First variation

(one-variable problem)

January 21, 2015

Contents

Since, however, the rules ⟨for isoperimetric curves or, in modern terms, 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

1 Stationarity of an integral functional

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: assuming that the optimal curve $u(x)$ exists among smooth (twice-differentiable curves), we compare the optimal curve with close-by trajectories $u(x) + \delta u(x)$, where variation $\delta u(x)$ is small in some sense. Using the smallness of δu, we simplify the comparison, deriving necessary conditions for the optimal trajectory $u(x)$. The method is quite analogous to the computing of differential $d\,f(x) = f'(x)\,dx$ a function $f(x)$, a minimum point is a solution of the equation $f'(x) = 0$ because of the arbitrariness of differential dx.

Generally speaking, 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, these methods are applicable to a great variety of extremal problems called *variational problems.* Similarly to the calculus problems, the proof of optimality of a trajectory should be complemented by inequalities that distinguish minimum from maximum of saddle points and an existence theorem that either guarantees that the minimizer is indeed a twice-differentiable function or suggests an alternative approach (relaxating or regularization techniques)

1.1 Euler equation (Optimality conditions)

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

$$\min_u I(u), \quad I(u) = \int_a^b F(x, u, u')dx, \quad u(a) = u_a, \; u(b) = u_b, \tag{1}$$

Here integrant F is called *Lagrangian*, it is assumed to be twice a differentiable function of its three arguments, $I(u)$ is called the cost functional. It is not known *a priori* whether the minimizer $u_0(x)$ is smooth, but let us assume that it is twice differentiable function of x.

Example 1.1 Consider the area of the surface of revolution around the axes OX that is supported by two parallel coaxial circles of radii R_a and R_b, the distance between the centers of circles is $b-a$. According to the calculus, the area J of the surface is

$$A(r) = \pi \int_a^b r(x)\sqrt{1+r'(x)^2}\, dx,$$

where $r(x)$ is the variable distance from the OX-axes. The problem of minimal area of such surface

$$I = \min_{r(x)} A(u), \quad r(a) = R_a,\ r(b) = R_b$$

is a variational problem, $r(x)$ is an unknown function.

To derive necessary condition of optimality of a minimizer u_0 we use the ideas of calculus, computing an analog of the derivative of I with respect to u (called the *functional derivative*) and setting it to zero. We suppose that function $u_0 = u_0(x)$ is a minimizer and replace u_0 with a test function $u_0 + \delta u$. The test function $u_0 + \delta u$ satisfies the same boundary conditions as u_0.

If indeed u_0 is a minimizer, then the increment of the cost is $\delta I(u_0) = I(u_0 + \delta u) - I(u_0)$ is nonnegative:

$$\delta I(u_0) = \int_0^1 (F(x, u_0 + \delta u, (u_0 + \delta u)') - F(x, u_0, u_0'))dx \geq 0. \qquad (2)$$

If δu is not specified, the equation (2) is not too informative. However, a particular form of the variation δu simplifies (2) and allows for finding an equation for the minimizer. Calculus of variations suggests a set of tests that differ by various form of variations δu.

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

$$\delta u = \begin{cases} \rho(x) & \text{if} \quad x \in [x_0, x_0 + \varepsilon], \\ 0 & \text{if} \quad x \text{ is outside of } [x_0, x_0 + \varepsilon]. \end{cases} \qquad (3)$$

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

$$|\rho(x)| < \varepsilon, \quad |\rho'(x)| < \varepsilon \ \ \forall x, \quad \rho(x_0) = \rho(x_0 + \varepsilon) = 0 \qquad (4)$$

The integrand at the perturbed trajectory can be expanded into Taylor series,

$$\begin{aligned} F(x, u_0 + \delta u, (u_0 + \delta u)') \ &= F(x, u_0, u_0') + \frac{\partial F(x, u_0, u_0')}{\partial u}\delta u \\ &+ \frac{\partial F(x, u_0, u_0')}{\partial u'}\delta u' + o(\delta u, \delta u') \end{aligned}$$

Here, $\delta u'$ is derivative of the variation δu, $\delta u' = (\delta u)'$, $o(\delta u, \delta u')$ denotes higher order terms which norms are smaller than $\|\delta u\|$ and $\|\delta u'\|$ when $\varepsilon \to 0$. Substituting this expression into (2) and collecting linear (with respect to ε) terms, we rewrite (2) as

$$\delta I(u_0) = \int_a^b \left(\frac{\partial F}{\partial u}(\delta u) + \underline{\frac{\partial F}{\partial u'}(\delta u)'} \right) dx + o(\varepsilon) \geq 0. \tag{5}$$

where The F is calculated at the examined trajectory u_0. To simplify notations, we omit index ($_0$) below.

The variations δu and $(\delta u)'$ are mutually dependent and $(\delta u)'$ can be expressed in terms of δu. Integration by parts of the underlined term in (5) gives

$$\int_a^b \frac{\partial F}{\partial u'}(\delta u)' dx = \int_a^b \left(-\frac{d}{dx}\frac{\partial F}{\partial u'} \right) \delta u \, dx + \frac{\partial F}{\partial u'}\delta u \Big|_{x=a}^{x=b}$$

and we obtain

$$0 \leq \delta I(u_0) = \int_a^b S_F(x, u, u')\delta u \, dx + \frac{\partial F}{\partial u'}\delta u \Big|_{x=a}^{x=b} + o(\varepsilon), \tag{6}$$

where S_F denotes the functional derivative,

$$S_F(x, u, u') = -\frac{d}{dx}\frac{\partial F}{\partial u'} + \frac{\partial F}{\partial u}. \tag{7}$$

The nonintegral term in the right-hand side of (6) is zero, because the boundary values of u are prescribed as $u(a) = u_a$ and $u(b) = u_b$; therefore their variations $\delta u|_{x=a}$ and $\delta u|_{x=b}$ equal zero,

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

Due to the arbitrariness of δu, we arrive at the following

Theorem 1.1 (Stationarity condition) Any differentiable and bounded minimizer u_0 of the variational problem (1) is a solution to the boundary value problem

$$S_F(x, u, u') = \frac{d}{dx}\frac{\partial F}{\partial u'} - \frac{\partial F}{\partial u} = 0 \quad \forall x \in (a, b); \quad u(a) = u_a, \quad u(b) = u_b, \tag{8}$$

called the *Euler–Lagrange equation*.

The Euler–Lagrange equation is also called the *stationary condition of optimality* because it expresses stationarity of the variation.

Using the chain rule, the left-hand side of equation (8) can be rewritten to the form of explicit second-rank differential equation.

$$S_F(x, u, u') = \frac{\partial^2 F}{\partial u'^2}u'' + \frac{\partial^2 F}{\partial u' \partial u}u' + \frac{\partial^2 F}{\partial x \partial u'} - \frac{\partial F}{\partial u} \tag{9}$$

In this derivation, it is indirectly assumed that the extremal $u(t)$ is a twice differentiable function of x.

Example 1.2 Compute the Euler equation for the problem

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

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

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

The minimizer $u_0(x)$ is

$$u_0(x) = \cosh(x) + \frac{a - \cosh(1)}{\sinh(1)} \sinh(x).$$

Remark 1.1 The stationarity test alone does not allow to conclude whether u is a true minimizer or even to conclude that a solution to (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 **??**.

Remark 1.2 The relation (6) request that differential equation (8) is satisfied in each infinitesimal interval but not in in each point. The equation may be not defined in isolated points, if F or its partial derivatives are not defined in these points. The minimizer can change its values at several points, or even at a set of zero measure without effecting the objective functional.

Such solutions are called *weak solutions* **??** of differential equation. The definition of the weak solution naturally arises from the variational formulation that does not check the behavior of the minimizer in every p[oint but only at intervals of nonzero measure. In ambiguous cases, one should specify in what sense (Riemann, Lebesgue) the integral is defined and define of variation accordingly.

1.2 First integrals: Three special cases

In several cases, the second-order Euler equation (8) can be integrated at least once. These are the cases when Lagrangian $F(x, u, u')$ does not depend on one of the arguments.

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

$$I(u) = \min_{u(x)} \int_0^1 F(x, u)dx \tag{10}$$

In this case, Euler equation (8) becomes an algebraic relation for u

$$\frac{\partial F}{\partial u} = 0 \tag{11}$$

The variation does not involve integration by parts, and the minimizer does not need to be continuous.

Monimizer $u(x)$ is determined in each point independently of neighboring points. The boundary conditions in (8) are satisfied by possible jumps of the extremal $u(x)$ in the end points; these conditions do not affect the objective functional at all.

Example 1.3 Consider the problem

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

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

$$u_0(x) = \begin{cases} \sin x & \text{if } \ 0 \le x \le 1 \\ 1 & \text{if } \ x = 0 \\ 0 & \text{if } x = 1 \end{cases}$$

Remark 1.3 Formally, the discontinuous minimizer contradicts the assumption posed when the Euler equation was derived. To be consistent, we need to repeat the derivation of the necessary condition for the problem (10) without any assumption on the continuity of the minimizer. This derivation is quite obvious.

Lagrangian is independent of u If Lagrangian does not depend on u, $F = F(x, u')$, Euler equation (8) becomes

$$S_F(x, u, u') = \frac{d}{dx} \frac{\partial F}{\partial u'} = 0 \quad \forall x \in (a, b); \quad u(a) = u_a, \quad u(b) = u_b, \tag{12}$$

it can be integrated once:

$$\frac{\partial F}{\partial u'} = \text{ constant} \tag{13}$$

The differential equation (13) for u is the *first integral* of equation (8); it defines a quantity that stays constant everywhere along an optimal trajectory. It remains to integrate the first order equation (13) and determine the constants of integration from the boundary conditions.

Example 1.4 Consider the problem

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

(compare with example 1.3) The first integral is

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

Integrating, we find the minimizer,

$$u(x) = \sin x + C_1 x + C_2.$$

The constants C_1 and C_2 are found from and the boundary conditions:

$$C_2 = 1, \quad C_1 = -1 - \sin 1,$$

minimizer u_0 and the cost of the problem become, respectively

$$u_0(x) = \sin x - (\sin 1 + 1)x + 1 \quad I(u_0) = (\sin 1 + 1)^2.$$

Notice that the Lagrangian in the example (1.3) is the square of difference between the minimizer u and function $\sin x$, and the Lagrangian in the example (1.4) is the square of difference of their derivatives. In the problem (1.3), the minimizer coincides with $\sin x$, and jumps to satisfy the prescribed boundary values. The minimizer u in the example (1.4) cannot jump. Indeed, a continuous approximation of a derivative u' of a discontinuous functionis unfunded in the proximity of the point of discontinuity, such a behavior increases the objective functional, and therefore it is nonoptimal.

Lagrangian is independent of x If $F = F(u, u')$, equation (8) has the first integral:

$$W(u, u') = \text{ constant} \tag{14}$$

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 (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

$$u' \left(\frac{\partial^2 F}{\partial u'^2} u'' + \frac{\partial^2 F}{\partial u' \partial u} u' - \frac{\partial F}{\partial u} \right) = 0 \tag{15}$$

The expression in parenthesis coincides with the left-hand-side term $S(x, u, u')$ of the Euler equation in the form (9), simplified for the considered case (F is independent of x, $F = F(u, u')$).

W is constant at any solution $u(x)$ of Euler equation. Instead of solving the Euler equation, we may solve the first-order equation $W = C$.

Example 1.5 Consider the Lagrangian for a harmonic oscillator:

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

The Euler equation is

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

The first integral is

$$W = \frac{1}{2}\left(\omega^2 u^2 + (u')^2\right) = C^2 = \text{ constant}$$

Let us 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

$$\begin{aligned} W = (u')^2 + \omega^2 u^2 \; &= \left[-A\omega\sin(cx) + B\omega\cos(\omega x)\right]^2 \\ &+\omega^2\left[A\cos(\omega x) + B\sin(\omega x)\right]^2 = \omega^2(A^2 + B^2) \end{aligned}$$

We have shown that W is constant at the optimal trajectory. Notice that W is the total energy (sum of the potential and kinetic energy) of the oscillator.

1.3 Variational problem as a limit of a finite-dimensional problem

Here, we derive Euler equation for a finite-dimensional problem that approximate the simplest variational problem

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

Consider a class of piece-wise constant discontinuous 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 $\{u_1, \dots u_N\}$.

Rewriting the variational problem for this class of minimizers, we replace the derivative $u'(x)$ with a finite difference $\mathrm{Diff}\,(u_i)$

$$\mathrm{Diff}\,(u_i) = \frac{1}{\Delta}(u_i - u_{i-1}), \quad \Delta = \frac{b-a}{N}; \tag{16}$$

when $N \to \infty$, this operator tends to the derivative. The variational problem is replaced by finite-dimensional optimization problem:

$$\min_{u_1,\dots,u_{N-1}} I_N \quad I_N = \Delta\sum_{i=1}^{N} F_i(u_i, z_i), \quad z_i = \mathrm{Diff}\,(z_i) = \frac{1}{\Delta}(z_i - z_{i-1}) \tag{17}$$

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

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

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 $z_i = \mathrm{Diff}\,(u_i)$, and the second– only through $z_{i+1} = \mathrm{Diff}\,(u_{i+1})$:

$$\frac{dF_i}{du_i} = \frac{\partial F_i}{\partial u_i} + \frac{\partial F_i}{\partial z_i}\frac{1}{\Delta}, \quad \frac{dF_{i+1}}{du_i} = -\frac{\partial F_{i+1}}{\partial z_{i+1}}\frac{1}{\Delta}.$$
$$\frac{dF_k}{du_i} = 0, \quad k \neq i, k \neq i+1$$

Collecting the terms, we write the stationary condition with respect to u_i :

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

or, recalling the definition (16) of Diff-operator, the form

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

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 $\to \frac{d}{dx}$, $z \to u'$ replacing 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

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

and

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

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

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

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 $\{x_1, \dots x_N\}$ 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.

Remark 1.4 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 1.5 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 **??**

2 Stationarity of boundary terms

2.1 Variation of boundary conditions

Variational conditions and natural conditions The value of minimizer may not be specified on one or both ends of the interval $[a, b]$. In this case, these values are calculated by the minimization of the goal functional together with the minimizer. Consider a variational problem where the boundary value at the right end b of the interval is not defined and the functional directly depends on this value,

$$\min_{u(x):u(a)=u_a} I(u), \quad I(u) = \int_a^b F(x, u, u')dx + f(u(b)) \tag{18}$$

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 comes from the requirement of the stationarity of the minimizer with respect to variation of the boundary term. This term is

$$\left(\frac{\partial F}{\partial u'} + \frac{\partial f}{\partial u}\right)\Bigg|_{x=b} \delta u(b)$$

The first term comes from the integration by part in the derivation of Euler equation, see (6), and the second is the variation of the last term in the objective functional (18): $\delta f(u) = \frac{\partial f}{\partial u}\delta u$. Because the sign of the variation $\delta u(b)$ is arbitrary, the stationarity condition has the form

$$\left(\frac{\partial F}{\partial u'} + \frac{\partial f}{\partial u}\right)_{x=b} = 0. \qquad (19)$$

This equality provides the missing boundary condition at the endpoint $x = b$ for the second order Euler equation. Similar condition can be derived for the point $x = a$ if the value at this point is not prescribed.

Example 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 (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 $f = 0$, the condition (19) becomes

$$\left.\frac{\partial F}{\partial u'}\right|_{x=b} = 0 \qquad (20)$$

and it is called the *natural boundary condition.*

Example 2.2 Consider the Lagrangian $F = a(x)(u')^2 + \phi(x, u)$ where $a(x) \neq 0$ The natural boundary condition is $u'|_{x=b} = 0$.

2.2 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 some problems, F 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. We consider the question: How to join these pieces together?

The first continuity condition is continuity of the (differentiable) minimizer $u(x)$

$$[u]_-^+ = 0 \quad \text{along the optimal trajectory } u(x) \qquad (21)$$

Here $[z]_-^+ = z_+ - z_-$ denotes the jump of the variable z.

The extremal u is differentiable, the first derivative u' exists at all points of the trajectory. This derivative does not need to be continuous. Instead, Euler equation requests the differentiability of $\frac{\partial F}{\partial u'}$ to ensure the existence of the term $\frac{d}{dx}\frac{\partial F}{\partial u'}$ in the Euler equation.

Integrating the stationarity condition (8), we obtain stationarity in the integral form

$$\int_a^x S_F(x,u,u')dx = \int_a^x \left(\frac{d}{dx}\frac{\partial F(x,u,u')}{\partial u'} - \frac{\partial F(x,u,u')}{\partial u} \right) dx = 0$$

or

$$\frac{\partial F(x,u,u')}{\partial u'} = \int_{x_0}^x \frac{\partial F(x,u,u')}{\partial u} dx + \left. \frac{\partial F(x,u,u')}{\partial u'} \right|_{x=a} \tag{22}$$

If $\frac{\partial F}{\partial u}$ is bounded at the optimal trajectory, the right-hand side is a continuous function of x, and so is the left-hand side. This requirement of continuity of an optimal trajectory is called the *Weierstrass-Erdman condition on broken extremal*.

Theorem 2.1 At any point of the optimal trajectory, the Weierstrass-Erdman condition must be satisfied:

$$\left[\frac{\partial F}{\partial u'} \right]_-^+ = 0 \quad \text{along the optimal trajectory } u(x). \tag{23}$$

Example 2.3 (Broken extremal) Consider the Lagrangian

$$F = \frac{1}{2}c(x)(u')^2 + \frac{1}{2}u^2, \quad c(x) = \begin{cases} c_1 & \text{if } x \in [a, x_*) \\ c_2 & \text{if } x \in (x_*, b] \end{cases},$$

where x_* is point in (a,b). The Euler equation is held everywhere in (a,b) except of the point x_*,

$$\frac{d}{dx}[c_1 u'] - u = 0 \quad \text{ if } x \in [a,\, x_*)$$
$$\frac{d}{dx}[c_2 u'] - u = 0 \quad \text{ if } x \in (x_*,\, b],$$

At $x = x_*$, the continuity conditions hold,

$$u(x_* - 0) = u(x_* + 0), \quad c_1\, u'(x_* - 0) = c_2\, u'(x_* + 0).$$

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

$$\frac{u'(x_* + 0)}{u'(x_* - 0)} = \frac{c_1}{c_2}$$

These conditions together with the Euler equation and boundary conditions determine the optimal trajectory.

3 Functional 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

$$\delta J = \int_a^b \left(\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u'} + \frac{d^2}{dx^2}\frac{\partial F}{\partial u''} \right) \delta u\, dx + \left[\frac{\partial F}{\partial u'}\delta u + \frac{\partial F}{\partial u''}\delta u' - \frac{d}{dx}\frac{\partial F}{\partial u''}\delta u \right]_{x=a}^{x=b} \tag{24}$$

The stationarity condition becomes the fourth-order differential equation

$$\frac{d^2}{dx^2}\frac{\partial F}{\partial u''} - \frac{d}{dx}\frac{\partial F}{\partial u'} + \frac{\partial F}{\partial u} = 0 \tag{25}$$

supplemented by two boundary conditions on each end,

$$\delta u' \frac{\partial F}{\partial u''} = 0, \quad \delta u \left[\frac{\partial F}{\partial u'} - \frac{d}{dx}\frac{\partial F}{\partial u''} \right] = 0 \quad \text{at } x = a \text{ and } x = b \tag{26}$$

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

Example 3.1 The equilibrium of an elastic bending beam correspond to the solution of the variational problem

$$\min_{u(x)} \int_0^L (\frac{1}{2}(E(x)u'')^2 - q(x)u)dx \tag{27}$$

where $u(x)$ 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: $u(a) = 0, \quad u'(a) = 0$
(2) a simply supported end $u(a) = 0$.
(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 (25) becomes

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

The equations (26) become

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

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 $(Eu'')' = 0$; the last expresses vanishing of the bending force at the free end (the bending momentum vanishes here as well).

Generalization The Lagrangian

$$F\left(x, u, u', \ldots, u^{(n)}\right)$$

dependent on first k derivatives of $u$dependent on higher derivatives of u is considered similarly. The stationary condition is the $2k$-order differential equation

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

supplemented at each end $x = a$ and $x = b$ of the trajectory by k boundary conditions

$$\left[\frac{\partial F}{\partial u^{(k)}}\right] \delta u^{(k-1)}|_{x=a,b} = 0$$

$$\left[\frac{\partial F}{\partial u^{(k-1)}} - \frac{d}{dx}\frac{\partial F}{\partial u^{(k)}}\right] \delta u^{(k-2)}|_{x=a,b} = 0$$

$$\ldots$$

$$\left[\frac{\partial F}{\partial u'} - \frac{d}{dx}\frac{\partial F}{\partial u''} + \ldots + (-1)^k \frac{d^{(k-1)}}{dx^{(k-1)}}\frac{\partial F}{\partial u^{(k)}}\right] \delta u|_{x=a,b} = 0$$

If u is a vector minimizer, u can be replaced by a vector but the structure of the necessary conditions stay the same.

4 Non-fixed interval

4.1 Transversality condition

Free boundary Consider now the case when the interval $[a, b]$ is not fixed, but the end point is to be chosen so that it minimizes the functional. Let us compute the difference between the two functionals over two different intervals

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

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

The second integral is estimated as

$$\int_b^{b+\delta x} F(x, u + \delta u, u' + \delta u')dx = F(x, u, u')|_{x=b}\, \delta x + o(\|\delta u\|, |\delta x|)$$

and the first integral is computed as before with integration by parts:

$$\int_a^b S_F(x, u, u')\delta u\, dx + \left.\frac{\partial F}{\partial u'}\right|_{x=b} \delta u(b) = 0$$

1. Suppose that no boundary conditions are imposed at the minimizer at the point $x = b$. Because of arbitrariness of δx and δu, we arrive at the conditions:

$$S_F(x, u, u') = 0 \quad x \in (a, b), \quad \left.\frac{\partial F}{\partial u'}\right|_{x=b} = 0,$$

and

$$F(x, u, u')|_{x=b} = 0. \tag{28}$$

Euler equation for the extremal satisfies an extra boundary condition (28), but has also an additional degree of freedom: unknown coordinate b.

Example 4.1 Consider the problem

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

The Euler equation $u'' + 1 = 0$ and the condition at $u(0) = 0$ corresponds to the extremal

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

where A is a parameter. The condition $\frac{\partial F}{\partial u'} = u' = 0$ at the unknown right end $x = s$ gives $s = A$. The transversality condition $F = 0$ or

$$(-u + x)|_{x=A=s} = \frac{1}{2}s^2 - s^2 + s = s\left(1 - \frac{1}{2}s\right) = 0$$

We find $s = 2$, $u = -\frac{1}{2}x^2 + 2x$.

2. Next, consider the problem in which the boundary data at $x = b$ is prescribed, $u = \beta$, but the value of b is not known. In the perturbed trajectory, the

boundary condition is $u(b+\delta x) = \beta$. The value of $u(b+\delta x)$ is an extrapolation of $u(x)$ as follows

$$u(b+\delta x) = u(b) + u'(b)\delta x + o(\|\delta u\|, |\delta x|)$$

Therefore, the value $(u+\delta u)_{x=b}$ depends on δx, $u(b) = \beta - u'(b)\delta x$ or $\delta u(b) = -u'(b)\delta x$. Combining the depending on δx terms, we obtain the condition

$$\left(F - \frac{\partial F}{\partial u'} u'(b)\right)\bigg|_{x=b} \delta x$$

Because δx is arbitrary, the boundary conditions are: $u = \beta$ and

$$\left(F(x,u,u') - u' \frac{\partial F}{\partial u'}\right)\bigg|_{x=b} = 0. \tag{29}$$

Remark 4.1 Notice that the left-hand side expression in (29) at the unknown end is identical to the expression for the first integral (14) of the problem in the case when $F(u,u')$ is independent of x. This integral is constant along an optimal trajectory which means that the problem with Lagrangian $F(u,u')$ does not satisfy (29) at an isolated point.

3. Finally, consider the problem when the rajectory ends at a curve. If the boundary value depends on b,

$$u(b) = \phi(b) \tag{30}$$

then the variations δu and δx are bounded: $\delta u = \phi' \delta x$. The two stationarity conditions at the end point

$$\frac{\partial F}{\partial u'}\delta u = 0 \quad \text{and} \quad \left(F(x,u,u') - u'\frac{\partial F}{\partial u'}\right)\delta x = 0$$

together (30) gives the conditions

$$\left(F - (u' - \phi')\frac{\partial F}{\partial u'}\right)\bigg|_{x=b} = 0 \quad \text{and } u(b) = \phi(b). \tag{31}$$

The next example deals with a constraint at the unknown length of the interval and the boundary data.

Example 4.2 Find the shortest path between the origin and a curve $\phi(x)$.
The path length is given by

$$I = \min_{y(x),s} \int_0^s \sqrt{1+y'^2}dx, \quad u(0) = 0$$

At the end point x the path meets the curve, therefore $y(s) = \phi(s)$ or

$$\delta y = \phi'(s)\delta s \tag{32}$$

The Euler equation

$$\frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1+y'^2}} = C$$

shows that $y' =$ constant, therefore the path is a straight line, $y = Ax$ as expected. At the point s, the variation is

$$\left(u' \frac{\partial F}{\partial y'} - F\right) \delta x + y' \frac{\partial F}{\partial y'} \delta y = \frac{1}{\sqrt{1+y'^2}} \delta x + \frac{y'}{\sqrt{1+y'^2}} \delta u$$

The stationarity gives the relation $-\delta x + y' \delta u = 0$.

Comparing it with the constraint (32), we conclude that $y'(s)\phi'(s) = -1$, or that the shortest path is a straight line orthogonal to the curve $\phi(x)$, as it is expected.

4.2 Extremal broken at an 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

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

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

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

The stationarity conditions at the unknown point ξ consist of stationarity of the trajectory

$$\frac{\partial F_+}{\partial u'} = \frac{\partial F_-}{\partial u'} \tag{33}$$

and stationarity of the position of the transit point

$$F_+(u) - u'_+ \frac{\partial F_+}{\partial u'} = F_-(u) - u'_- \frac{\partial F_-}{\partial u'}. \tag{34}$$

or

$$F_+(u) - F_-(u) = (u'_+ - u'_-) \frac{\partial F}{\partial u'}. \tag{35}$$

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

In particular, if the Lagrangian is independent of x, the condition (34) expresses the constancy of the first integral (14) at the point ξ.

Example 4.3 Consider the problem with Lagrangian

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

and boundary conditions

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

The Euler equation is

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

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 three conditions at the unknown point ξ which express
(1) continuity of the extremal

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

(2) Weierstrass-Erdman condition

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

(3) transversality condition

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

The transversality condition states the equality of two first integral. It is simplified to

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

From the Weierstrass-Erdman condition, we find

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

The first condition and the definition of q 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

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

and transversality conditions:

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

5 Several minimizers

5.1 Euler equations and first integrals

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

$$I(u) = \min_u \int_a^b F(x, u, u')dx, \tag{36}$$

where x is a point in the interval $[a, b]$ and $u = (u_1(x), \ldots, u_n(x))$ 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 variations of each component of vector u applying variations $\delta_i u = (0, \ldots, \delta u_i \ldots, 0)$. The increment of the objective functional should be zero for each of these variation, otherwise the functional can be decreased by one of them. The stationarity condition for any of considered variations coincides with the one-minimizer case.

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

Proceeding as before, we obtain the system of n second-order differential equations,

$$\frac{d}{dx}\frac{\partial F}{\partial u_i'} - \frac{\partial F}{\partial u_i} = 0, \quad i = 1, \ldots n \tag{37}$$

and the boundary term

$$\sum_{i=1}^n \frac{\partial F}{\partial u_i'}\delta u_i \Bigg|_{x=a}^{x=b} = 0 \tag{38}$$

If the value of $u_i(a)$ or $u_i(b)$ is not prescribed, the natural boundary conditions $\frac{\partial F}{\partial u_i'}|_{x=a}$ or $\frac{\partial F}{\partial u_i'}|_{x=b}$, respectively, must be satisfied.

The vector form of the system (37),

$$S_F(u) = \frac{d}{dx}\frac{\partial F}{\partial u'} - \frac{\partial F}{\partial u} = 0, \quad \delta u^T \frac{\partial F}{\partial u'}\Bigg|_{x=a}^{x=b} = 0 \tag{39}$$

is identical to the scalar Euler equation. This system corresponds to an definition of differentiation with respect to a vector argument u.

Example 5.1 Consider the problem with the integrand

$$F = \frac{1}{2}u_1'^2 + \frac{1}{2}u_2'^2 - u_1 u_2' + \frac{1}{2}u_1^2 \tag{40}$$

The system of stationarity conditions is computed to be

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

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

First integrals The first integrals that are established for the special cases of the scalar Euler equation, can also be derived for the vector equation.

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 one of differential equation in (37) becomes an algebraic one. 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 discontinuities of $u_k(x)$ 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 5.1 can be integrated and replaced by
$$u_2' - u_1 = \ \text{constant}$$

3. If F is independent of x, $F = F(u, u')$ then a first integral exist
$$W = u'^T \frac{\partial F}{\partial u'} - F = \text{constant} \tag{41}$$
Here
$$u'^T \frac{\partial F}{\partial u'} = \sum_{i=1}^{n} u_i' \frac{\partial F}{\partial u_i'}$$
For the Example 5.1, this first integral is computed to be
$$\begin{aligned}
W \ &= \ u_1'^2 + u_2' \left(u_2' - u_1'\right) - \left(\frac{1}{2}u_1'^2 + \frac{1}{2}u_2'^2 - u_1 u_2' + \frac{1}{2}u_1^2\right) \\
&= \ \frac{1}{2}\left(u_1'^2 + u_2'^2 - u_1^2\right) = \ \text{constant}
\end{aligned}$$

These three cases do not exhaust all possible first integrals for vector case. For example, if the functional depends only on, say $(u_1 + u_2)$, one can hope to find new invariants by changing the variables. We discuss this matter below in Sections **??** and **??**.

Transversality and Weierstrass-Erdman conditions These conditions are quite analogous to the scalar case and their derivation is straightforward. We simply list them here.

The expressions $\frac{\partial F}{\partial u_i'}$, $i = 1 \ldots, n$ 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

$$u^T \frac{\partial F}{\partial u'} - F = 0 \text{ or } \sum_{i=1}^{n} u_i' \frac{\partial F}{\partial u_i'} - F = 0$$

at the end point is satisfied.

5.2 Variational boundary conditions

Consider the variation of the boundary term (38) which we rewrite here for convenience

$$\frac{\partial F}{\partial u_1'} \delta u_1 + \ldots + \frac{\partial F}{\partial u_n'} \delta u_n \Big|_{x=a}^{x=b} = 0 \tag{42}$$

If all variations $\delta u_i(a)$, $\delta u_i(b)$ are free, it produces $2n$ boundary conditions

$$\frac{\partial F}{\partial u_i'} = 0, \quad x = a \text{ and } x = b$$

for Euler equations (37). In the opposite case, when the values of all minimizers are prescribed at the end points,

$$u_i(a) = u_i^a, \quad u_i(b) = u_i^b, \; i = 1, \ldots, n$$

then the equation (42) is satisfied, because all variations are zero.

$$\delta u_i(a) = 0, \quad \delta u_i(b) = 0, \; i = 1, \ldots, n$$

If the values of several components of $u(a)$ or $u(b)$ are not given, the variations of these components are free and the corresponding natural boundary condition supplements the boundary conditions: For each $i = 1, \ldots, n$ one of the two conditions holds

$$\text{Either } \left. \frac{\partial F}{\partial u_i'} \right|_{x=a,b} = 0 \quad \text{ or } \delta u_i|_{x=a,b} = 0. \tag{43}$$

The total number of the conditions at each endpoint is n. The missing main boundary conditions are supplemented by the natural conditions that express the optimality of the trajectory. This number agrees with the number of boundary conditions needed to solve the boundary value problem for Euler equation.

In a general case, p relations ($p < 2n$) between boundary values of u are prescribed,

$$\beta_k(u_1(a), \ldots, u_n(a), u_1(b), \ldots, u_n(b),) = 0 \tag{44}$$

at the end points $x = a$ and $x = b$. In this case, we need to find $2n - p$ supplementary variational constraints at these points that together with (44) give $2n$ boundary conditions for the Euler equation (38) of the order $2n$.

The conditions (44) are satisfied at all perturbed trajectories,

$$\beta_k(w + \delta w) = 0$$

where $2n$ dimensional vector w is the direct sum of $u(a)$ and $u(b)$ defined as:

$$\begin{array}{ll} w_k = u_k(a) & \text{if } k = 1, \ldots, n \\ w_k = u_{k-n}(b) & \text{if } k = n+1, \ldots, 2n. \end{array}$$

The variation δw_i is constraint by a linear system

$$\frac{\partial \beta_k}{\partial w}^T \delta w = 0, \quad k = 1, \ldots, p$$

which shows that the constraints (44) are satisfied at the varied trajectory. A matrix form of these conditions is

$$P \delta w = 0,$$

where

$$P = \begin{pmatrix} \frac{\partial \beta_1}{\partial u_1} & \cdots & \frac{\partial \beta_1}{\partial u_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial \beta_p}{\partial u_1} & \cdots & \frac{\partial \beta_p}{\partial u_n} \end{pmatrix}$$

and we may also assume that these conditions are linearly independent. Then the solution to the system is a vector $\delta w_{\rm ad}$ of the form

$$\delta w_{\rm ad} = Q v$$

where v is an arbitrary $(2n - p)$-dimensional arbitrary vector and $(2n - p) \times n$ matrix Q is a supplement (orthogonal matrix) to P that is defined as a solution to the matrix equation $PQ = 0$ which leads to $P \delta w_{\rm ad} = P\, Q \delta v = 0$ for any v.

Any admissible variation $\delta w_{\rm ad}$ makes the fist variation (39) of the objective functional vanish; correspondingly, we have

$$\left(\frac{\partial F}{\partial u'} \right)^T \delta w_{\rm ad} = 0$$

Using the representation of $\delta w_{\rm ad}$ and the arbitrariness of the potentials v, we conclude that the fist variation vanishes is the coefficient by each of these potentials is zero or

$$\left(\frac{\partial F}{\partial u'} \right)^T Q = 0 \qquad (45)$$

This representation provides the $2n - p$ missing boundary conditions.

Example 5.2 Consider again the variational problem with the Lagrangian (40) 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

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

Since the condition$u_1(a) = 1$ is prescribed, it becomes

$$u_2'(a) = 1$$

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 ($\delta u = Q \delta v$)

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

where δv is an arbitrary scalar. The variational condition at $x = b$ becomes

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

or

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

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 obtained from the minimization requirement.

Periodic boundary conditions Consider a variational problem with periodic boundary conditions $u(a) = u(b)$. The variational boundary conditions are

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

They are obtained from the expression (42) and equalities $\delta u(a) = \delta u(b)$.

6 Geometric optics

6.1 Geometric optics problem.

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). The Fermat 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

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

Then,

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

Consider minimization of travel time T by the trajectory. The corresponding Lagrangian has the form

$$F(y, y') = \psi(y)\sqrt{1 + y'^2}, \quad \psi(x, y) = \frac{1}{v(x, y)},$$

where $\psi(x, y)$ is the slowness, and Euler equation is

$$\frac{d}{dx}\left(\psi(x, y)\frac{y'}{\sqrt{1 + y'^2}}\right) - \frac{\partial\psi(x, y)}{\partial y}\sqrt{1 + y'^2} = 0, \quad y(a) = y_a, \; y(b) = y_b$$

Assume that the medium is layered and the speed $v(y) = \frac{1}{\psi(y)}$ of travel varies only along the y axes. Then, the problem allows for the first integral, (see (**??**))

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

which simplifies to

$$\psi(y) = -c\sqrt{1 + y'^2}. \tag{46}$$

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

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

that has an explicit solution in quadratures:

$$x = \pm \Phi(u) = \int \frac{c\, dy}{\sqrt{\psi^2(y) - c^2}}. \tag{47}$$

Notice that equation (46) 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 (46) assumes the form

$$\psi(y)\ \cos \alpha = c \tag{48}$$

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.

6.2 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

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

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

$$\frac{\sin \alpha_+}{v_+} = \frac{\sin \alpha_-}{v_-}$$

Snell's law for an inclined plane 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 (**??**) 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.

6.3 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.

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:

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

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 (46), we obtain

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

and

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

To compute the quadrature, we change the variable in the integral:

$$y = y_0 + 2a\sin^2\frac{\theta}{2}, \quad dy = 2a\sin\frac{\theta}{2}\cos\frac{\theta}{2}d\theta$$

and find

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

The optimal trajectory is a parametric curve

$$\begin{aligned} x &= x_0 + a(\theta - \sin\theta), \\ y &= y_0 + a(1 - \cos\theta), \end{aligned} \tag{49}$$

We recognize the equation of the cycloid in (49). 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 6.1 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.

Notice that brachistochrone is in fact an optimal design problem: the trajectory must be chosen by a designer to minimize the time of travel.

6.4 Minimal surface of revolution

Another classical example of design problem solved by variational methods is the problem of minimal surface. 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

$$J = \pi \int_0^a r\sqrt{1 + r'^2}\, dx \tag{50}$$

This problem is again a special case of the geometric optic, corresponding to $\psi(r) = r$. Equation (47) becomes

$$x = \int \frac{dr}{\sqrt{c^2 r^2 - 1}} = \frac{1}{C} \cosh^{-1}(Cr)$$

and we find

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

Assume for clarity that the surface is supported by two equal circles of radius R located symmetrically to OX axis; then by symmetry $x_0 = 0$ and $c_1 = 0$ and the surface is

$$r = \frac{1}{C} \cosh\left(Cx\right)$$

The extremals $Cr = \cosh(Cx)$ form a C-dependent family of functions varying only by simultaneously rescaling of the variables r and x. The angle towards a point on the curve from the origin is independent on the scale. This implies that all curves of the family lie inside the triangular region which boundary is defined by a minimal values of that angle, or by the equation $\frac{r}{x} = f'(x)$ of the enveloping curve. This equation reads

$$\frac{\cosh(r)}{r} = \sinh(r)$$

and leads to $r = 1.199678640...$ All extremals lie inside the triangle $\frac{r}{|x|} \leq 1.50887956$.

Analysis of this formula reveals unexpected features: The solution may be either unique, or the problem may have 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 from the common sense viewpoint because the problem of minimal area seems to have 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 (Goldschmidt solution) that explores an alternative: Two circles and an infinitesimal bar between them. The objective functional is $I_0 = 2\pi R^2$.

Goldschmidt solution is a limit of a minimizing sequence that geometrically is a sequence of two cones supported by the circles $y(x_0) = R$ and turned toward each other,

$$r_n(x) = \begin{cases} n\left(x - x_0 + \frac{1}{n}\right) R & \frac{n\,x_0 - 1}{n} < x \leq x_0 \\ n\left(-x + x_0 - \frac{1}{n}\right) R & \frac{-nx_0 + 1}{n} > x \geq -x_0 \\ \frac{1}{n} & +\frac{-x_0 n + 1}{n} > x \geq \frac{nx_0 - 1}{n} \end{cases}$$

or

$$r_n'(x) = \begin{cases} n & x_0 - \frac{1}{n} < x \leq x_0 \\ -n & -x_0 + \frac{1}{n} > x \geq -x_0 \\ 0 & -x_0 + \frac{1}{n} > x \geq -x_0 - \frac{1}{n} \end{cases}$$

and an infinitesimally thin cylinder $r = \frac{1}{n^2}$ that join them.

When $n \to \infty$, the contribution of central part disappears, the cones flatten and become circles. r_u tents to a discontinuous function

$$\lim_{n \to \infty} r_n(x) = \begin{cases} 0, & |x| \leq x_0 \\ R, & |x| = x_0 \end{cases}$$

The values on functional at the minimizing sequence are the areas of the cones, they obviously tend to the area of the circle. The derivative of the minimizing sequence growth indefinitely in the proximity of the endpoints, however, the intregal (6.4) stays finite. 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 $[x(t), r(t)]$ then $r' = \tan\alpha$ where α is the angle of inclination to OX axis. The equation (48) that takes the form

$$r \cos\alpha = C$$

admits either the regular solution $r = C \sec \alpha$, $C \neq 0$ which yields to the catenoid (**??**), or the singular solution $C = 0$ and either $r = 0$ 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 $v = 1/r$ is inverse proportional to the distance to the axis OX. The optimal path between the two close-by points lies along the arch of an appropriate catenoid $r = \cosh(Cx)$ 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 traveler goes straight to the OX-axis where the speed is infinite, then is transported 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.

Reformulation of the minimal surface problem The paradox in the solution was caused by the infinitisimally thin connection path between the supports. One way to regularize it is to request that this path has a finite radius everywhere that is to impose an additional constraint $r(x) \geq r_0$ in (50). With such constraint, the solution splits into a cylinder $r(x) = r_0$ and the catenoid

$$r(x) = \begin{cases} r_0 & \text{if } 0 \leq x \leq s \\ \frac{\cosh(C(x - C_1))}{C} & \text{if } s \leq x \leq a \end{cases}$$

where

$$R = \frac{\cosh(C(a - C_1))}{C}, \quad r_0 = \frac{\cosh(C(s - C_1))}{C}.$$

the minimal area is given by

$$A(s_0, C) = 2 \min_{s,C} \left(2\pi r_0 s - \frac{\sinh(C\,s) - \sinh(C\,x_0)}{C^2} \right)$$

It remains to find (numerically) optimal values os s and C and C_1.