Contents

Chapter 2

Stationarity

The famous Euler undertook the task of reducing all such investigations [determination of optimal curves] *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

2.1 Euler equation

2.1.1 Derivation of Euler equation

Simplest problem of Calculus of Variations Consider the extremal problem called *the simplest problem of the calculus of variations*

$$\min_{u(x), x\in[a,b]} I(u), \quad I(u) = \int_a^b F(x, u, u')dx, \quad u(a) = u_a,\ u(b) = u_b. \tag{2.1}$$

where integrant F is twice a differentiable function of its three arguments, $x \in [a, b]$ is the real variable, $u(x)$ is a differentialble scalar function, and $u'(x)$ is its derivative. Assume that the boundary values $u(a) = u_a$ and $u(b) = u_b$ of $u(x)$ are fixed. The problem is to find $u(x)$ that minimizes the objective functional I.

The follwing examples of the simplest problems are geometically evident. They use the definitions known from the standard calculus books:

1. The problem of the shortest path S between two points (a, A) and (b, B) at the plane (which of cause has a geometically trivial solution). The objective functional (the length of the path) L is

$$L = \int_a^b \sqrt{1 + u'^2}\, dx; \quad u(a) = A, u(b) = B \tag{2.2}$$

where u represent the ordinate and x the absciss of the path $u(x)$.

2. The problem of minimal surface of revolution supported by two circles. In cylindrical coordinates, the surface $r(h)$ minimizes its area A, that is

$$A = \pi \int_a^b r\sqrt{1+r'^2}\, dx; \quad r(a) = A, r(b) = B \tag{2.3}$$

where a and b are the h coordinates of the centers of the supporting circles and A and B are their radii.

3. The problem of the quickest path T (problem of geometric optics) between two points (a, A) and (b, B) at the plane. It is assumed that the speed v of the motion depends on the point which the path is crossing, $v = v(x, u)$. This problem remains the problem of the shortest path but is he objective (the time of travel) T depends on the speed is

$$T = \int_a^b \frac{\sqrt{1+u'^2}}{v(x,u)}\, dx; \quad u(a) = A, u(b) = B \tag{2.4}$$

Variational method Calculus of variation derives the equations for the minimizer $u(x)$. The analysis of conditions of optimality of $u(x)$ is similar to the optimality conditions in finite dimensional calculus where the functional has an extremum if its derivative is zero.

The condition is derived as follows: Assume that an 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, 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*. The technique was pioneered 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.

Supposing that function $u_0 = u_0(x)$ is a minimizer, we replace u_0 with a test function $u_0 + \delta u$. 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:

$$\delta I(u_0) = \int_a^b (F(x, u_0 + \delta u, (u_0 + \delta u)') - F(x, u_0, u_0'))dx \geq 0. \tag{2.5}$$

When δu is not specified, the equation (2.5) 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 (2.5).

Stationarity condition The simplest variational condition (the Euler–Lagrange equation) is derived assuming that that the test function $u_0 + \delta u$ satisfies the same boundary conditions as u_0, so that $\delta u(a) = \delta u(b) = 0$ and that the variation δu is infinitesimal and localized:

$$\delta u = \begin{cases} \rho(x) & \text{if} \quad x \in [x_0, x_0 + \varepsilon], \\ 0 & \text{if} \quad x \text{ is outside of } [x_0, x_0 + \varepsilon]. \end{cases} \tag{2.6}$$

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

$$|\rho(x)| < \varepsilon, \quad |\rho'(x)| < \varepsilon \quad \forall x. \tag{2.7}$$

As an example, one can consider

$$\rho(x)) = \pm\varepsilon^2 \left(x - (x_0 + \frac{1}{2}\varepsilon) \right)^2. \tag{2.8}$$

Using the smallness of the variation, we expand the integrant at the perturbed trajectory into Taylor series neglecting the smaller terms,

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

where $o(\delta u, \delta u')$ means terms that are smaller than δu and $\delta u'$ when $\varepsilon \to 0$. Substituting this expression into (2.5) and collecting the linear (with respect to ε) terms, we rewrite (2.5) as

$$\delta I(u_0) = \int_a^b \left(\frac{\partial F}{\partial u}(\delta u) + \underline{\frac{\partial F}{\partial u'}(\delta u)'} \right) dx + o(\varepsilon) \geq 0. \tag{2.9}$$

The increment depends on the two connected variations, δu and $\delta u'$. We exclude $\delta u'$ by integration by parts of the underlined term in (2.9):

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

The increment becomes

$$0 \leq \delta I(u_0) = \varepsilon \int_a^b S_F(u) \delta u \, dx + \left. \frac{\partial F}{\partial u'} \delta u \right|_{x=a}^{x=b} + o(\varepsilon), \tag{2.10}$$

where we S_F denotes the differential operator

$$S_F(u) = -\frac{d}{dx} \frac{\partial F}{\partial u'} + \frac{\partial F}{\partial u}. \tag{2.11}$$

The nonintegral term in the right-hand side of (2.10) is zero, because the boundary values of u are prescribed, hence 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 conclude the following

Theorem 2.1.1 (Stationarity condition) Any twice continuously differentiable and bounded minimizer u_0 of the variational problem (2.1) is a solution to the boundary value problem

$$S_F(u) = \frac{d}{dx}\frac{\partial F}{\partial u'} - \frac{\partial F}{\partial u} = 0 \quad \forall x \in (a,b); \quad u(a) = u_a, \quad u(b) = u_b, \qquad (2.12)$$

called the *Euler–Lagrange equation*, assuming that the solution to this equation exists.

Proof: Assume the contrary, $S_F(u) > 0$ and $x = x_* \in (a,b)$. By assumption, minimizer u_0 us continuously twice differentiable function of x and F is continuously twice diffeentiable function of its arguments, therefore $S_F(u)$ is a continuous function of x, and there exists an interval $x_*, x_* + \varepsilon$, such that $S_F(u(x)) > 0$ at every point $x \in (x_*, x_* + \varepsilon)$ of this interval. The variation δu can be chosen so that $x_0 = x_*$ and $\rho(x) < 0$ in $x_*, x_* + \varepsilon$, see (2.6). The increment δI, see (2.10), becomes negative, $\delta I < 0$, hence u is not a minimizer. A similar analysis excludes the other remaining option, $S_F(u) < 0$ and $x = x_* \in (a,b)$. Therefore (2.12) holds.

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

Remark 2.1.1 The stationarity test alone does not allow to conclude whether u_0 is a true minimizer or even to conclude that a solution to (2.12) 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 **??**.

Indirectly, we assume in this derivation that u_0 is a twice differentiable function of x. Using the chain rule, the left-hand side of equation (2.12) can be rewritten as

$$S(x,u,u') = u''\frac{\partial^2 F}{\partial u'^2} + u'\frac{\partial^2 F}{\partial u' \partial u} + \frac{\partial^2 F}{\partial u' \partial x} - \frac{\partial F}{\partial u} \qquad (2.13)$$

Example 2.1.1 Compute the Euler equation for the problem

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

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

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

The minimizer $u_0(x)$ is

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

Remark 2.1.2 (The weak form of the stationarity conditions) The construction of Euler-Lagrange equation does not require that $S_F(u)$ be defined in all points x; it may stay undefined of take nonzero values in isolated point x of the interval (a, b). Later we consider a broken extremals that do not have the second derivative at some points.

The Euler equation is understood in the *weak sense*, as the integral identity

$$\int_a^b \left(\frac{\partial F}{\partial u} v + \frac{\partial F}{\partial u'} v' \right) dx = 0 \quad \forall v(x) \in V \tag{2.14}$$

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 several points, or at a set of zero measure without alternation the objective functional.

2.1.2 Natural Boundary conditions

Natural conditions The condition $u(b) = u_b$ on one or both ends of extremal may not be specified, but the extremal problem still has a meaning. In this case, a *variational boundary condition* is posed that is derived from the requirement of the stationarity of the minimizer with respect to variation of the boundary term. Assume for definiteness that $u(b)$ can vary and $u(a)$ is fixed. Recall, that the variation (2.10) of the simplest variational functional has the form

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

In the right-hand side of the last equation, the integral term is zero by virtue of Euler equation and the last term is zero because of assumed boundary condition $u(a) = u_a$, which implies $\delta u(a) = 0$. The stationary condition $\delta I = 0$ implies that

$$\left. \frac{\partial F}{\partial u'} \right|_{x=b} = 0 \tag{2.15}$$

because $\delta u(b)$ is arbitrary. This condition is called is the *natural* or *variational* boundary condition. Similar condition can be derived for the point $x = a$ if the value of $u(a)$ is not prescribed.

Example 2.1.2 The natural boundary condition for the problem with the Lagrangian $L = (u')^2 + \phi(x, u)$ is $u'|_{x=b} = 0$

In canonic variables, the natural boundary condition takes a specially simple form

$$p|_{x=b} = 0 \tag{2.16}$$

Problem with a boundary cost Also, the objective functional may contain terms defined on the boundary only in which case the problem becomes

$$\min_{u(x):u(a)=u_a} I(u), \quad I(u) = \int_a^b F(x,u,u')dx + f(u(b)) \tag{2.17}$$

The Euler equation for the problem remains 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 (2.17). The stationarity condition with respect to the variation of $\delta u(b)$ becomes

$$\left(\frac{\partial F}{\partial u'} + \frac{\partial f}{\partial u} \right) \Bigg|_{x=b} = 0 \tag{2.18}$$

Null-Lagrangian approach We may approach the problem (2.17) transforming the term $f(u(b))$ into an integral term

$$f(u(b)) = \int_a^b \frac{df}{dx} dx + f(u(a)) = \int_a^b \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial u} u' \right) dx + f(u(a))$$

The term $f(u(a))$ is constant because $u(a)$ is prescribed; it does not change the optimality conditions. The problem (2.17) is equivalent to the simplest variational problem with a Lagrangian $\breve{F}$ equal to

$$\breve{F}(x,u,u') = F(x,u,u') + \frac{\partial f}{\partial x} + \frac{\partial f}{\partial u} u' \tag{2.19}$$

The Euler equation of the two problems are identical because the contribution of the last two terms are zero. Indeed,

$$\left(\frac{d}{dx} \frac{\partial}{\partial u'} - \frac{\partial}{\partial u} \right) \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial u} u' \right) = \frac{d}{dx} \frac{\partial f}{\partial u} - \frac{\partial^2 f}{\partial x \partial u} - u' \frac{\partial^2 f}{\partial u^2} = 0$$

because $\frac{d}{dx} = \frac{\partial}{\partial x} + u' \frac{\partial}{\partial u}$ (the chain rule).

The natural boundary condition for $\hat{F}$ is

$$\frac{\partial F}{\partial u'} = \frac{\partial F}{\partial u'} + \frac{\partial f}{\partial u'} = 0$$

The Lagrangians that correspond to identical zero operator $S(u)$ are called Null-Lagrangians. They influence the boundary conditions only.

Example 2.1.3 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. Stationary conditions are

$$u'' = 0 \quad u(0) = 0, \quad u'(1) + A = 0$$

The extremal is a straight line, $u = -Ax$. The cost of the problem is $I = -\frac{1}{2}A^2$.

2.1.3 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

$$\delta J = \int_a^b \left(\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u'} + \frac{d^2}{dx^2}\frac{\partial F}{\partial u''} \right) \delta u \, dx + \left[\frac{\partial F}{\partial u'}\delta u + \frac{\partial F}{\partial u''}\delta u' - \frac{d}{dx}\frac{\partial F}{\partial u''}\delta u \right]_{x=a}^{x=b} \tag{2.20}$$

The stationarity condition becomes the fourth-order differential equation

$$\frac{d^2}{dx^2}\frac{\partial F}{\partial u''} - \frac{d}{dx}\frac{\partial F}{\partial u'} + \frac{\partial F}{\partial u} = 0 \tag{2.21}$$

supplemented by two natural boundary conditions on each end,

$$\delta u' \frac{\partial F}{\partial u''} = 0, \quad \delta u \left[\frac{\partial F}{\partial u'} - \frac{d}{dx}\frac{\partial F}{\partial u''} \right] = 0 \quad \text{at } x = a \text{ and } x = b \tag{2.22}$$

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

Example 2.1.4 The equilibrium of an elastic bending beam correspond to the solution of the variational problem

$$\min_{w(x)} \int_0^L (\frac{1}{2}(E(x)w'')^2 - q(x)w)dx \tag{2.23}$$

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 (2.21) becomes

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

The equations (2.22) 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 bending momentum vanishes here as well).

Generalization The Lagrangian

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

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

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

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

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

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

2.2 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.

2.2.1 Quadratic penalties

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

$$\min_u J(u), \quad J(u) = \int_a^b \frac{1}{2} \left(\alpha u'^{\,2} + (h-u)^2 \right) dx \tag{2.24}$$

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 (2.25)) 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. Before the discussion of methods for solving this problem, we introduce several other penalty functionals

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

$$\min_u J(u), \quad J(u) = \int_a^b \frac{1}{2} \left(\alpha (u'')^2 + (h-u)^2 \right) dx \tag{2.25}$$

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 (2.25)) 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.

2.2.2 Supplement: Linear boundary-value problem

The Euler equation for the approximation problem is described as a linear problem

$$\mathcal{L}_\alpha(u) = f$$

where $\mathcal{L}$ is the linear operator of the Sturm-Liuville type, which consists of the homogeneous linear differential equation in the interval $[a, b]$ of the approximation, and the homogeneous boundary conditions posed at its ends. The approximating function f stays in the right-hand side of the equation.The operator $\mathcal{L}_\alpha$ depends on the magnitude α of the penalty, and becomes identical operator with $\alpha \to 0$,

$$\lim_{\alpha \to 0} \mathcal{L}_\alpha(u) = u$$

The problem of an optimal approximation becomes the problem of inverting of this operator:

$$u = \mathcal{L}_\alpha^{-1} f$$

Here we describe two approaches to this problem. The first of them is based on the Green's function – the resolvent for the operator $\mathcal{L}_\alpha$. The solution assumes the form

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

where $G(x, y)$ is the Green's function for the operator $\mathcal{L}$. The second method is besed on the expansion of u into series of *eigenfunctions* of the operator λ_α.

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

$$L(x)u(x) = f(x) \quad x \in [a, b], \quad B_a(u, u')|_{x=a} = 0, \quad B_b(u, u')|_{x=a} = 0. \tag{2.26}$$

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 (2.26) one solves first the problem for a single concentrated load $f_0 = \delta(x-\xi)$ applied at the point $x = \xi$

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

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. This problem is simpler than (2.26) because it correspond to the simplest right-hand-side term. Formally, the Green's function can be expressed as

$$g(x,\xi) = L(x)^{-1}\delta(x-\xi) \tag{2.27}$$

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 (2.26). We multiply both sides of (2.27) 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-dimensional 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 $e^{\frac{x-\xi}{\alpha}}$.

Solution by expansion into series of eigenfunctions To solve the problem, we find the specter of the operator $L\alpha = a^2 u'' - u$ solving the Sturm-Liouville problem:

$$\mathcal{L}_\alpha u = \lambda u, \quad u'(0) = u'(1) = 0.$$

This problem has nontrivial solutions ($u \neq 0$ everywhere) for only some special values of λ – the eigenvalues $\lambda_0, \lambda_1, \lambda_2, \ldots$. The solutions – the eigenfunctions $u_n(x)$ – transform the operator $\mathcal{L}$ as

$$\mathcal{L}(u_n) = \lambda_n u_n.$$

The particular solutions are defined up to a multiplier and are mutually orthogonal,

$$\langle u_k u_m \rangle = 0, \quad \text{if } k \neq m$$

where the operation $\langle z \rangle$ means the average

$$\langle z \rangle = \frac{1}{b-a} \int_a^b z \, dx.$$

A general solution has the form

$$u(x) = \sum_{n=0}^{\infty} \alpha_n u_n$$

where α_n are arbitrary coefficients.

With these representations, the equation (**??**) becomes

$$\mathcal{L}(u) = \mathcal{L}\left(\sum_{n=0} \alpha_n u_n \right) = \sum_{n=0} \alpha_n \lambda_n u_n = f$$

The coefficients α_n are found from the orthogonality conditions. We compute

$$\langle u_m f \rangle = \left\langle u_m \sum_{n=0} \alpha_n \lambda_n u_n \right\rangle$$

apply the orthogonality conditions, and obtain

$$\alpha_n \lambda_n \left\langle u_m^2 \right\rangle = \langle u_m f \rangle,$$

or, finally,

$$\alpha_n = \frac{\langle u_m f \rangle}{\lambda_n \langle u_m^2 \rangle}$$

and the approximate u becomes

$$u(x) = \sum_{n=0}^{\infty} \frac{\langle u_m f \rangle}{\lambda_n \langle u_m^2 \rangle} u_n(x)$$

We observe that the approximate linearly depends on f as expected. Also, the eigenvalues λ_n that correspond to fast oscillation eigenfunctions $u_n(x)$ are large, and the contribution of the corresponding harmonic is small because of the multiplier $\frac{1}{\lambda_n}$ in the expansion.

Example For the operator (**??**) we have

$$a^2 u'' - u = \lambda u, \quad u'(0) = u'(1) = 0.$$

Particular solutions (the eigenfunctions) are $u_n = \cos(\pi n x)$ and the eigenvalues are $\lambda_n = (1 - \pi n^2 a^2)$. The general solution is

$$u(x) = \sum_{n=0}^{\infty} \alpha_n \cos(\pi n x)$$

where α_n are arbitrary coefficients. Solving, we obtain

$$u(x) = \sum_{n=0}^{\infty} \frac{c_n}{\lambda_n} \cos(\pi n x), \quad c_n = \int_0^1 \phi(x) \cos(\pi n x) dx$$

The next graph illustrates dependence on a of the approximation of $f(x) = |x|$.

2.2.3 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

$$\min_u J(u), \quad J(u) = \int_a^b \frac{1}{2} \left(\gamma |u'| + (h-u)^2\right) dx \tag{2.28}$$

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:

$$(\gamma \text{sign}\,(u'))' + u = h, \quad \text{sign}\,(u'(a)) = \text{sign}\,(u'(b)) = 0 \tag{2.29}$$

is not too helpful because it requires the differentiation of a discontinuous function sign; besides, the Lagrangian (2.28) 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 (2.28). We fix $\epsilon > 0$, derive the necessary conditions and analyze them assuming that $\epsilon \to +0$. The Euler equation is more regular,

$$(k(u', \epsilon))' = h - u + O(\epsilon), \quad u'(a) = u'(b) = 0 \tag{2.30}$$

where

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

Remark 2.2.1 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 (2.30) 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 (2.30) 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 **??**.

2.3 Several minimizers and first integrals

2.3.1 Several minimizers

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

$$I(u) = \min_u \int_a^b F(x, u, u')dx, \quad u(a) = u_a, u(b) = u_b \tag{2.31}$$

where x is a point in the interval $[a, b]$, $u = (u_1(x), \ldots, u_n(x)]$ is a vector function, and u_a and u_b and n-dimensional real vectors. We suppose that F is a twice differentiable function of its arguments.

As before, we 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. In order to compute the Lagrangian at the perturbed trajectory $u+\delta u$, we use the expansion

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

The increment δI is computed as before

$$\delta I(u) = \int_a^b \left(\sum_{i=1}^{n} \left(\frac{\partial F}{\partial u_i} - \frac{d}{dx}\frac{\partial F}{\partial u_i'} \right) \delta u_i \right) dx + \left(\frac{\partial F}{\partial u_i'} \right)^T \delta u \Bigg|_{x=a}^{x=b}$$

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 decrease by one of them. This way, we obtain n stationary conditions; each condition coincides with the one-variable case. The boundary terms vanish,

$$\sum_{i=1}^{n} \frac{\partial F}{\partial u_i'}\delta u_i \Bigg|_{x=a}^{x=b} = 0 \tag{2.32}$$

because the boundary vaues of u are fixed hence $\delta u_i = 0, i = 1, \ldots, n$.

Proceeding as before, we arrive at the theorem.

Theorem 2.3.1 Any twice-differentiable vector minimizer of the functional (2.31) satisfies the system of differential equations of the order $2n$,

$$\frac{d}{dx}\frac{\partial F}{\partial u_i'} - \frac{\partial F}{\partial u_i} = 0, \quad i = 1, \ldots n \tag{2.33}$$

and the boundary conditions $u(a) = u_a,\ u(b) = u_b$.

Remark 2.3.1 Notice the conditions of the theorem: It is assumed that a solution to the system (2.33) exists and is twice-differentiable. These conditions is often hard to check without actually finding the minimizer. Notice also that the theorem says noting about the case when the minimizer is not differentiable.

The vector form of the system (2.33),

$$S_F(u) = \frac{d}{dx}\frac{\partial F}{\partial u'} - \frac{\partial F}{\partial u} = 0, \quad \delta u^T \frac{\partial F}{\partial u'}\Bigg|_{x=a}^{x=b} = 0 \tag{2.34}$$

is analogous to the scalar Euler equation. This system is obtained by simply applying an algebraic definition of differentiation with respect a vector arguments u and u' to the scalar form of the Euler equation.

Example 2.3.1 (Quadratic Lagrangian) Consider the problem with the integrand

$$F = \frac{1}{2}u^T A u + u^T B u' - \frac{1}{2}(u')^T C u' \tag{2.35}$$

where $u = (u_1, \ldots, u_n)^T$ is the vector of minimizers [1], A, B, and C are $n \times n$ matrices. The stationarity conditions is the system of linear second-order equations

$$\begin{aligned} A\,u'' + C\,u = 0 \quad &\text{in } [a,b] \\ (A\,u' + B\,u)^T \delta u|_{x=a}^{x=b} &= 0 \end{aligned}$$

When A is a diagonal positive matrix, and C is nonnegative, the system describes the system of linear oscillators or the system of masses joined by linear springs. We should ask ourselves: What quantity do we need to minimize to arrive at the equation of motion of a physical system?

Problem 2.3.1 Notice that matrix B does not enter the Euler equation. What conditions correspond to the case $A = C = 0$?

2.3.2 First integrals

In several cases, the Euler equation (2.12) 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

$$J(u) = \int_a^b F(x, u)dx \tag{2.36}$$

In this case, the variation does not involve integration by parts, and the minimizer does not need to be continuous. Euler equation (2.12) becomes an algebraic relation for u

$$\frac{\partial F}{\partial u} = 0 \tag{2.37}$$

Curve $u(x)$ is determined in each point independently of neighboring points. The boundary conditions in (2.12) are in a sense irrelevant: they are satisfied by jumps of the extremal $u(x)$ in the end points and these conditions do not affect the objective functional.

Example 2.3.2 Consider the problem

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

[1] The unexpected transposition u^T is used for typographical reason: We need a column vector u but we want to write it neatly in a text line

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 was derived. To be consistent, we need to repeat the derivation of the necessary condition for the problem (2.36) without any assumption on the continuity of the minimizer. This derivation is quite obvious.

Lagrangian is independent of u If Lagrangian is independent of u, $F = F(x, u')$, Euler equation (2.12) can be integrated once:

$$\frac{\partial F}{\partial u'} = \text{ constant} \tag{2.38}$$

The first order differential equation (2.38) 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 (2.38) and determine the constants of integration from the boundary conditions.

Example 2.3.3 Consider the problem

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

The first integral is

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

Integrating, we find the minimizer,

$$u(x) = \sin x + C_1 x + C_0.$$

The constants C_0 and C_1 are found from and the boundary conditions:

$$C_0 = 1, \quad C_1 = -1 - \sin 1,$$

minimizer u_0 and the cost of the problem become, respectively

$$u_0(x) = \sin x - x(\sin 1 + 1) + 1 \quad J(u_0) = \int_0^1 u_0(x) dx = (\sin 1 + 1)^2$$

Notice that the Lagrangian in Example 2.3.2 is the square of difference between the minimizer u and function $\sin x$, and the Lagrangian in Example 2.3.3 is the square of difference of their derivatives. In Example 2.3.2, the minimizer coincides with $\sin x$, and jumps to the prescribed boundary values. The minimizer u in Example 2.3.3 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. The discontinuous 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 **??**.

Lagrangian is independent of x

Theorem 2.3.2 (Constancy of Hamiltionian) If $F = F(u, u')$, equation (2.12) has the first integral:

$$u' \frac{\partial F}{\partial u'} - F = \text{ constant} \tag{2.39}$$

Proof:
Compute the derivative of the left-hand side of (**??**) at the minimizer

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

because of equalities ??????????????????????????

Example 2.3.4 (Linear Oscillator) Consider the Lagrangian

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

The Euler equation is

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

one recognizes the equation of the *linear oscillator*. The Hamiltonian is constant,

$$\hat{H} = u' \frac{\partial F}{\partial u'} - F = (u')^2 + \omega^2 u^2 = C^2 = \text{ constant}$$

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

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

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

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

We have shown that $\hat{H}$ is constant at the optimal trajectory. In this mechanical problem, $\hat{H}$ is the whole energy of the oscillator.

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

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

$$\hat{H} = u'^T \frac{\partial F}{\partial u'} - F = \text{constant} \tag{2.40}$$

Here

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

and $\hat{H}$ is the Hamiltonian of the system with a vector minimizer. Again, the Hamiltonian stays constant along the optimal trajectory if the Lagrangian F is independent of x.

For the Lagrangian in Example 2.3.1, this first integral is

$$\hat{H} == \frac{1}{2} u^T A u + \frac{1}{2} (u')^T C u';$$

it expressed the sum of kinetic energy of the masses and elastic energy of the springs. The whole energy of this system remains constant. Notice that $\hat{H}$ is independent of matrix B in Example 2.3.1.

Remark 2.3.2 (First integrals) The analogs of the first integrals that are established for the special cases of the scalar Euler equation are also valued for the vector equation.

These three cases do not exhaust all possible first integrals for vector case; for instance one can try to find new invariants by changing the variables. The theory of first integrals will be briefly discussed below in Section ??.

2.4 Geometric optics and Geodesics

2.4.1 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

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

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

$$\psi(y) = -c\sqrt{1+y'^2} \tag{2.41}$$

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

$$x = \pm\Phi(u) = \int \frac{c\,dy}{\sqrt{\psi^2(y)-c^2}} \tag{2.42}$$

Notice that equation (2.41) 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 (2.41) assumes the form

$$\psi(y)\ \cos\alpha = c \tag{2.43}$$

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.

2.4.2 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.

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 (2.41), 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 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

$$\begin{aligned} x &= x_0 + a(\theta - \sin\theta), \\ y &= y_0 + a(1 - \cos\theta), \end{aligned} \tag{2.44}$$

We recognize the equation of the cycloid in (2.44). 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 2.4.1 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 3.4.3.

Remark 2.4.2 Notice that brachistochrone is in fact solution to the problem of optimal design: the trajectory must be chosen by a designer to minimize the time of travel.

2.4.3 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

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

This problem is again a special case of the geometric optic, corresponding to $\psi(y) = y$. Equation (2.42) 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 (**??**) becomes

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

The family of extremals with various C 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 from the common sense viewpoint 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 (2.43) that takes the form

$$y\cos\alpha = C$$

admits either the regular solution $y = C\sec\alpha$, $C \neq 0$ which yields to the catenoid (**??**), 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 **??**.

2.4.4 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:

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

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 (**??**))

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

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 that the geodesics is a great circle.

Remark 2.4.3 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 $g_{ij}(x_1, x_2)$ is called the metric tensor. The problem of geodesics is: Minimize the path

$$\int ds = \int \sqrt{g_{ij}(x_1, x_2)\dot{x}_i \dot{x}_j}\, dt.$$

Here, $x_i = x_i(t)$ is the parameterized path at the surface. 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

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

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

Chapter 3

Developments

This Chapter deals with analysis of the Euler equation, its generalizations to the vector case, and accounting for various boundary conditions. The given examples illustrate the algebra of the derivation. Also, we introduce two areas of application. The homogenization theory deals with solutions of equations with fast oscillating coefficients. The Lagrangian mechanics describe the motion of systems of particles that satisfy Newton's laws of motion.

3.1 Structure of Euler equations

3.1.1 Canonic form

The structure of Euler equations (2.31)

$$\frac{d}{dx}\frac{\partial L}{\partial u_i'} - \frac{\partial L}{\partial u_i} = 0, \quad i = 1, \ldots, N \tag{3.1}$$

can be simplified and unified if we rewrite them as $2N$ first-order differential equations instead of N second-order ones. A first-order system can be obtained from (3.1) if the new variables p_i are introduced,

$$p_i(x) = \frac{\partial L(x, u, u')}{\partial u_i'}, \quad i = 1, \ldots, N \tag{3.2}$$

In mechanics, $p = (p_1, \ldots, p_N)$ is called the vector of *impulse.* The Euler equation takes the form

$$p' = \frac{\partial L(x, u, u')}{\partial u} = f(x, u, u'), \tag{3.3}$$

where f is a function of x, u, u'. The system (3.2), (3.3) becomes symmetric with respect to p and u if we algebraically solve (3.2) for u' as follows:

$$u' = \phi(x, u, p), \tag{3.4}$$

and substitute this expression (3.3):

$$p' = f(x, u, \phi(x, u, p)) = \psi(x, u, p) \tag{3.5}$$

where ψ is a function of the variables u and p but not of their derivatives.

In summary, system (3.2), (3.3) is transferred to the canonic from (or Cauchy form)

$$\begin{array}{l} u' = \phi(x, u, p) \\ p' = \psi(x, u, p) \end{array} : \tag{3.6}$$

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 3.1.1 The equation (3.2) can be solved for u' and (3.4) can be obtained if the Lagrangian is convex function of u' of a superlinear growth. As we will see, (Chapter (**??**)), this condition is to be satisfied if the problem has a classical minimizer.

Example 3.1.1 (Quadratic Lagrangian) Assume that

$$L = \frac{1}{2}a(x)u'^2 + \frac{1}{2}b(x)u^2.$$

We introduce p as in (3.2)

$$p = \frac{\partial L(x, u, u')}{\partial u'} = au'$$

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.

The equations of Lagrangian mechanics (see below Section **??**) correspond to stationarity of the action

$$L(t, q, q') = \frac{1}{2}\dot{q}^T R(q)\dot{q} - V(q)$$

where R is the matrix of inertia, and $V(q)$ is a convex function called the potential energy. The impulses $p = \frac{\partial L}{\partial \dot{q}}$ are equal to $p = R(q)\dot{q}$. The canonic system becomes

$$\begin{array}{l} \dot{q} = R^{-1}p, \\ \dot{p} = p^T R^{-1} \frac{dR}{dq} R^{-1} p - \frac{\partial V}{\partial q} \end{array}$$

The last equation is obtained by excluding $\dot{q}$ from the $\frac{\partial L}{\partial q} = \dot{q}^T \frac{dR}{dq}\dot{q} - \frac{\partial V}{\partial q}$.

3.1.2 Hamiltonian

We can rewrite the system (3.6) in a more symmetric form introducing a special function called *Hamiltonian*. The Hamiltonian is defined by the formula (see (3.6)):

$$H(x,u,p) = pu'(x,u,p) - L(x,u,u'(u,p)) = p\phi(x,u,p) - L(x,u,\phi(x,u,p)) \tag{3.7}$$

where u is a stationary trajectory – the solution of Euler equation. If the original variables u, u' are used instead of instead of u, p the expression for the Hamiltonian becomes

$$\hat{H} = u'\frac{\partial F}{\partial u'} - L(x,u,u') \tag{3.8}$$

To distinguish the function (3.8) from the conventional expression (**??**) for the Hamiltonian, we use the notation $\hat{H}$.

Let us compute the partial derivatives of H:

$$\frac{\partial H}{\partial u} = p\frac{\partial \phi}{\partial u} - \frac{\partial L}{\partial u} - \frac{\partial L}{\partial \phi}\frac{\partial \phi}{\partial u}$$

By the definition of p, $p = \frac{\partial L}{\partial u'} = \frac{\partial L}{\partial \phi}$, hence 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

$$p' = -\frac{\partial H}{\partial u} \tag{3.9}$$

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

$$u' = \frac{\partial H}{\partial p} \tag{3.10}$$

The system (3.9), (3.10) is called the canonic system, it is remarkable symmetric.

Introducing $2n$ dimensional vector (u,p) of the variables, we combine the equations (3.9) and (3.10) as

$$\frac{d}{dx}\begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \nabla H, \quad \nabla H = \begin{pmatrix} \frac{\partial}{\partial u} \\ \frac{\partial}{\partial p} \end{pmatrix} H \tag{3.11}$$

3.1.3 The first integrals through the Hamiltonian

The expressions for the first integrals become very transparent when expressed through impulses the Hamiltonian. If the Lagrangian is independent of u_i', the

Hamiltonian is also independent of it. The first integral is $\frac{\partial L}{\partial u'_i} =$ constant. In our notations, this first integral reads

$$p_i = \text{ constant}, \quad \text{if } H = \text{ constant}(u_i)$$

Similarly, we observe that

$$u_i = \text{ constant}, \quad \text{if } H = \text{ constant}(p_i)$$

Lagrangian is independent of x

Theorem 3.1.1 (Constancy of Hamiltionian) If $F = F(u, u')$, equation (2.12) has the first integral:

$$\hat{H}(u,p) = \text{ constant} \tag{3.12}$$

Proof:

Compute the derivative of the Hamiltonian at the miminizer

$$\frac{d}{dx}H(x,u,p) = \frac{\partial H}{\partial x} + \frac{\partial H}{\partial u}u' + \frac{\partial H}{\partial p}p' = \frac{\partial H}{\partial x}$$

because of equalities (3.9), $u' = \frac{\partial H}{\partial p}$ and (3.10), $p' = -\frac{\partial H}{\partial u}$. If Lagrangian does not depend on x, the Hamiltonian is independent of x as well, $\frac{\partial H}{\partial x} = 0$ and we arrive at (3.12)

In Lagrangian mechanics, the Hamiltonian H is equal to the sum of kinetic and potential energy, $H = T + V$ where $\dot{q}$ is expressed through p, and q, $\dot{q} = R(q)^{-1}p$

$$H(q,p) = \frac{1}{2}p^T R^{-1} p + \frac{\partial}{\partial q}(p^T R^{-1} p + V)$$

Here, we use the equality $\frac{\partial T}{\partial \cdot q} = 2T$ for the kinetic energy, a second degree homogeneous function of $\dot{q}$.

Example 3.1.2 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 the Hamiltonian is

$$H = p\left(\frac{p}{a}\right) - L = \frac{1}{2}\left(\frac{1}{a(x)}p^2 - b(x)u^2\right)$$

and the canonic system is

$$\frac{\partial H}{\partial u} = -b(x)u = -p', \quad \frac{\partial H}{\partial p} = \frac{1}{a(x)}p = u'$$

which coincides with the previous example.

3.1.4 Hamiltonian for geometric optics

The results of study of geometric optics (Section **??**) 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:

$$y' = \frac{p}{\sqrt{w^2 - p^2}} = \phi(x,y,p), \tag{3.13}$$

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:

$$p' = -\frac{w}{\sqrt{w^2 - p^2}}\frac{dw}{dy} \tag{3.14}$$

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 (**??**))

$$p = \psi(x,y)\sin\alpha, \quad H = -\psi(x,y)\cos\alpha.$$

3.2 Broken extremals and nonfixed interval

3.2.1 Weierstrass-Erdman condition

The classical derivation of the Euler equation requires the existence of all second partial derivatives 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. Consider the question how to join these solutions together.

We always assume that the extremal u is differentiable everywhere so that the first derivative u' exists at all point of the trajectory. However, 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. To illustrate this, consider *the stationarity requirement in the integral form.* Namely, integrating (2.12) with respect to x, we obtain

$$\int_{x_0}^{x} S_F(u)dx = \int_{x_0}^{x}\left(\frac{d}{dx}\frac{\partial F}{\partial u'} - \frac{\partial F}{\partial u}\right)dx = 0$$

or

$$\frac{\partial F}{\partial u'} = \int_{x_0}^{x}\frac{\partial F}{\partial u}dx \tag{3.15}$$

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 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:
>
> $$\left[\frac{\partial F}{\partial u'}\right]_-^+ = 0 \quad \text{along the optimal trajectory } u(x) \tag{3.16}$$
>
> Here $[z]_-^+ = z_+ - z_-$ denotes the jump of the variable z.

Another way to derive the Weierstrass-Erdman conditions requires a division of the interval of integration $[a,b]$ into two, $[a,b] = [a,x_0]\cup[x_0,b]$ and computing the stationarity of the left and the right part of the trajectory. Doing this, we integrate by parts as and obtain the out-of-integral terms at the point x_0 of breakage as in (2.10)

$$\left[\frac{\partial F}{\partial u'}\right]_+ \delta u|_{x=x_0+0} + \left[\frac{\partial F}{\partial u'}\right]_- \delta u_{x=x_0-0}$$

The trajectory is continuous at the point of breakage, therefore $\delta u(x_0) = \delta u|_{x=x_0+0} = -\delta u_{x=x_0-0}$. The value of u is not prescribed, therefore the the coefficient by the variation $\delta u(x_0)$ must vanish; the condition (3.16) follows.

Example 3.2.1 (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.

Example 3.2.2 (Snell's law of refraction) Consider again the problem of geometric optics. 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

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

????????????????????????

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 (**??**) 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.

3.2.2 Paradox of geometric optics

Consider propagation of light through a glass barrier shaped as a parallelepiped. The light's path joints the light source (at the origin) and the target that is located at a line parallel to the glass barrier (see Figure..). Assume for simplicity that the barrier is cut along the line of a ray's path.

When the line between the source and target is located outside the barrier (above the corner), it represent the shortest path. When the line crosses the barrier that path has a kink. It reaches the targets that are distant from the corner path. Question: what is the shortest path to the points just below the corner.

3.2.3 Variable interval. Transversality condition

First variation Consider again the variational problem (**??**). Consider now the case when the interval (a, b) of integration is not fixed, and its end point b is to be chosen to minimize the functional, together with the minimizer $u(x)$, $x \in [a, b]$. Assume for simplicity in notations,, that point a is fixed.

The optimal value of the end point b is found by the variational method. Consider the variation of b which assumes the position $b + \delta x$. The variation of the cost of the variational problem (**??**) is the difference

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

caused by the variation δx of the interval and the related variation δu of the minimizer. The first integral in the right-hand side is computed over the varied interval $[a, b + \delta x]$, and the minimizer $u(x)$ in it is replaced by and admissible function $u(x) + \delta u(x)$ where $\delta u(x)$ is a small variation. Parting the interval of integration in the first integral into the intervals $[a, b]$ and $[b, b+\delta x]$, we compute

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

The last expression includes three variations: variation $\delta u(x)$ of the trajectory in the inner points $x \in (a, b)$ of the interval, variation $\delta u|_b$ at the boundary point b and variation δx of the length of the interval. All these variations are small. As before, we compute the increment by expanding the increment into the Taylor series and Keeping only the linear terms with respect to the variations.

We simplify the first integral in the right-hand side by using the same technique as in Section **??**,

$$\int_a^b [F(x, u + \delta u, u' + \delta u') - F(x, u, u')]dx = \int_a^b S(F(u))\delta u(x)\, dx + \frac{\partial F}{\partial u'}\delta u|_b$$

and compute the second integral using the smallness of interval δx

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

Thus, the linear part of the increment is split into three parts

$$\delta I = \int_a^b S(F(u))\delta u(x)\, dx + \frac{\partial F}{\partial u'}\delta u|_b + \delta x F(x, u, u') \tag{3.17}$$

Stationarity of the first term requires that the Euler equation $S(F(u)) = 0$ is satisfied for all $x \in (a, b)$. This second-order differential equation fixes the trajectory up to boundary conditions. We also have assumed that the boundary condition at the unvaried left side of the interval are fixed, $u(a) = u_a$. Therefore the stationary trajectory $u(x)$ depends on one parameter – the boundary condition at the end point, $u = u(x, C)$. We also implicitly assume that the trajectory that $u(x, C)$ can be defined at an interval $[a, d]$ larger than $[a, b]$. $(d > b)$.

The stationarity of the non-integral terms of (3.17) defines both the boundary condition at $u(b)$ and the length b of the interval. The variation $\delta u|_b$ is the difference between the value of a varied $(u + \delta u)$ and optimal (u) trajectories at the point $x = b$. Recall that the varied trajectory is defined on a larger interval $[a, b+\delta x]$ and b is not its end point. This variation should be expressed through variation δx of the interval and variation δu_{end} of the trajectory $u + \delta u$ at the end $b + \delta x$ of its interval. The variation δu_{end} of the trajectory on its end point is computed as the difference

$$\delta u_{\text{end}} = (u + \delta u)|_{b+\delta x} - u|_{x=b}$$

between the values of the varied trajectory $u + \delta u$ at its end point $b + \delta x$ and the stationary trajectory u at its end b. This variation δu_{end} can be either set free or be subjected to the constrains. For example, the condition $u(b) = u_b$ which fixes u at the end point implies that $\delta u_{\text{end}} = 0$.

We need to connect δu_{end} and the variation $\delta u|_b$ in the right-hand side of (3.17). We use the expansion

$$\delta u_{\text{end}} = (u + \delta u)|_{x=b+\delta x} = u|_{x=b} + \delta u|_{x=b} + u'|_{x=b}\delta x = o(\|v\|) + o(|\delta x|).$$

We keep only linear term and rewrite it as

$$\delta u|_{x=b} = \delta u_{\text{end}} - u'(b)\delta x.$$

The derived representation together with the Euler equation allows for presenting the stationarity of δI in (3.17) in the form

$$\delta I = \frac{\partial F}{\partial u'}\delta u_{\text{end}} + \left[F(x,u,u') - u'\frac{\partial F}{\partial u'}\right]_{x=b} \delta x = 0 \tag{3.18}$$

The first term is the familiar variational boundary condition at $x = b$. The second term gives an additional condition at the unknown end point. The variation δx is arbitrary; the stationarity with respect to it requires the additional *transversality condition* at the end point b

$$\left.\left[F(x,u,u') - u'\frac{\partial F}{\partial u'}\right]\right|_{x=b} = 0 \tag{3.19}$$

that is used to determine the unknown length of the interval.

One recognizes the negative of Hamiltonian (see (**??**) in the left-hand side expression. The transversality condition can be rewritten in the form

$$H(x,u,p)_{x=b} = 0 \tag{3.20}$$

Special cases Suppose first that no conditions on the end point are imposed and the variation δu_{end} is arbitrary. Stationarity condition requires the satisfaction of natural variational condition

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

The natural condition simplifies the transversality condition which becomes

$$F(x,u,u')|_{x=b} = 0 \tag{3.21}$$

Example 3.2.3 Consider the problem

$$\max_{u(x),b} \int_0^b F(x,u,u')dx, \quad F = \frac{1}{2}(u')^2 + u, \quad u(0) = \alpha$$

Here parameter $b > 1$ and function $u(x), x \in [0, b]$ are the unknowns.

The Euler equation
$$u'' - 1 = 0, \quad u(1) = \alpha$$
produces the family of the solutions
$$u(x) = \frac{1}{2}x^2 + cx + \alpha.$$
The two remaining unknowns – parameter c and the end point b are found from the conditions at the free end. The natural boundary condition
$$\left.\frac{\partial F}{\partial u'}\right|_{x=b} = u'|_{x=b} = b + c = 0$$
gives $c = -b$, so that $u(x) = \frac{1}{2}x^2 - bx + \alpha$. The transversality condition in the form (3.21)
$$F = \left.\left(\frac{1}{2}(u')^2 + u\right)\right|_{x=b} = -\frac{1}{2}b^2 + \alpha = 0$$
gives $b = \sqrt{2\alpha}$.

We find that the optimal interval is $[0, \sqrt{2\alpha}]$ and the minimizer is
$$u(x) = x^2 - \sqrt{2\alpha}\,x + \alpha.$$

2. If the Lagrangian is x-independent, the Euler equation admits the first integral. The Hamiltonian H is constant along the optimal trajectory, see (3.12),
$$H(u, p) = C$$
If, in addition, the length of the interval is undetermined, this condition becomes
$$H(u, p) = 0 \quad \text{along the optimal tranjectory}$$

3. The boundary condition $u(b) = u_b$ can be imposed at the unknown end point b which replaces the natural boundary condition.

Example 3.2.4 Consider the problem
$$\max_{u(x),b} \int_0^b F(x, u, u')dx, \quad F = (u')^2 + u^2, \quad u(1) = 1,\ u(b) = A,$$
for unknown parameter $b > 0$ and function $u(x), x \in [1, b]$.

The Lagrangian is x-independent, therefore the Hamiltonian is equal zero along an optimal trajectory
$$\hat{H}(u, u') = u'\frac{\partial F}{\partial u'} - F = (u')^2 - u^2 = 0.$$
Integration gives $u' = \pm u$. Applying the boundary conditions and integrating, we find that
$$\begin{array}{ll} u(x) = \exp(x), & b = \log A, \quad \text{if } A > 1; \\ u(x) = \exp(-x), & b = -\log A, \quad \text{if } A < 1. \end{array}$$

Non-integral term Another generalization is the problem with the functional (2.17) with an additional non-integral cost $f(x,u)$. Variation δf of this term gives

$$\delta f = f(b+\delta x, (u+\delta u)|_{x=b+\delta x}) - f(b, u|_{x=b}) = \left.\left(\frac{\partial f}{\partial x} - \frac{\partial f}{\partial u}u'\right)\right|_{x=b} \delta x + \frac{\partial f}{\partial u}\delta u_{\text{end}}$$

(recall that $\delta u_{\text{end}} = (u+\delta u)|_{x=b+\delta x} - u_{x=b}$) The variation of the functional in the problem (2.17) is the sum of the variations of the integral and the non-integral terms.

For this problem, the natural boundary condition coincides with (2.18). The transversality condition becomes

$$\left.\left[F(x,u,u') + \frac{\partial f}{\partial x} - u'\left(\frac{\partial F}{\partial u'} + \frac{\partial f}{\partial u}\right)\right]\right|_{x=b} = 0$$

The derivation of these conditions can be easily done using the Null-Lagrangian approach transforming the boundary terms into integral ones as in (2.19).

Main boundary condition A variant of the free boundary problem asks to an optimal trajectory that ends at the curve $\phi(u,x) = 0$. In this case, the end point variations δu_{end} and δx are related:

$$\frac{\partial \phi}{\partial u}\delta u_{\text{end}} + \frac{\partial \phi}{\partial x}\delta x = 0$$

This condition together with the stationary (3.18) of the increment δI results in a linear system for the variations δu_{end} and δx

$$\begin{aligned} -\hat{H}\delta u_{\text{end}} + p\delta x &= 0 \\ \frac{\partial \phi}{\partial x}\delta u_{\text{end}} + \frac{\partial \phi}{\partial u}\delta x &= 0 \end{aligned}$$

or

$$A\,z = 0, \quad z = \begin{pmatrix} \delta u_{\text{end}} \\ \delta x \end{pmatrix}, \quad A = \begin{pmatrix} -H(u,p) & p \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial u} \end{pmatrix}$$

This linear homogeneous system must admit a nonzero solution for z, which yields to the optimality condition

$$\det A = 0. \tag{3.22}$$

Example 3.2.5 (Distance to a curve) Find the path of minimal length between the origin and the curve $y = \phi(x)$.

Assume that the path is the graph of a function $y(x)$. The length L of the path is

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

The path of the minimal length d is found from

$$d = \min_{y(x)} L(y), \quad y(0) = 0, y(x_{\text{end}}) = \phi(x_{\text{end}})$$

The Lagrangian $F = \sqrt{1 + y'^2}$ is independent of x, therefore the first integral exists

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

which implies that $y'(x) = \text{ constant}$ and the trajectory is a straight line, as expected. The condition (3.22) determines the slop of this straight line,

$$\det \begin{pmatrix} \frac{-1}{\sqrt{1+y'^2}} & \frac{y'}{\sqrt{1+y'^2}} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} \end{pmatrix} = 0 \quad \text{or} \quad -\frac{1}{\sqrt{1 + y'^2}} \left(\frac{\partial \phi}{\partial y} + y' \frac{\partial \phi}{\partial x} \right) = 0$$

This condition states that the extremal $y(x)$ is orthogonal to the curve $\phi(u, x)$ at the point where they meet,

$$\frac{\partial \phi}{\partial y} \left(\frac{\partial \phi}{\partial x} \right)^{-1} = -y'$$

3.2.4 An extremal broken at the unknown point

Combining the two above techniques, we derive optimality conditions at the extremal broken in an unknown point, when the position of this point is determined from the minimization requirement. Consider the Lagrangian in 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, $\xi \in [a, b]$ is an unknown point in the interval $[a, b]$ of 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 $x = \xi$ are

$$\frac{\partial F_+}{\partial u'} = \frac{\partial F_-}{\partial u'} \quad \text{or } p(\xi_-) = p(\xi_+) \tag{3.23}$$

where p is the impulse, (the stationarity of the slope u' of the trajectory at the transit point) and

$$H|_{x=\xi_-} = H|_{x=\xi_+} \tag{3.24}$$

(the stationarity of the position ξ 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 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 (3.24) express the continuity of the Hamiltonian at the point ξ of breakage.

Example 3.2.6 Consider the problem with the Lagrangian

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

and boundary conditions

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

The Euler equation is

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

The solution to this equation that satisfies the boundary conditions is

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

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

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

(2) the Weierstrass-Erdman condition

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

(3) the transversality condition

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

The transversality condition is simplified to

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

From the Weierstrass-Erdman condition, we have

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

The condition of the continuity of the extremal 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

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

3.2.5 Boundary conditions for a vector minimizer

Variational boundary conditions The variational condition (2.32) which we rewrite here for convenience

$$\left.\frac{\partial F}{\partial u_1'}\delta u_1 + \ldots + \frac{\partial F}{\partial u_n'}\delta u_n\right|_{x=a}^{x=b} = 0 \tag{3.25}$$

produces $2n$ boundary conditions for the Euler equations (2.33). 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 (3.25) is satisfied, because all variations are zero. 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

$$\text{Either } \left.\frac{\partial F}{\partial u_i'}\right|_{x=a,b} = 0 \quad \text{or } \delta u_i|_{x=a,b} = 0. \tag{3.26}$$

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

Consider a general case when $p < 2n$ boundary conditions of the form

$$\beta_k(u_1(a), \ldots, u_n(a), u_1(b), \ldots, u_n(b),) = 0 \tag{3.27}$$

are prescribed and the both end points $x = a$ and $x = b$. We need to find $2n - p$ supplementary variational constraints at these points that together with (3.27) give $2n$ boundary conditions for the Euler equation (2.32) of the order $2n$.

The conditions (3.27) 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 $u(b)$. It is defined as follows:

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

The variation δw_i is constraint by a linear system

$$\frac{\partial \beta_k}{\partial w}^T \delta w = 0, \quad k = 1, \ldots, p \tag{3.28}$$

These conditions have a matrix form,

$$P\delta w = 0,$$

where

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

is $n \times p$ matrix. We also assume that these conditions are linearly independent, rank P =p.

The solution to the constraints (3.25) is a $(2n-p)$-dimensional vector δw_{ad} of the form

$$\delta w_{\text{ad}} = Qv$$

where v is arbitrary $(2n-p)$-dimensional arbitrary vector (potential) and $(2n-p) \times n$ matrix Q that is orthogonal P: $PQ = 0$.

One can choose Q as a projector onto the zero subspace of P

$$Q = I - P^T (P\,P^T)^{-1} P \tag{3.29}$$

Here, I is the unit $2n \times 2n$ matrix. Indeed, one can check that he condition $PQ = 0$ is satisfied. In addition, matrix Q has the eigenvalues equal either to one of to zero; its rank is equal to $2n - p$.

Any admissible variation δw_{ad} makes the fist variation (2.34) of the objective functional vanish; correspondingly, we have

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

Using the representation of δw_{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. Using the definition (3.29) of Q, we obtain the variational conditions in the form

$$\left(\frac{\partial F}{\partial u'}\right)^T \left(I - P^T (P\,P^T)^{-1} P\right) = 0. \tag{3.30}$$

This representation provides the $2n - p$ linearly independent boundary conditions. Together with p imposed conditions (3.25), we obtain exactly $2n$ needed boundary conditions.

Example 3.2.7 Consider again the variational problem with the Lagrangian (2.35) 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 supplementary variational condition is

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

Since $u_1(a) = 1$, the condition 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 ($\delta u = Qv$)

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

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

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

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.

Periodic boundary conditions The variational boundary conditions for the problem with a periodic solution are obtained from the expression (3.25) of the variation of the functional when we use the equalities $\delta u(a) = \delta u(b)$. These conditions have the form

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

Boundary conditions at the unknown end point If the minimizer is a vector, the transversality condition (**??**) preserves its form, but the term $u'\frac{\partial F}{\partial u'}$ becomes the scalar product of the vectors u' and $\frac{\partial F}{\partial u'}$. The consideration of this case is quite analogous to the scalar case. Likewise, the next conditions preserve their form if u is a vector-function.

The case when the minimizer $u = u_1, \ldots, u_n$ is an n-dimensional vector function that meets the end manifold described by a k-dimensional condition $\phi_i(x, u) = 0$ $i = 1, \ldots, k$ is handled similarly. At the end of an optimal trajectory, the stationarity condition (see (3.18))

$$\delta I = \sum_{k=1}^{n} \frac{\partial F}{\partial u_k'} z_k + \left[F(x, u, u') - u'\frac{\partial F}{\partial u'}\right]_{x=b} z_{n+1} = 0$$

is satisfied for all $n + 1$ dimensional vectors $z = [\delta u_{\text{end}}, \delta x]$ which satisfy the homogeneous linear system

$$Rz = 0, \quad \text{where } R_{ij} = \frac{\partial \phi_i}{\partial u_j}, \ j = 1, \ldots, n, \quad R_{i,n+1} = \frac{\partial \phi_i}{\partial x}.$$

Here. R is $(n + 1) \times k$ matrix that constraints the variations. The optimality conditions coincide with (3.30) where P is replaced by R.

3.3 Fast oscillating coefficients. Homogenization

The canonic form allows for handling of Lagrangians with fast oscillating coefficients of the type $L(\frac{x}{\epsilon}, u, u')$ where ϵ is a "small parameter" Physically, it may represent the scale of fast oscillating material properties or a fast external excitation. In such problems, it is important to describe the evolution of an average $\hat{u}$

$$u_\delta(x) = \frac{1}{0}2\delta \int_{x-\delta}^{x+\delta} u(\xi)\, d\xi$$

where δ – the interval of the average – it larger than the scale ϵ of oscillations but smaller than the interval $[a, b]$ of the variation of x. The average is defined in subinterval $[a+\delta, b-\delta]$ of $[a, b]$. The averaged variables are denoted by a subindex $_\delta$.

In many cases, it is possible to show that the average $u_\delta((x) =$ satisfies the new variational problem that is derived from the original one using the procedure called *homogenization.* For such problem, it is convenient to use the mentioned continuity of the canonic variables u and p.

Consider the stationarity conditions in the *canonic* form (3.4) and (3.5). Assume that the functions ϕ and psi in the right-hand-sides of them are of Lipschitz type, e.a

$$|\phi(x, a, b) - \phi(y, c, d)| \leq C(|x - y| + |a - c| + |b - d|).$$

Then the derivatives $u' = \phi$ and $p' = psi$ are bounded and their variation over a small interval is proportional to the length of that interval,

$$|u(\xi) - u(x)| \leq C_1\delta, \quad \forall \xi \in [x = \delta, x + \delta]$$

When this length 2δ is small, we may pass to the homogenized stationarity condition for the smooth variables u_δ and p'_δ

$$\begin{aligned} u'_\delta &= \phi_\delta(x, u_\delta, p_\delta) + o(\delta), \\ p'_\delta &= \psi_h(x, u_\delta, p_\delta) + o(\delta), \end{aligned}$$

where

$$\begin{aligned} \phi_h(x, u_\delta, p_\delta) &= \frac{1}{0}2\delta \int_{x-\delta}^{x+\delta} \phi(\xi, u_\delta, p_\delta) d\xi \\ \psi_h(x, u_\delta, p_\delta) &= \frac{1}{0}2\delta \int_{x-\delta}^{x+\delta} \psi(\xi, u_\delta, p_\delta) d\xi \end{aligned}$$

are the averaged functions ϕ and ψ, respectively, computed with the frozen values of smooth variables u and p.

In turn, the stationarity system corresponds to the homogenized Hamiltonian that is reconstructed from (**??**) and (**??**). Further, one can formulate the homogenized variational problem for the average $\hat{u}$ and its derivative.

Example: A quadratic Lagrangian Let us show an example of the homogenization procedure. Consider again the quadratic Lagrangian

$$F = \frac{1}{2} a\left(\frac{x}{\epsilon}\right) u'^2 + \frac{1}{2} b\left(\frac{x}{\epsilon}\right) u^2.$$

and assume that $a_\epsilon(x)$ and $b_\epsilon(x)$ are ϵ-periodic functions and $a\epsilon(x)$ is nonnegative,

$$a_\epsilon(x) = a_\epsilon(x+\epsilon), \quad b_\epsilon(x) = b_\epsilon(x) \quad a_\epsilon(x) > 0 \quad \forall x$$

Accordingly, the solution $u(x)$ is also an oscillating function. We want to find a variational formulation of the averaged Lagrangian.

Let us average the equations (3.6) 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 $[u']_\epsilon$ is not a smooth or even continuous function of x and it is not clear how to compute the term $[a(x)u'^2]_\epsilon$ which is a product of two oscillatory variables.

To find the homogenized equations, 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 averaged 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 u' is the form

$$[u']_\epsilon = \left[\frac{1}{a(x)}\right]_\epsilon^{-1} p_\epsilon$$

and 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 (homogenized) 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

$$u'_\epsilon = \left[\frac{1}{a(x)}\right]_\epsilon^{-1} p_\epsilon \quad p'_\epsilon = [b(x)]_\epsilon u_\epsilon \tag{3.31}$$

Example 3.3.1 Let us specify the oscillating coefficients $a(x)$ and $b(x)$ as follows

$$a_\epsilon(x) = a_0 + a_1 \sin\left(\frac{x}{\epsilon}\right)$$
$$b_\epsilon(x) = b_0 + b_1 \sin\left(\frac{x}{\epsilon}\right)$$

where $0 < a_1 < a_0$ and $0 < b_1 < b_0$

$$a(x) = \alpha_1 + \alpha_2 \sin^2\left(\frac{x}{\epsilon}\right), \quad b(x) = \beta_1 + \beta_2 \sin^2\left(\frac{x}{\epsilon}\right)$$

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.

3.4 Introduction to 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 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.

3.4.1 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 3.4.1 (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 (**??**) exist

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

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

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

Example 3.4.2 (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,

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

and P is the $n \times (n-1)$ matrix that corresponds to the operation of difference,

$$P = \begin{pmatrix} 1 & -1 & 0 & \ldots & 0 \\ 0 & 1 & -1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix},$$

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

$$\begin{aligned} m_1\ddot{x}_1 &+ C(x_1 - x_2) = 0 \\ m_2\ddot{x}_2 &+ C(-x_1 + 2x_2 - x_3) = 0 \\ \ldots & \quad \ldots \\ m_n\ddot{x}_n &- C(x_{n-1} - x_n) = 0 \end{aligned}$$

or in vector form,

$$\ddot{x} + kP_2x = 0$$

where $k = \frac{C}{m}$ is the positive parameter, and $P_2 = P^TP$ is the $n \times n$ matrix of second differences,

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

3.4.2 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

$$\frac{\partial}{\partial \dot{\boldsymbol{q}}} T_q(\boldsymbol{q}, \dot{\boldsymbol{q}}) \cdot \dot{\boldsymbol{q}} = 2T_q(\boldsymbol{q}, \dot{\boldsymbol{q}}) \tag{3.32}$$

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

$$\min_{\boldsymbol{q}} \int_{t_0}^{t_1} (T_q - V_q)dt \tag{3.33}$$

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

$$\frac{d}{dt}\frac{\partial T_q}{\partial \dot{\boldsymbol{q}}} - \frac{\partial}{\partial \boldsymbol{q}}(T_q - V_q) = 0 \tag{3.34}$$

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 (3.32))

$$\dot{\boldsymbol{q}}\frac{\partial T_q}{\partial \dot{\boldsymbol{q}}} - (T_q - V_q) = T_q + V_q = \text{constant}(t) \tag{3.35}$$

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 3.4.3 (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 3.4.4 (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.

3.4.3 More examples: Two degrees of freedom.

Example 3.4.5 (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 3.4.6 (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

$$\dot\theta\cos\phi=\mu \tag{3.36}$$

(conservation of angular momentum) and

$$m(\dot\phi^2+\dot\theta^2\cos\phi)+g\cos\phi=\Pi \tag{3.37}$$

(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 $\phi(t) = \phi_0 =$ constant, the pendulum oscillates around a horizontal circle. In this case, the speed of the pendulum is constant (see (3.36)) and the generalized coordinate – the angle θ is

$$\theta = \frac{\mu}{\cos \phi_0} t + \theta_0$$

The motion is periodic with the period

$$T = \frac{2\pi \cos \phi_0}{\mu}$$

3.5 Comments

3.5.1 Variational problem as the limit of a vector problem

The variational problem can be considered as a limit of a finite-dimensional minimization problem, if the interval $[a, b]$ is partitioned by the points $u_1, \dots, u_N$, the function $u(x)$ replaced by an unknown vector $(u_1, \dots, u_N)$, integral is replaced by the sum and derivative with finite difference. Solving the finite-dimensional problem and then passing to the limit $N \to \infty$, we should came to the Euler-Lagrange equation.

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 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 $\{u_1, \dots u_N\}$.

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

$$\text{Diff}\,(u_i) = \frac{1}{\Delta}(u_i - u_{i-1}), \quad \Delta = \frac{b-a}{N}; \tag{3.38}$$

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

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

$$\min_{u_1, \dots, u_{N-1}} I_N \quad I_N = \sum_{i=1}^{N} F_i(u_i, \text{Diff}\,(u_i)), \quad \text{Diff}\,(z_i) = \frac{1}{\Delta}(z_i - z_{i-1}) \tag{3.39}$$

Compute the stationary conditions for the minimum of $I_N(u)$ treating the vector components u_i as minimizers

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

Assume that u_i is not a boundary point, $i \neq 0$ and $i \neq N$. 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 Diff(u_i), and the second– only through Diff(u_i):

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

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

$$\frac{\partial I_N}{\partial u_i} = \frac{\partial F_i}{\partial u_i} + \frac{1}{\Delta}\left(\frac{\partial F_i}{\partial \mathrm{Diff}\,(u_i)} - \frac{\partial F_{i+1}}{\partial \mathrm{Diff}\,(u_{i+1})}\right) = 0 \tag{3.40}$$

or, recalling the definition (3.38) 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:

$$0 = \frac{\partial F_{N+1}}{\partial \mathrm{Diff}\,(u_{N+1})} \to \frac{\partial F}{\partial u'(a)}, \quad 0 = \frac{\partial F_o}{\partial \mathrm{Diff}\,(u_0)} \to \frac{\partial F}{\partial u'(b)},$$

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 $\{u_k\}$ exists and does not depend on the partition $\{x_1, \dots x_N\}$ if only $|x_k - x_{k-1}| \to 0$ for all k , and the limit of the sequence of finite-dimensional problems. We also assume that the limits of the components of the vector $\{u_k\}$ represent values of a smooth function in the close-by points $x_1, \dots, x_N$. This is a strong assumption; recall, that u_k are solutions of optimization problems with the coefficients that slowly vary with the number k. We need to answer the question: In what case the sequence $\{u_i\}$ of solutions of minimization problems (3.40) with slowly varied coefficients F_i tends varies slowly as well. This is not always true: the minimal point of slightly perturbed curve may vary significantly. For example, the minimum point of $f(x, \epsilon) = \cos x + \epsilon$, $x \in [-\pi, \pi]$ is a discontinuous function of ϵ,

$$\min_{x \in [-\pi, \pi]} f(x, \epsilon) = \begin{cases} -\pi & \text{if } \epsilon < 0 \\ \pi & \text{if } \epsilon > 0 \end{cases}$$

The continuity of the minimal points is related to the convexity of the function.

Moreover, we need to know whether the limit

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

exists that would approximate the derivative of the minimizer; that is not always the case. We briefly address this question later in Chapter **??**.

3.5.2 Euler equation and the true minimizers

Stationary solution and a "true 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. 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 there is no other close-by classical curve that corresponds 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.