$$\frac{dG}{dx} = \frac{\partial G}{\partial x} + [G,H] \tag{8.13}$$

where $[G,H]$ are the Poisson brackets:

$$[G,H] = \frac{\partial G}{\partial p}\cdot\frac{\partial H}{\partial u} - \frac{\partial G}{\partial u}\cdot\frac{\partial H}{\partial p} \tag{8.14}$$

or, in coordinates,

$$[G,H] = \sum_{i=1}^{n}\left(\frac{\partial G}{\partial p_i}\frac{\partial H}{\partial u_i} - \frac{\partial G}{\partial u_i}\frac{\partial H}{\partial p_i}\right)$$

The function G stays constant at the trajectory if the right-hand side of (8.13) is zero. To clarify the use of the Poisson brackets, we derive the already discussed first integrals by their means.

Example 8.2.1 (Time-independent Hamiltonian) Assume that H does not explicitly depend on x: $H = H(u,p)$ or $\frac{\partial H}{\partial x} = 0$. Then H is constant along the trajectory. Indeed, we set $G = H$ and compute (8.13):

$$\frac{dH}{dx} = [H,H] = 0 \quad \Rightarrow \ H = \ \mathrm{constant}(t)$$

The equality $[H,H] = 0$ immediately follows from the definition (8.14) of the Poisson brackets.

Example 8.2.2 (Conservation of impulse) Assume that H is independent of u_k: $\frac{\partial H}{\partial u_k} = 0$. Then p_k is constant

$$p_k = \ \mathrm{constant} \tag{8.15}$$

To prove, set $G = p_k$ and compute (8.13) using the definition (8.14)

$$\frac{dp_k}{dx} = [p_k, H] = \frac{\partial H}{\partial u_k}$$

Since by assumption $\frac{\partial H}{\partial u_k} = 0$, the result (8.15) follows.

Example 8.2.3 (Conservation of coordinate) Similarly, if H is independent of the impulse p_k: $\frac{\partial H}{\partial p_k} = 0$, the coordinate u_k is constant along the extremal trajectory:

$$u_k = \ \mathrm{constant}(t)$$

Again, set $G = u_k$ and compute (8.13)

$$\frac{du_k}{dx} = [u_k, H] = -\frac{\partial H}{\partial p_k} = 0$$

The technique does not tell how to guess the quantity G from a special form of Hamiltonian but provides a method to check a guess.

8.2.2 Nöther's Theorem

Nöther's Theorem proves a relationship between symmetries and conservation principles: "Every symmetry gives a conserved quantity." Assume, for example, that motion of a system of particles is described by a Lagrangian $L(q, q')$ that depends on distances between the particles but is independent of their absolute locations and of time. According to Nöther's theorem, a quantity must be conserved; here, the conserved quantities are the whole energy, the main moment and the main angular moment. The theorem, proved in 1915 by Emmy Nöther, was praised by Einstein as a piece of "penetrating mathematical thinking."

Transformation Suppose that the system is invariant to a transformation: Lagrangian $L(x,u,u')$ doesn't change its value under some family of transformations that sends u and x to some new positions $\hat{u}$ and $\hat{x}$:

$$\hat{x} = \Phi(x,u,\alpha) = x + \alpha\phi(x,u) + o(\alpha^2) \tag{8.16}$$

$$\hat{u} = \Psi(x,u,\alpha) = x + \alpha\psi(x,u) + o(\alpha^2) \tag{8.17}$$

where α is a parameter of the transformation. It is assumed for definiteness that the transformation is identical (nothing is transformed) when $\alpha = 0$ and that it smoothly depend on the parameter α of transformation. For example, the independence of time can be viewed as the invariance to the transformation $\hat{t} = t + \alpha$, the independence of shift – as invariance to the transformation $\hat{x} = x + \alpha$, the independence of rotation – as invariance to the transformation $\hat{x} = R(\alpha)x$, where R is the matrix of rotation, and vector α are is composed from angles of rotation.

Lorentz transform preserves the quantity $x^2 - c^2t^2$ and is invariant to the transformation

$$\hat{x} = \cosh(\alpha)x + c\sinh(\alpha)t$$
$$\hat{t} = -\sinh(\alpha)x + c\cosh(\alpha)t$$

Theorem 8.2.1 (Nöther) If the system with Hamiltonian H is invariant to the transforms (8.16), (8.17), than the quantity W is conserved.

$$W = p \cdot \psi - H\phi = \text{ constant} \tag{8.18}$$

Proof We observe that $\delta x = \phi da$, and $\delta u = \psi \cdot da$. Substituting these expression into the formula for the first variation (8.8) we obtain (8.18) because da is an arbitrary number.

Example 8.2.4 If Hamiltonian is independent of x, it is invariant to translation (13.21); in this case $\phi = 1, \psi = 0$. By (8.18), we compute $H = \text{ constant}$.

Example 8.2.5 The invariance to the shift (13.21) is expressed as in this case $\phi = 0, \psi_i = 1$. By (8.18), we compute $\sum p_i = \text{ constant}$.

Another proof for time-independent symmetries Assume that s is independent of x. Compute the x derivative of W by the chain rule:

$$W' = p'\frac{du}{d\alpha} + p\frac{du'}{d\alpha}$$

Recalling the definitions $p' = \frac{\partial L}{\partial u}, \quad p = \frac{\partial L}{\partial u'}$ we obtain the result:

$$W' = \frac{\partial L}{\partial u}\frac{dq}{ds} + \frac{\partial L}{\partial q'}\frac{dq'}{ds} = \frac{d}{ds}L$$

Since L is independent of s by assumption, $\frac{d}{ds}L = 0$, then $W' = 0$ and W is constant.

> **Related anecdote**
>
> In 1915, Emma Nöther arrived in Göttingen but was denied the private-docent status. The argument was that a woman cannot attend the University senate (essentially, the faculty meetings). Hilbert's reaction was: "Gentlemen! There is nothing wrong to have a woman in the senate. Senate is not a bath."

First integrals of the double pendulum What is conserved? The time shift. The shift in space is not applicable because the first pendulum is fasten. If we allow it to move along the OX axis, the shift in x-direction is invariant and the corresponding first integral becomes

$$(0,1)(p_1 + p_2) = \text{ constant}$$

Also, if the gravity is neglected (the motion is only due to the inertia) then the system is invariant to rotation around the hitch and the corresponding angular momentum is preserved.

System of particles with central forces Consider a system of N particles with the forces between them directed along the line between particles and had a magnitude that depends only of the distance between particles

The Hamiltonian is

$$H = \frac{1}{2}\sum \frac{p_i^2}{m_i} + \sum_{i,j} \phi(|r_i - r_j|)$$

Euler equations are

$$m_i \ddot{r} = f_{ij}, \quad f_{ij} = \sum_j \left(\frac{\phi'}{|r_i - r_j|} \right) (r_i - r_j)$$

The motion is the system is invariant to the shifts in time and space, and to the rotation of the whole system. Let us find first integrals of this system.

The invariance to the time shift implies the constancy of the Hamiltonian

$$H = \text{ constant}$$

The invariance to the space shift $\cap r_i = r_i + \alpha$ implies the conservation of the total impulse (three scalar equations)

$$\sum p_i = \text{ constant}$$

The invariance to the space rotation $\cap r_i = \alpha\Omega \times r_i$ implies the conservation of the total angular momentum (three scalar equations)

$$\sum \omega \times p_i = \text{ constant}$$

The preservation of these integral allows for viewing the particle system as a simpler object.

8.2.3 Kepler's laws in Celestial mechanics

The studied variational principles enable to obtain Kepler's laws in celestial mechanics. In cylindrical coordinates chosen so that the position of the Sun in in the origin and the position of the planet and the vector of the planet's speed are in the plane $z = 0$, the kinetic energy T of the planet is

$$T = \frac{m}{2}\left(\dot{r}^2 + \frac{\dot{\theta}^2}{r^2}\right)$$

and the potential energy V is defined by the Newton's law

$$V = -\frac{1}{2}\frac{mM}{\gamma}\frac{1}{r}$$

The Lagrangian L depends on two unknown functions, $r(t)$ and $\theta(t)$:

$$L(r, \theta) = T - V = \frac{m}{2}\left(\dot{r}^2 + \dot{\theta}^2 r^2 - \frac{M}{\gamma}\frac{1}{r}\right)$$

Two first integrals are: The rule of the areas:

$$\frac{\partial L}{\partial \dot{\theta}} = m\dot{\theta}r^2 = C_1$$

(because the Lagrangian does not depend on θ) and conservartion of energy

$$T + V = \frac{m}{2}\left(\dot{r}^2 + \dot{\theta}^2 r^2 + \frac{M}{\gamma}\frac{1}{r}\right) = C_2$$

(because the Lagrangian does not depend on t)

From the first we obtain the conservation of the angular momentum (the motion is in that plane all the time).

Then we have

$$\ddot{r} - \dot{\theta}^2 r - \frac{M}{\gamma r^2} = 0$$

and, ()

Denote $u = \frac{1}{r}$, $u' = -\frac{r'}{r^2}$ and Find that

$$\frac{d^2u}{d\theta^2} + u = C_5$$

and find the equation for the trajectory $r(\theta)$:

$$r = \frac{A}{1 + \varepsilon \cos\theta}$$

or, in Cartesian coordinates,

$$(1 - p)x^2 + ..x + y^2 = p$$

ellipse.

Finally, let us apply Nöther's theory to the Hamiltonian

$$H =$$

the invariants

the groups of transformations: $\bar{t} = t + c$ and $\bar{\theta} = \theta + c$

8.2.4 General case of central forces: Invariants

$$L = m_i|\dot{r}_i|^2 - \sum_j \gamma m_i m_j V(|r_i - r_j|)$$

we compute $p_i = m_i|\dot{r}_i|$ and

$$H = \frac{|p_i|^2}{m_i} + \sum_j \gamma m_i m_j V(|r_i - r_j|)$$

Conservation of energy H is independent of time, therefore,

$$H = \text{constant} \quad \forall t$$

The Hamiltonian is equal (see (13.21)) to the whole energy $T+V$ which remains constant.

Conservation of momentum The shift of all positions by the same vector $\bar{r_i} = r_i + a$ keeps the Lagrangian invariant; therefore, the whole moment

$$P = p_1 + \dots p_N = \sum m_i \dot{r}_i = \text{ constant}$$

is constant.

Conservation of angular momentum The transform $\bar{r_i} = r_i + (r_i - r_0) \times \omega -$ the rotation across the center of mass $r_0 = m_i r_r$ keeps the Lagrangian invariant; therefore, the whole angular moment

$$PA = \sum m_i \dot{r}_i \times = \text{ constant}$$

is constant.

see Gelfand

8.2.5 Lorentz transform and invariants of relativity

8.3 Duality

8.3.1 Duality as solution of a constrained problem

The variational problem

$$I = \min_u \int_0^1 F(x, u, u')dx \tag{8.19}$$

can be rewritten as the minimization of the constrained problem

$$I = \min_{u,v} \int_0^1 F(x,u,v)dx \quad \text{subject to } u' = v$$

where the constraint specifies the differential dependence between arguments of the Lagrangian. The last problem is naturally rewritten using Lagrange multiplier $p(x)$:

$$I = \min_{u,v} \max_{p} \int_0^1 \left[F(x,u,v) + p(u' - v) \right] dx \tag{8.20}$$

Let us analyze this problem. First, we integrate by parts the term $p\,u'$ and interchange the sequence of extremal operations using the min-max theorem **??**

$$\min_x \max_y f(x,y) \geq \max_y \min_x f(x,y). \tag{8.21}$$

We obtain the inequality:

$$I \geq I_D \tag{8.22}$$

where the functional I_D is

$$I_D = \max_p \int_0^1 F_D(x,p,p')dx + p\,u|_0^1 \tag{8.23}$$

and

$$F_D(x,p,p') = \min_{u,v} \left[F(x,u,v) - p'u - p\,v \right] \tag{8.24}$$

Notice that F_D depends on u and v but not on their derivatives; therefore the variation with respect to these is performed independently in the each point of the trajectory (under the integral):

The first variation of I_D with respect to u and v is zero,

$$\delta I_D = \int_0^1 \left[\left(\frac{\partial F}{\partial v} - p \right) \delta v + \left(\frac{\partial F}{\partial u} - p' \right) \delta u \right] dx = 0.$$

The coefficients by variations δu and δv vanish which gives the stationarity conditions

$$p = \frac{\partial F}{\partial v}, \quad p' = \frac{\partial F}{\partial u}. \tag{8.25}$$

Now, we may transform the problem in three different ways.

1. Excluding p and p' from (8.25), we obtain the conventional Euler equation:

$$\frac{d}{dx}\frac{\partial F}{\partial v} - \frac{\partial F}{\partial u} = 0, \quad u' = v$$

2. Excluding u and v from (8.25): $u = \phi(p, p')$, $v = \psi(p, p')$, we obtain the dual variational problem

$$I \geq I_D; \quad I_D = \max_p \int_0^1 [L_D(x, p, p')]\, dx + p\,\phi|_0^1 \tag{8.26}$$

where

$$L_D(p, p') = F(x, \phi, \psi) - p'\phi - p\psi$$

The dual problem depends on *dual variable* p and its derivative p' instead of u and u'. If the Lagrangian F is convex, the minimax theorem delivers to equality sign in (8.22) and both primary an dual problem have the same cost.

3. Excluding v and p' from (8.25) as follows: $v = \rho(u, p)$ we obtain the Hamiltonian H

$$H(u, p) = L(x, u, \rho(u, p)) - \rho(u, p)p;$$

the Hamiltonian is independent of the derivatives of the arguments; the variational problem becomes

$$I^H = \min_u \max_p \int_0^1 [H(u, p) + u'p]\, dx$$

Example 8.3.1 (Quadratic Lagrangian) Find a conjugate to the Lagrangian

$$F(u, u') = \frac{1}{2}\sigma(u')^2 + \frac{\gamma}{2}u^2. \tag{8.27}$$

Rewrite the Lagrangian using the Lagrange multiplier (impulse) to account for the differential constraint,

$$\bar{F} = \frac{1}{2}\sigma v^2 + \frac{1}{2}\gamma u^2 + p(u' - v).$$

1. The impulse p is

$$p = \frac{\partial F}{\partial v} = \sigma v.$$

Derivative $u' = v$ is expressed through p as

$$u' = \frac{p}{\sigma}.$$

2. The Hamiltonian H is

$$H = \frac{1}{2}\frac{p^2}{\sigma} - \gamma u^2.$$

The canonical system is

$$u' = \frac{p}{\sigma}, \quad p' = \gamma u,$$

3. The dual form F^* of the Lagrangian is obtained from the Hamiltonian using canonical equations to exclude u, as follows: $u = \frac{p'}{\gamma}$; substituting this into the expression for the Hamiltonian, we obtain

$$F^*(p, p') = \frac{1}{2}\left(\frac{p^2}{\sigma} - \frac{1}{\gamma}(p')^2\right).$$

The Legendre transform is an involution: The variable dual to the variable $\boldsymbol{p}$ is equal to $\boldsymbol{u}$.

8.3.2 Legendre and Young-Fenchel transforms

Duality in calculus of variation is closely related to the duality in the theory of convex function; both use the same algebraic means to pass to the dual representation. Here we review the Legendre and Young-Fenchel transforms that serve to compute the dual Lagrangian. Namely, the inner minimization problem (8.24) in the problem (8.22) is an algebraic one

$$F_D(x, p, p') = \min_z (F(z) - z \cdot q) \quad z = (u, u'),\ q = (p, p') \tag{8.28}$$

Here, z and q are two vector arguments that in the variational problem represent the minimizer and its derivative. We view arguments p and p' as independent variables and the problem (8.28) as a special algebraic transform. This transform is studied in convex analysis and it is called Young-Fenchel transform. If F is convex and differentiable everywhere, the transform is called Legendre transform.

Definition 8.3.1 Function $L^*(\boldsymbol{z}^*)$–the *conjugate* to the $L(z)$– is defined by the relation

$$L^*(\boldsymbol{z}^*) = \max_{\boldsymbol{z}} \{\boldsymbol{z}^*\, \boldsymbol{z} - L(\boldsymbol{z})\}, \tag{8.29}$$

The definition implies that $\boldsymbol{z}^*$ is an analog of p (compare with (**??**)).

Geometric interpretation Consider graph of a convex function $y = f(x)$ of a scalar argument x. Assume that a straight line $x^*x + b$ touches the graph approaching it from below moving up (that is, increasing b. When the line touches the graph, we register the tangent x^* of its angle and the coordinate b of the intersection of the line with the axes OY. Then we change the angle and repeat the experiment for all angles that is for all $x^* \in R$. Clearly, any convex curve can be found if we know the set of all such curves. This curve is simply an envelope of the family of straight lines.

The relation $f^*(x^*) = -b(x^*)$ between negative of b and x^* is called the dual or Young-Fenchel transform of the original function $f(x)$. If $f(x)$ is a convex function it can be recovered back from its Young-Fenchel transform.

The multidimensional case is treated in the same way: a hyperplane $b = \sum_i x_i^* x_i$ is used instead of the straight line and the transform is given by the relation (8.29) where $\boldsymbol{z}$ is a vector $x_1, ..x_n$.

Legendre transform Let us compute the conjugate to the Lagrangian $L(x, \boldsymbol{y}, \boldsymbol{z})$ with respect to $\boldsymbol{z}$, treating $x, \boldsymbol{y}$ as parameters. If L is a convex and differentiable function of $\boldsymbol{z}$, then (8.29) is satisfied if

$$z^* = \frac{\partial L(\boldsymbol{z})}{\partial \boldsymbol{z}}. \tag{8.30}$$

This formula is exactly the transform from the Lagrangian that depends on $z = u'$ to the Hamiltonian which depends on the impulse p.

The similarity suggests that the Legendre transform $u' \to \boldsymbol{p}$ and the Young–Fenchel transform $\boldsymbol{z}^*$ coincide if the Legendre transform is applicable, that is if L is a convex and differentiable function of u'.

Example 8.3.2 (A conjugate to a quadratic)

$$F(x) = \frac{1}{2}(x-a)^2 \tag{8.31}$$

We have

$$F^*(x^*) = \frac{1}{2}(|x^*| + a)^2 - \frac{1}{2}a^2 \tag{8.32}$$

In particular, $F(x) = \frac{1}{2}x^2$ is stable to the transform:

$$F^*(x^*) = \frac{1}{2}(x^*)^2$$

The Young–Fenchel transform is well defined and finite for a larger class of non-differentiable functions, namely, for any Lagrangian that grows not slower than an affine function:

$$L(\boldsymbol{z}) \geq c_1 + c_2\|\boldsymbol{z}\| \quad \forall \boldsymbol{z}, \tag{8.33}$$

where c_1 and $c_2 > 0$ are constants.

Example 8.3.3 (A conjugate to a function with discontinuous derivative) Consider

$$F(x) = \exp(|x|). \tag{8.34}$$

From (8.29) we have

$$F^*(x^*) = \begin{cases} (|x^*|(\log|x^*| - 1) & \text{if} \quad |x^*| > 1, \\ 0 & \text{if} \quad |x^*| > 1. \end{cases} \tag{8.35}$$

Example 8.3.4 (Additional example)

$$F(x) = \begin{cases} \frac{1}{2}(|x| - a)^2 & \text{if} \quad |x| \geq a \\ 0 & \text{if} \quad |x| \leq a \end{cases}. \tag{8.36}$$

The conjugate is ....

A more telling example involves a function that growth linearly and is discontinuous.

Example 8.3.5 (A conjugate to $|x|$) Consider

$$F(x) = |x|. \tag{8.37}$$

From (8.29) we have

$$F^*(x^*) = \begin{cases} 0 & \text{if} \quad |x^*| < 1, \\ \infty & \text{if} \quad |x^*| > 1. \end{cases} \tag{8.38}$$

Finally, consider the function with sublinear growth.

Example 8.3.6 (A conjugate to $\sqrt{|x|}$)

$$F(x) = \sqrt{|x|} \tag{8.39}$$

The function is not convex and the Legendre transform does not exist. The Young-Fenchel transform gives

$$F^*(x^*) = 0 \quad \forall x^*. \tag{8.40}$$

A multivariable example?

Observe that a corner point corresponds to a straight interval and vice versa. Nonconvex parts of the graph of $F(x)$ do not affect the conjugate.

8.3.3 Second conjugate and convexification

It is easy to estimate minimum of a function from above:

$$f(x_a) \geq \min_x f(x)$$

where x_a is any value of an argument. The lower estimate is much more difficult. Duality can be used to estimate the minimum from below. The inequality

$$x\, x^* \leq f(x) + f^*(x^*)$$

provides the lower estimate:

$$f(x) \geq x\, x^* - f^*(x^*) \quad \forall x, \forall x^*$$

Choosing a trial value x^* we find the lower bound.

Example 8.3.7 Revisit above examples for the lower bound

Second conjugate We can compute the conjugate to $F^*(\boldsymbol{z}^*)$, called the *second conjugate* F^{**} to F,

$$F^{**}(\boldsymbol{z}) = \max_{\boldsymbol{z}^*} \{\boldsymbol{z}^* \cdot \boldsymbol{z} - F^*(\boldsymbol{z}^*)\} . \tag{8.41}$$

We denote the argument of F^{**} by $\boldsymbol{z}$.

If $F(\boldsymbol{z})$ is convex, then the transform is an involution. If $F(\boldsymbol{z})$ is not convex, the second conjugate is the convex envelope of F (see [**?**]):

$$F^{**} = \mathcal{C}F. \tag{8.42}$$

The convex envelope of F is the maximal of the convex functions that does not surpass F.

Proof:

8.4 Variational principles of classical mechanics

> All that is superfluous displeases God and Nature
> All that displeases God and Nature is evil.
>
> *Dante*

There are two kinds of variational principles: Differential principles that characterize properties of a motion in each time instance and integral or global principles that characterize the action of motion at finite time interval.

Lagrange principle of minimal action Principle of virtual deflections: The system is in equilibrium when the sum of work of all acting forces F_i on kinematically possible deflections δr_i is zero (starting from Galileo, J.Bernoulli – Lagrange)

$$\sum F_i \delta r_i = 0$$

If the forces have a potential $\boldsymbol{F} = \nabla V$, this principle says that in the equilibrium $\delta V = 0$.

d'Alambert-Legendre principle The same principle works for dynamics, if inertial forces $-m_i \ddot{r}$, where $\ddot{r}$ is the acceleration of the kinematically possible deflections, are added to the forces: (d'Alambert-Legendre principle)

$$\sum (F_i - m_i \ddot{r}) \cdot \delta r_i = 0$$

Notice that there is no requirement for a quantity to reach minimum.