Part I

Preliminaries

1
Relaxation of One-Dimensional Variational Problems

This introductory chapter gives a brief review of nonconvex variational problems. We examine stability of solutions to one-dimensional extremal problems associated with ordinary differential equations. The reader familiar with nonconvex variational problems can skip this chapter. However, this material is necessary to understand the one-dimensional analogue of the multidimensional ill-posed problems that are in the focus of this book.

Optimizing design is a variational problem. Such problems ask for the minimization of a functional that measures the quality of a structure choosing by a proper control function (the materials' layout). In this chapter we detect and describe unstable solutions of these *extremal problems* in a one-dimensional setting. The solutions to these problems are characterized by fine-scale oscillations. To deal with these oscillations, we introduce a relaxation procedure. *Relaxation* essentially means the replacement of an unstable optimization problem with another that has a classical (differentiable) solution.

1.1 An Optimal Design by Means of Composites

Let us start with an example that demonstrates why composites appear in optimal design. Here we find an optimal solution using only commonsense arguments.

The Elastic Beam

Consider an elastic beam with variable stiffness $d(x)$. The beam is loaded by the load $f(x)$, and its ends $x = 0$ and $x = 1$ are simply supported. The deflection $w = w(x)$ of the points of the beam satisfies the classical boundary value problem (see, for example (Timoshenko, 1970))

$$\begin{aligned} \theta &= dw'', \quad \theta'' = f, \\ w(0) &= w(1) = 0, \quad \theta(0) = \theta(1) = 0, \end{aligned} \tag{1.1.1}$$

where $\theta = \theta(x)$ is the bending moment and $d = d(x)$ is the material's stiffness at point x. Suppose that the beam can be made of two materials with effective stiffness d_1 and d_2, so that the stiffness takes one of these two values, $d(x) = d_1$ or $d(x) = d_2$ at each point $x \in [0, 1]$. The deflection w depends on the layout $d = d(x)$ and the loading $f = f(x)$: $w = w(f, d)$.

Optimization Problem

Let us state the following optimal design problem: Lay out the given materials with the stiffness d_1 and d_2 along the beam to approximate in the best way some desired function w_* with the deflection $w(f, d)$. Specifically, we want to minimize the square of the L_2-norm of the difference between the actual displacement $w(d, f)$, which depends on the layout $d = d(x)$ and the loading f, and the desired function[1] $w_*(x)$:

$$I = \min_d \int_0^1 (w(d, f) - w_*)^2. \tag{1.1.2}$$

Let us assume that the desired function w_* is the deflection of a homogeneous beam of an intermediate stiffness d_*,

$$d_* = \frac{d_1 + d_2}{2}, \tag{1.1.3}$$

which is subject to the same boundary conditions and the same loading f: $w_* = w_*(d_*, f)$. The deflection w_* satisfies the equation

$$\begin{aligned} \theta &= d_*(x) w_*'', \qquad \theta_*'' = f, \\ w_*(0) &= w_*(1) = 0, \quad \theta_*(0) = \theta_*(1) = 0, \end{aligned} \tag{1.1.4}$$

similar to (1.1.1). The optimization problem becomes

$$I = \min_{d(x)} \int_0^1 (w(d, f) - w(d_*, f))^2. \tag{1.1.5}$$

[1]The symbol "dx" of the differential is omitted in the integrals over the explicitly defined interval of the independent variable x.

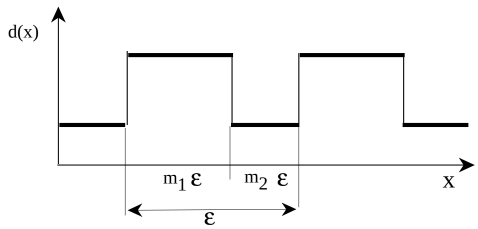

FIGURE 1.1. Oscillation of the pointwise stiffness of an optimal inhomogeneous beam.

The Minimizing Sequence

The solution to the optimization problem (1.1.4) is intuitively obvious: One should mix the given materials d_1 and d_2 in special proportions m_1 and $m_2 = 1 - m_1$ to imitate the intermediate stiffness d_* of the beam and therefore to make the nonnegative cost I (see (1.1.2)) arbitrarily close to zero. The stiffness becomes an oscillatory function of x that alternates between the values d_1 and d_2. The approximation improves when the frequency of the oscillations increases. Therefore, an optimal design does not exist: the higher the frequency, the better (see Figure 1.1). Formally, the minimizing layout of the material corresponds to the limit $\lim_{\varepsilon\to 0} d_\varepsilon(x)$, where

$$d_\varepsilon(x) = \begin{cases} d_1 & \text{if } x \in [n\varepsilon, (n+m_1)\varepsilon], \\ d_2 & \text{if } x \in [(n+m_1)\varepsilon, (n+1)\varepsilon], \end{cases} \qquad n = 1, \dots N,$$

$\varepsilon \ll 1$ is a small parameter, and $N = \left[\frac{1}{\varepsilon}\right]$. The remaining problem is the computation of the needed proportion m_1. We will demonstrate that $m_1 \neq \frac{1}{2}$, contrary to the intuitive expectation.

Homogenization

This consideration poses the question of an adequate description of rapidly oscillating sequences of control. To describe these sequences we use the method of homogenization, which simplifies the problem: Details of the behavior of minimizing sequences become intractable, and the equations depend only on average characteristics of them.

Let us derive equations for an average deflection $\langle w \rangle$ of the beam. The averaging operator $\langle\ \rangle$ is introduced by the formula

$$\langle z(x) \rangle = \frac{1}{2\varepsilon'} \int_{x+\varepsilon'}^{x-\varepsilon'} z(\xi) d\xi, \tag{1.1.6}$$

where $[x - \varepsilon',\ x + \varepsilon']$ is the interval of the averaging and $z = z(x)$ is the averaged variable.

We suppose that the interval ε' is much larger than the period ε of oscillation of the control but much smaller than the length of the beam:

$$0 < \varepsilon \ll \varepsilon' \ll 1. \tag{1.1.7}$$

Homogenized Equation

To derive the homogenized equation for the average deflection $\langle w \rangle$, we mention that the variable $\theta(x)$ is twice differentiable (see (1.1.1)); therefore it is continuous even if $d(x)$ is discontinuous:

$$\langle \theta \rangle (x) = \theta(x) + O(\varepsilon').$$

This implies that discontinuities in $d(x)$ are matched by discontinuities in $w''(x)$, leaving the product $\theta = d(x)w''(x)$ continuous. Therefore $\langle w''(x) \rangle$ is computed as

$$\langle w''(x) \rangle = \left\langle \frac{\theta(x)}{d(x)} \right\rangle = \theta(x) \left\langle \frac{1}{d(x)} \right\rangle + O(\varepsilon'). \tag{1.1.8}$$

Notice that we take the smooth function $\theta(x)$ out of the averaging because its derivation is arbitrarily small in the small interval of averaging.

Note also that the function $\frac{1}{d(x)}$ takes only two values, and it alternates faster than the averaging (1.1.7). Therefore its average is found (up to terms of the order of ε') as

$$\left\langle \frac{1}{d(x)} \right\rangle = \frac{m_1}{d_1} + \frac{m_2}{d_2}, \quad m_2 = 1 - m_1. \tag{1.1.9}$$

The homogenized equation for the average value $\langle w \rangle''$ of the deflection of the beam can easily be found from (1.1.1), (1.1.8), and (1.1.9):

$$\begin{aligned} \theta &= \left\langle \tfrac{1}{d} \right\rangle^{-1} \langle w \rangle'', & \theta'' &= f, \\ \langle w(0) \rangle &= \langle w(1) \rangle = 0, & \theta(0) &= \theta(1) = 0 \end{aligned} \tag{1.1.10}$$

(these equations are satisfied up to ε'). They are called the *homogenized equations* for the composite beam.

Homogenized Solution

We are able to approximate the desirable deflection w_* by the deflection of an inhomogeneous beam. Comparing (1.1.4) and (1.1.10), we conclude that the solutions to these two equations are arbitrarily close to each other if $\varepsilon' \to 0$ and if the fraction m_1 corresponds to the equality

$$\frac{1}{d_*} = \frac{m_1}{d_1} + \frac{m_2}{d_2}. \tag{1.1.11}$$

Then the cost of (1.1.5) goes to zero together with ε'. Strictly speaking, the minimizing layout does not exist: the smaller the period, the better. The actual minimum of the functional I is not achievable. Nothing bounds the period ε of oscillation of $d(x)$ from zero.

The minimizing sequence corresponds to the optimal volume fraction m_1 that can be found from (1.1.3), (1.1.11):

$$m_1 = \frac{d_1}{d_1 + d_2}.$$

Notice that $m_1 \neq \frac{1}{2}$ and $\langle d(x) \rangle = \frac{d_1^2 + d_2^2}{d_1 + d_1} \neq d_*$.

Remark 1.1.1 *In a general situation the fraction m_1 may vary from point to point. Then the outlined homogenization procedure introduces a smoothly varying quantity $m(x)$ (the volume fraction of the material in the composite) that describes the fine-scale oscillating sequence of control.*

Remark 1.1.2 *The described solution with fine-scale oscillations is an example of so-called chattering controls, which are well known in the theory of one-dimensional control problems (Gamkrelidze, 1962; Young, 1969). Chattering regime of control occurs when the interval of the independent variable* x *is split into infinitely many subintervals, and each of them is characterized by alternation of the value of control. The value of the minimizing functional decreases as the scale of alternation of intervals becomes finer.*

1.2 Stability of Minimizers and the Weierstrass Test

1.2.1 Necessary and Sufficient Conditions

Extremal Problems

Consider an extremal problem:

$$I(u) = \min_{u(x)} \int_0^1 F(x, u(x), u'(x)), \quad u(0) = u_0,\ u(1) = u_1, \tag{1.2.1}$$

where x is a point of the interval $[0,\ 1]$, $u(x)$ is a function that is differentiable almost everywhere in $[0, 1]$, and F is a function of three arguments called the Lagrangian. We assume that the Lagrangian F is a continuous and almost everywhere differentiable function of its arguments. Problem (1.2.1) asks for a function $u_0(x)$ that delivers the minimum of $I(u)$:

$$I(u_0) \leq I(u) \quad \forall u(x).$$

This function is called the *minimizer*.

There are several approaches to the solution of this extremal problem (see, for example (Ioffe and Tihomirov, 1979)). The simplest approach is based on sufficient conditions. One guesses solutions using special algebraic properties of the Lagrangian F; typically, the convexity of F is used.

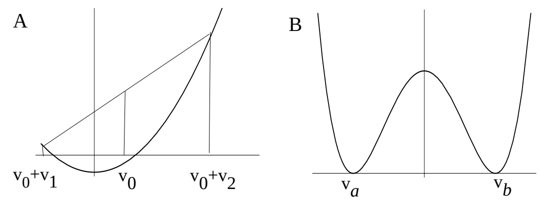

FIGURE 1.2. (A) The definition of convexity; (B) graph of a nonconvex function; $(\mathbf{v}_a, \mathbf{v}_b)$–interval of nonconvexity.

Convexity

Let us briefly discuss convexity. For a detailed exposition of properties of convex functions and convex functionals the reader is referred to (Krasnosel′skiĭ and Rutickiĭ, 1961; Rockafellar, 1997; Ekeland and Temam, 1976; Hardy, Littlewood, and Pólya, 1988).

Here we define convexity through Jensen's inequality. We consider a continuous function $F(\mathbf{v})$ of an n-dimensional vector argument $\mathbf{v} = [v_1, \ldots, v_n]$. Suppose that $\mathbf{v}$ varies in the whole space R^n.

Definition 1.2.1 The function $F(\mathbf{v})$ is convex at the point $\mathbf{v}_0$ if the following inequality (called Jensen's inequality) holds:

$$F(\mathbf{v}_0) \leq \frac{1}{r} \sum_{k=1}^{r} F(\mathbf{v}_0 + \mathbf{v}_k) \quad \forall \mathbf{v}_k : \sum_k \mathbf{v}_k = 0. \tag{1.2.3}$$

The function $F(\mathbf{v})$ is strongly convex at the point $\mathbf{v}_0$ if (1.2.3) becomes a strong inequality.

This inequality expresses the geometrical fact that the graph of a convex function F lies below any secant hyperplane to that graph. The secant is supported by the graph of F at points $\mathbf{v}_0+\mathbf{v}_k$ (see Figure 1.2). For example, the function $F_1(v) = v^2$ of a scalar argument[2] v is convex everywhere, and $F_2(v) = (v^2 - 1)^2$ is convex at the points of the intervals $[-\infty, -1]$ and $[1, \infty]$.

We list here several properties of the convex function that will be used (for a detailed exposition, we refer to the mentioned books):

- A strongly convex function has only one minimum.
- For convexity of F, it is necessary and sufficient that for any point $\mathbf{v}$ there exists an affine function (supporting hyperplane)

$$l(\mathbf{v}) = a_1 v_1 + \ldots + a_n v_n + a_0$$

[2]As a rule, we use Roman letters for scalar and boldface letters for vectors.

such that the graph $F(\mathbf{v}$ lies above the graph of $l(\mathbf{v})$ or coincide with it,

$$F(\mathbf{v}) = l(\mathbf{v}), \quad F(\mathbf{v}') \geq l(\mathbf{v}') \quad \forall \mathbf{v}'.$$

- If F is twice differentiable, then the eigenvalues of its Hessian

$$H = \{H_{ij}\}, \quad H_{ij} = \frac{\partial^2 F}{\partial v_i \partial v_j}, \quad i, j = 1, \ldots, n$$

  are nonnegative: $H \geq 0$. The limiting case when F belongs to the boundary of the domain of the convexity corresponds to the vanishing of an eigenvalue of H or to the condition $\det H = 0$.

Integral Form. We also use an integral form of the definition of convexity. In the limit $r \to \infty$ Jensen's inequality takes the form of an integral inequality. In this case, the vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ are replaced by a vector function $\boldsymbol{\xi}(x) = [\xi_1(x), \ldots, \xi_n(x)]$ of a scalar argument $x \in [0,\ l]$. In this notation, (1.2.3) yields the following inequality.

The function $F(\mathbf{v})$ is convex at point $\mathbf{v}_0$ if and only if the following inequality holds:

$$F(\mathbf{v}_0) \leq \frac{1}{l} \int_0^l F(\mathbf{v}_0 + \boldsymbol{\xi}), \quad \forall \boldsymbol{\xi} : \quad \int_0^l \boldsymbol{\xi} = 0. \tag{1.2.4}$$

Of course, we assume existence of the integrals in (1.2.4).

An equivalent form of Jensen's inequality is obtained by setting $\mathbf{v}(x) = \mathbf{v}_0 + \boldsymbol{\xi}(x)$ and using the identity $l\, F(\mathbf{v}_0) = \int_0^l F(\mathbf{v}_0)$. The inequality is

$$\int_0^l F(\mathbf{v}_0) \leq \int_0^l F(\mathbf{v}(x)) \quad \text{if} \quad \frac{1}{l} \int_0^l \mathbf{v}(x) = \mathbf{v}_0. \tag{1.2.5}$$

This inequality states that the integral of a convex Lagrangian $F(\mathbf{v})$ takes its minimal value if the minimizer $\mathbf{v}$ is constant.

Inequality (1.2.5) introduces a convex functional of $\mathbf{v}$. The properties of convex functionals are discussed in many classical books, such as (Hardy et al., 1988). Particularly, if $f_1(\mathbf{v})$ and $f_2(\mathbf{v})$ are convex functionals and α_1 and α_2 are positive numbers, then $\alpha_1 f_1(\mathbf{v}) + \alpha_2 f_2(\mathbf{v})$ is also a convex functional.

Convexity and the Extremal Problems

For some problems, the convexity of the Lagrangian can be immediately used to find the solution.

Example 1.2.1 Consider the problem of the shortest path between two points in a plane. Suppose that the coordinates of these points are $A =$

$(0,0)$ and $B = (c,d)$ and that the path between them is given by a curve $y = u(x)$. The minimal length of the path is the solution to the problem

$$I = \min_{u(x)} \int_0^c \sqrt{1 + (u'(x))^2}, \quad u(c) - u(0) = \int_0^c u'(x) = d. \tag{1.2.6}$$

The function $f(v) = \sqrt{1+v^2}$ is convex. The integral $\int_0^c v$ is fixed. Therefore (see (1.2.5)) the minimal value of I corresponds to the constant minimizer $v(x)$, $v(x) = \text{constant}(x)$. Applying inequality (1.2.5) to (1.2.6) and using the constraint in (1.2.6), we find that the solution to (1.2.6) is a straight line with slope $u'(x) = \frac{d}{c}$ that passes through the prescribed starting point. We have $u(x) = \frac{d}{c}x$. The cost is $I = \sqrt{c^2 + d^2}$.

More advanced sufficient conditions yield to isoperimetric inequalities (Pólya and Szegö, 1951), symmetrization, Lyapunov functions, etc. If applicable, these conditions immediately detect the true minimizer. However, they are applicable to a very limited number of problems.

Generally, there is no guarantee that sufficient conditions result in strict inequalities that are realizable by a function $u(x)$. If the inequalities are not strict, they can still serve as a lower bound of the cost, but in this case they do not lead to the minimizer.

1.2.2 Variational Methods: Weierstrass Test

More general methods are based on an analysis of infinitesimal variations of a minimizer. We suppose that the function $u_0 = u_0(x)$ is a minimizer and replace u_0 with a varied function $u_0 + \delta u$, assuming that the norm of δu is infinitesimal. The varied function $u_0 + \delta u$ satisfies the same boundary conditions as u_0. If indeed u_0 is a minimizer, the increment of the cost $\delta I(u_0) = I(u_0 + \delta u) - I(u_0)$ is nonnegative:

$$\delta I(u_0) = \int_0^1 (F(x, u_0 + \delta u, (u_0 + \delta u)') - F(x, u_0, u_0')) \geq 0. \tag{1.2.7}$$

To effectively compute $\delta I(u_0)$, we also assume the smallness of δu. Variational methods yield to the necessary conditions of optimality because it is assumed that the compared trajectories are close to each other. On the other hand, variational methods are applicable to a wide variety of extremal problems of the type (1.2.1), called *variational problems*. Necessary conditions are the true workhorses of extremal problem theory, while exact sufficient conditions are rare and remarkable exceptions.

There are many books that expound the calculus of variations, including (Bliss, 1946; Courant and Hilbert, 1962; Gelfand and Fomin, 1963; Lavrent′ev, 1989; Weinstock, 1974; Mikhlin, 1964; Leitmann, 1981; Fox, 1987; Dacorogna, 1989).

Euler–Lagrange Equations

The calculus of variations suggests a set of tests that must be satisfied by a minimizer. These conditions express realization of (1.2.7) by various variations δu. To perform the test one must specify the type of perturbations δu. The simplest variational condition (the Euler–Lagrange equation) is derived by linearizing the inequality (1.2.7) with respect to an infinitesimal small and localized variation

$$\delta u = \begin{cases} \rho(x) & \text{if } x \in [x_0, x_0 + \varepsilon], \\ 0 & \text{otherwise.} \end{cases} \tag{1.2.8}$$

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

$$|\rho(x)| < \varepsilon, \quad |\rho'(x)| < \varepsilon \quad \forall x \in [x_0, x_0 + \varepsilon].$$

Linearizing with respect to ε and collecting main terms, we rewrite (1.2.7) as

$$\delta I(u_0) = \varepsilon \left(\int_0^1 \left(\frac{\partial F}{\partial u}(\delta u) + \frac{\partial F}{\partial u'}(\delta u)' \right) \right) + o(\varepsilon) \geq 0. \tag{1.2.9}$$

Integration by parts of the last term on the right-hand side of (1.2.9) gives

$$\delta I(u_0) = \varepsilon \int_0^1 S(u, u', x)\delta u + \frac{\partial F}{\partial u'}\delta u \bigg|_{x=0}^{x=1} + o(\varepsilon) \geq 0, \tag{1.2.10}$$

where

$$S(u, u', x) = -\frac{d}{dx}\frac{\partial F}{\partial u'} + \frac{\partial F}{\partial u}. \tag{1.2.11}$$

The second term on the right-hand side of (1.2.10) is zero, because the boundary values of u are prescribed

$$u(0) = u_0, \quad u(1) = u_1$$

and their variations $\delta u|_{x=0}$ and $\delta u|_{x=1}$ are zero.

Due to the arbitrariness of δu we conclude that a minimizer u_0 must satisfy the differential equation

$$S(u, u', x) = 0, \quad u(0) = u_0, \quad u(1) = u_1, \tag{1.2.12}$$

called the *Euler–Lagrange equation* and the corresponding boundary conditions. The Euler–Lagrange equation is also called the *stationary condition.* Indirectly, we assume in this derivation that u_0 is a twice differentiable function of x. We do not discuss here the properties of the Euler–Lagrange equations for different types of Lagrangians; we refer readers to mentioned books on the calculus of variations.

It is important to mention that the stationarity test alone does not allow us to conclude whether u_0 is a true minimizer or even to conclude that a solution to (1.2.12) exists. For example, the function u that *maximizes* $I(u)$ satisfies the same Euler–Lagrange equation.

The Weierstrass Test

In addition to being a solution to the Euler equation, the true minimizer satisfies necessary conditions in the form of inequalities. These conditions distinguish the trajectories that correspond to the minimum of the functional from trajectories that correspond either to its maximum or to a saddle point stationary solution. One of these conditions is the Weierstrass test; it detects stability of a solution to a variational problem against strong local perturbations.

Suppose that u_0 is the minimizer of variational problem (1.2.1) that satisfies the Euler equation (1.2.11). Additionally, u_0 should satisfy another test that uses a type of variation δu different from (1.2.8). The variation used in the Weierstrass test is an infinitesimal triangle supported on the interval $[x_0, x_0+\varepsilon]$ in a neighborhood of a point $x_0 \in (0,\ 1)$ (see Figure 1.3):

$$\Delta u(x) = \begin{cases} 0 & \text{if } \ x \notin [x_0, x_0+\varepsilon], \\ v_1(x - x_0) & \text{if } \ x \in [x_0, x_0+\alpha\varepsilon], \\ v_1\alpha\varepsilon + v_2(x - x_0 - \alpha\varepsilon) & \text{if } \ x \in [x_0+\alpha\varepsilon, x_0+\varepsilon], \end{cases}$$

where the parameters $\alpha,\ v_1,\ v_2$ are related by

$$\alpha v_1 + (1-\alpha)v_2 = 0.$$

This relation provides the continuity of $u_0+\Delta u$ at the point $x_0+\varepsilon$, because it yields to the equality $\Delta u(x_0+\varepsilon-0) = 0$.

Note that this variation (the Weierstrass variation) is localized and has an infinitesimal absolute value (if $\varepsilon \to 0$), but its derivative $(\Delta u)'$ is finite, unlike the variation in (1.2.8) (see Figure 1.3):

$$(\Delta u)' = \begin{cases} 0 & \text{if } \ x \notin [x_0, x_0+\varepsilon], \\ v_1 & \text{if } \ x \in [x_0, x_0+\alpha\varepsilon], \\ v_2 & \text{if } \ x \in [x_0+\alpha\varepsilon, x_0+\varepsilon]. \end{cases}$$

Computing δI from (1.2.7) and rounding up to ε, we find that

$$\begin{aligned} \delta I = \varepsilon[&\alpha F(x_0, u_0, u_0'+v_1) + (1-\alpha)F(x_0, u_0, u_0'+v_2) \\ &-F(x_0, u_0, u_0')] + o(\varepsilon) \geq 0 \end{aligned}$$

if u_0 is a minimizer.

The last expression yields to the Weierstrass test and the necessary Weierstrass condition. Any minimizer $u(x)$ of (1.2.1) satisfies the inequality

$$\alpha F(x_0, u_0, u_0'+v_1) + (1-\alpha)F(x_0, u_0, u_0'+v_2) - F(x_0, u_0, u_0') \geq 0.$$

Comparing this with the definition of convexity (1.2.2), we observe that the Weierstrass condition requires convexity of the Lagrangian $F(x, y, z)$ with respect to its third argument $z = u'$. The first two arguments $x,\ y = u$ here are the coordinates $x,\ u(x)$ of the testing minimizer $u(x)$. Recall that minimizer $u(x)$ is a solution to the Euler equation.

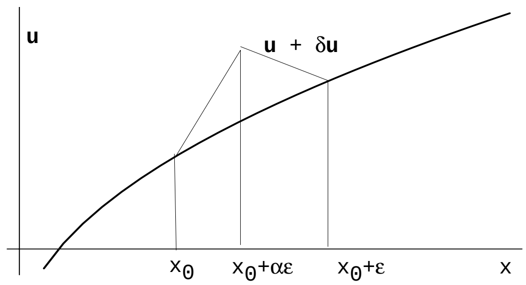

FIGURE 1.3. Weierstrass variation.

Vector-Valued Minimizer. The Euler equation and the Weierstrass condition can be naturally generalized to the problem with the vector-valued minimizer

$$I(u) = \min_{\mathbf{u}} \int_0^1 F(x, \mathbf{u}, \mathbf{u}'),$$

where x is a point in the interval $[0, 1]$ and $\mathbf{u} = [u_1(x), \ldots, u_n(x)]$ is a vector function. We suppose that F is a twice differentiable function of its arguments. The classical (twice differentiable) local minimizer $\mathbf{u}_0$ of the problem (1.2.1) is given by a solution to the vector-valued Euler equations,

$$\frac{d}{dx}\frac{\partial F}{\partial \mathbf{u}_0'} - \frac{\partial F}{\partial \mathbf{u}_0} = 0,$$

which expresses the stationarity requirement of a minimizer to small variations of the variable $\mathbf{u}$.

The Weierstrass test requires convexity of $F(x, \mathbf{y}, \mathbf{z})$ with respect to the last vector argument. Here again $\mathbf{y} = \mathbf{u}_0(x)$ represents a minimizer.

Remark 1.2.1 *Convexity of the Lagrangian does not guarantee the existence of a solution to a variational problem. It states only that the minimizer (if it exists) is stable against fine-scale perturbations. However, the minimum may not exist at all, see, for example (Ioffe and Tihomirov, 1979; Zhikov, 1993).*

If the solution of a variational problem fails the Weierstrass test, then its cost can be decreased by adding infinitesimal wiggles to the solution. The wiggles are the Weierstrass trial functions, which decrease the cost. In this case, we call the variational problem ill-posed, and we say that the solution is unstable against fine-scale perturbations.

1.3 Relaxation

1.3.1 Nonconvex Variational Problems

Typical problems of structural optimization correspond to a Lagrangian $F(x, \mathbf{y}, \mathbf{z})$ that is nonconvex with respect to $\mathbf{z}$. In this case, the Weierstrass test fails, and the problem is ill-posed.

Let us consider a problem of this type. Suppose that the Lagrangian $F(x, \mathbf{y}, \mathbf{z})$ is a nonconvex function of its third argument; is bounded from below (say, by zero),

$$F(x, \mathbf{y}, \mathbf{z}) \geq 0 \quad \forall x, \mathbf{y}, \mathbf{z};$$

and satisfies the condition

$$\lim_{|z| \to \infty} \frac{F(x, \mathbf{y}, \mathbf{z})}{|z|} = \infty.$$

Then the infimum I_0

$$I_0 = \inf_{\mathbf{u}} I(\mathbf{u}), \quad I(\mathbf{u}) = \int_0^1 F(x, \mathbf{u}, \mathbf{u}')$$

is nonnegative, $I_0 \geq 0$.

One can construct a minimizing sequence $\{\mathbf{u}^s\}$ such that $I(\mathbf{u}^s) \to I_0$. Due to the preceding condition, the minimizing sequence $\{\mathbf{u}^s\}$ consists of continuous functions with L_1-bounded derivatives; see (Dacorogna, 1989).

Because $F(.,.,\mathbf{z})$ is not convex, this minimizing sequence cannot tend to a differentiable curve in the limit. Otherwise it would satisfy the Euler equation and the Weierstrass test, but the last requires the convexity of $F(.,.,\mathbf{z})$.

We will demonstrate that a minimizing sequence tends to a "generalized curve." It consists of infinitesimal zigzags. The limiting curve has a dense set of points of discontinuity of the derivative. A detailed explanation of this phenomenon can be found, for example, in (Young, 1942a; Young, 1942b; Gamkrelidze, 1962; Young, 1969; Warga, 1972; Gamkrelidze, 1985). Here we give a brief description of it, mainly by working on several examples.

Example 1.3.1 Consider a simple variational problem that yields to the generalized solution (Young, 1969):

$$\inf_u \ I(u) = \inf_u \int_0^1 G(u, u'), \tag{1.3.1}$$

where

$$G(u, v) = u^2 + \min\{(v-1)^2, \ (v+1)^2\}, \quad u(0) = u(1) = 0. \tag{1.3.2}$$

The graph of the function $G(., v)$ is presented in Figure 1.2B.

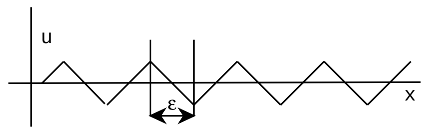

FIGURE 1.4. Oscillating minimizing sequence.

The Lagrangian G penalizes the trajectory u for having the speed $|u'|$ different from ± 1 and penalizes the deflection of the trajectory u from zero. These contradictory requirements cannot be resolved in the class of classical trajectories.

Indeed, a differentiable minimizer satisfies the Euler equation (1.2.12) that takes the form

$$u'' - u = 0 \quad \text{if} \;\; u' \neq 0. \tag{1.3.3}$$

Next, the Weierstrass test additionally requires convexity of $G(u, v)$ with respect to v; the Lagrangian $G(u, v)$ is nonconvex in the interval $v \in (-1,\ 1)$ (see Figure 1.2). The Weierstrass test requires that the extremal (1.3.3) is supplemented by the inequality (recall that $v = u'$)

$$u' \notin (-1,\ 1) \quad \text{at the optimal trajectory.} \tag{1.3.4}$$

and it is not clear how to satisfy it

On the other hand, the minimizing sequence for problem (1.3.1) can be immediately constructed. Indeed, the infimum of (1.3.1) obviously is nonnegative, $\inf_u I(u) \geq 0$. Therefore, a sequence u^s with the property

$$\lim_{s \to \infty} I(u^s) = 0$$

is a minimizing sequence.

Consider a set of functions $\tilde{u}^s(x)$ that belong to the boundary of the *forbidden interval* $\tilde{u}'(x) = -1$ or $\tilde{u}'(x) = 1$ of the nonconvexity of $G(., v)$. These functions make the second term in the Lagrangian (1.3.2) vanish, $\min\{(\tilde{u}' - 1)^2,\ (\tilde{u}' + 1)^2\} = 0$, and the problem becomes

$$I(\tilde{u}^s, (\tilde{u}^s)') = \int_0^1 (\tilde{u}^s)^2.$$

The term $\tilde{u}^s$ oscillates near zero if the derivative $(\tilde{u}^s)'$ changes its sign on intervals of equal length. The cost $I(\tilde{u}^s)$ depends on the density of switching points and tends to zero when the number of these points increases (see Figure 1.4). Therefore, the minimizing sequence consists of the saw-tooth functions $\tilde{u}^s$; the heights of the teeth tend to zero and their number tends to infinity as $s \to \infty$.

Note that the minimizing sequence $\{\tilde{u}^s\}$ does not converge to any classical function but rather to a distribution. This minimizer $\tilde{u}^s(x)$ satisfies the contradictory requirements, namely, the derivative must keep the absolute value equal to one, but the function itself must be arbitrarily close to zero:

$$|(\tilde{u}^s)'| = 1 \quad \forall x \in [0,1], \qquad \max_{x\in[0,1]} \tilde{u}^s \to 0 \quad \text{as } s \to \infty.$$

The limiting curve u_0 has zero norm in $C_0[0,\ 1]$ but a finite norm in $C_1[0,\ 1]$.

Remark 1.3.1 *If boundary values were different, the solution could correspond partly to the classical extremal (1.3.3), (1.3.4), and partly to the saw-tooth curve; in the last case u' belongs on the boundary of the forbidden interval $|u'| = 1$.*

This considered nonconvex problem is an example of an ill-posed variational problem. For these problems, the classical variational technique based on the Euler equation fails to work. Other methods are needed to deal with such problems. Namely, we replace an ill-posed problem with a *relaxed* one.

1.3.2 Convex Envelope

Consider a variational problem with a nonconvex Lagrangian F. We want to replace this problem with a new one that describes infinitely rapidly oscillating minimizers in terms of averages. This will be done by the construction of the *convex envelope* of a nonconvex Lagrangian. Let us give the definitions (see (Rockafellar, 1997)).

Definition 1.3.1 The *convex envelope* $\mathcal{C}F$ is a solution to the following minimal problem:

$$\mathcal{C}F(\mathbf{v}) = \inf_{\boldsymbol{\xi}} \frac{1}{l}\int_0^l F(\mathbf{v}+\boldsymbol{\xi}) \quad \forall\, \boldsymbol{\xi} : \int_0^l \boldsymbol{\xi} = 0. \tag{1.3.5}$$

This definition determines the convex envelope as the minimum of all parallel secant hyperplanes that intersect the graph of F; it is based on Jensen's inequality (1.2.4).

To compute the convex envelope $\mathcal{C}F$ one can use the Carathéodory theorem (see (Carathéodory, 1967; Rockafellar, 1997)). It states that the argument $\boldsymbol{\xi}(x) = [\xi_1(x), \ldots, \xi_n(x)]$ that minimizes the right-hand side of (1.3.5) takes no more than $n+1$ different values. This theorem refers to the obvious geometrical fact that the convex envelope consists of the supporting hyperplanes to the graph $F(\xi_1, \ldots, \xi_n)$. Each of these hyperplanes is supported by no more than $(n+1)$ arbitrary points.

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

(1.3.5) becomes:

$$CF(\mathbf{v}) = \min_{m_i \in M} \min_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}} \left(\sum_{i=1}^{n+1} m_i F(\mathbf{v} + \boldsymbol{\xi}_i) \right), \tag{1.3.6}$$

where

$$M = \left\{ m_i : \quad m_i \geq 0, \quad \sum_{i=1}^{n+1} m_i = 1 \right\} \tag{1.3.7}$$

and

$$\boldsymbol{\Xi} = \left\{ \boldsymbol{\xi}_i : \quad \sum_{i=1}^{n+1} m_i \boldsymbol{\xi}_i = 0 \right\}. \tag{1.3.8}$$

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

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

On the other hand, the convex envelope is the greatest convex function that does not exceed $F(\mathbf{v})$ in any point $\mathbf{v}$ (Rockafellar, 1997):

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

Example 1.3.2 Obviously, the convex envelope of a convex function coincides with the function itself, so all m_i but m_1 are zero in (1.3.6) and $m_1 = 1$; the parameter $\boldsymbol{\xi}_1$ is zero because of the restriction (1.3.8).

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

$$\Phi(\mathbf{v}) = \min \{F_1(\mathbf{v}), \ F_2(\mathbf{v})\},$$

where $F_1, \ F_2$ are convex functions of $\mathbf{v}$, either coincides with one of the functions $F_1, \ F_2$ or is supported by no more than two points for every $\mathbf{v}$; supporting points belong to different wells. In this case, formulas (1.3.6)–(1.3.8) for the convex envelope are reduced to

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

Indeed, the convex envelope touches the graphs of the convex functions F_1 and F_2 in no more than one point. Call the coordinates of the touching points $\mathbf{v} + \boldsymbol{\xi}_1$ and $\mathbf{v} + \boldsymbol{\xi}_2$, respectively. The restrictions (1.3.8) become $m_1 \boldsymbol{\xi}_1 + m_2 \boldsymbol{\xi}_2 = 0, \ m_1 + m_2 = 1$. It implies the representations $\boldsymbol{\xi}_1 = -(1-m)\boldsymbol{\xi}$ and $\boldsymbol{\xi}_2 = m\boldsymbol{\xi}$.

Example 1.3.3 Consider the special case of the two-well function,

$$F(v_1, v_2) = \begin{cases} 0 & \text{if } v_1^2 + v_2^2 = 0, \\ 1 + v_1^2 + v_2^2 & \text{if } v_1^2 + v_2^2 \neq 0. \end{cases} \tag{1.3.9}$$

The convex envelope of F is equal to

$$\mathcal{C}F(v_1, v_2) = \begin{cases} 2\sqrt{v_1^2 + v_2^2} & \text{if } v_1^2 + v_2^2 \leq 1, \\ 1 + v_1^2 + v_2^2 & \text{if } v_1^2 + v_2^2 > 1. \end{cases} \tag{1.3.10}$$

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

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

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

and

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

The calculation of the minimum gives (1.3.10).

Example 1.3.4 Consider the nonconvex function $F(v)$ used in Example 1.3.1:

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

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

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

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

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

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

1.3.3 Minimal Extension and Minimizing Sequences

The construction of the convex envelope is used to reformulate (relax) a nonconvex variational problem. Consider again the variational problem

$$I(\mathbf{u}) = \min_{\mathbf{u}} \int_0^1 F(x, \mathbf{u}, \mathbf{u}') \tag{1.3.11}$$

where F is a continuous function that is not convex with respect to its last argument. Recall that this problem does not satisfy the Weierstrass test on the intervals of nonconvexity of F.

Definition 1.3.2 We call the *forbidden region* $\mathcal{Z}_\mathrm{f}$ the set of $\mathbf{z}$ for which $F(x, \mathbf{y}, \mathbf{z})$ is not convex with respect to $\mathbf{z}$,

$$Z_\mathrm{f} = \{z : \ \mathcal{C}_z F(x, \mathbf{y}, \mathbf{z}) < F(x, \mathbf{y}, \mathbf{z})\} .$$

The notation $\mathcal{C}_z F(x, \mathbf{y}, \mathbf{z})$ is used to show the argument $\mathbf{z}$ of the convexification: The other two arguments are considered to be parameters when the convex envelope is calculated. (Later, we omit the subindex $(\ )_z$ when this does not lead to misunderstanding.)

Note that the derivative $\mathbf{u}'$ of a minimizer $\mathbf{u}$ of (1.3.11) never belongs to the region Z_f:

$$\mathbf{u}' \notin Z_\mathrm{f}.$$

This additional constraint on the minimizer is satisfied in the construction of a minimizer of a nonconvex problem.

To deal with a nonconvex problem, we "relax" it. *Relaxation* means that we replace the problem with another one that has the same cost but whose solution is stable against fine-scale perturbations; particularly, it cannot be improved by the Weierstrass variation. The relaxed problem has the following two basic properties:

- The relaxed problem has a classical solution.
- The infimum of the functional (the cost of the problem) in the initial problem coincides with the cost of the relaxed problem.

Here we will demonstrate two approaches to relaxation. Each of them yields to the same construction but uses different arguments to achieve it. In the next chapters we will see similar procedures applied to variational problems with multiple integrals; sometimes they also yield the same construction, but generally they result in different relaxations.

Minimizing Sequences

The first construction is based on local minimization. Consider the extremal problem (1.3.11) and the corresponding solution $\mathbf{u}_0(x)$. Let us fix two neighboring points $A = (x_0,\ \mathbf{u}_0(x_0))$ and $B = ((x_0 + \varepsilon),\ \mathbf{u}_0(x_0 + \varepsilon))$ on this solution. Using the smallness of ε, we represent B as

$$B = ((x_0 + \varepsilon), \mathbf{u}_0(x_0) + \varepsilon \mathbf{u}'(x_0) + o(\varepsilon))\,.$$

The impact to the cost of problem (1.2.1) due to this interval is

$$I_\varepsilon(\mathbf{u}_0) = \int_{x_0}^{x_0+\varepsilon} F(x, \mathbf{u}_0, \mathbf{u}_0') = \varepsilon F(x_0, \mathbf{u}_0(x_0), \mathbf{u}_0'(x_0)) + o(\varepsilon).$$

Let us examine a local variation of the solution $\mathbf{u}_0(x)$: We replace it with a zigzag piecewise linear curve that passes through the points A and B.

Consider a continuous curve $\mathbf{u}_\varepsilon$ that contains $p-1$ subintervals of the constancy of the derivative $\mathbf{v} = \mathbf{u}_\varepsilon'$. The variable $\mathbf{v}(x)$ takes several values $\mathbf{v}_1 + \mathbf{u}_0'(x_0), \ldots, \mathbf{v}_p + \mathbf{u}_0'(x_0)$; each value is kept on the subinterval of length εm_i, where $\sum m_i = 1$. The derivative $\mathbf{u}_\varepsilon'(x)$ is

$$\mathbf{u}_\varepsilon'(x) = \mathbf{u}_0'(x_0) + \mathbf{v}_k \quad \text{if} \quad x \in \left[x_0 + \varepsilon \sum_{i=1}^{k} m_i,\ x_0 + \varepsilon \sum_{i=1}^{k+1} m_i. \right],$$

where $k = 1, \ldots, p-1$. The saw-tooth curve $\mathbf{u}_\varepsilon$ is

$$\mathbf{u}_\varepsilon(x) = \mathbf{u}(x_0) + \int_{x_0}^{x} (\mathbf{u}'(x_0) + \mathbf{v}(x)) dx. \tag{1.3.12}$$

We require that any admissible solution $\mathbf{u}_\varepsilon$ passes through point B. More exactly, we require that its value at the point $x_0 + \varepsilon$ is equal to $\mathbf{u}_0(x_0 + \varepsilon)$ up to the terms of the order of $o(\varepsilon)$,

$$\mathbf{u}_\varepsilon(x_0 + \varepsilon) - \mathbf{u}_0(x_0 + \varepsilon) = \sum_{i=1}^{p} m_i \mathbf{v}_i = o(\varepsilon). \tag{1.3.13}$$

Let us compute the effect of replacing the differentiable curve $\mathbf{u}_0$ with the zigzag curve $\mathbf{u}_\varepsilon$. We estimate the integral of $F(x, \mathbf{u}_\varepsilon, \mathbf{u}_\varepsilon')$ over this interval, up to terms of order of $o(\varepsilon)$. To estimate, we use the smallness of the interval of variation. Replace $\mathbf{u}_\varepsilon(x)$ with $\mathbf{u}_0(x_0)$

$$\mathbf{u}_\varepsilon(x) = \mathbf{u}_0(x_0) + O(\varepsilon)$$

and compute

$$F(x, \mathbf{u}_\varepsilon(x), \mathbf{v}_i + \mathbf{u}_0'(x_0)) = F(x_0, \mathbf{u}_0(x_0), \mathbf{u}_0'(x_0) + \mathbf{u}_\varepsilon'(x)) + O(\varepsilon)$$

for any $x \in [x_0,\ x_0 + \varepsilon]$. The Lagrangian (rounded up to $O(\varepsilon)$) is piecewise constant in the interval $[x_0,\ x_0 + \varepsilon]$. The impact $I_\varepsilon(\mathbf{u}_\varepsilon)$ becomes

$$I_\varepsilon(\mathbf{u}_\varepsilon) = \varepsilon \sum_{i=1}^{p} m_i F(x_0, \mathbf{u}_\varepsilon(x_0), \mathbf{u}_0'(x_0) + \mathbf{v}_i) + o(\varepsilon). \tag{1.3.14}$$

Calculate the minimum of (1.3.14) with respect to the arguments $\mathbf{v}_1, \dots, \mathbf{v}_p$ and $m_1, \dots, m_p$, which are subject to the constraints (see (1.3.13))

$$m_i(x) \geq 0, \quad \sum_{i=1}^{p} m_i = 1, \quad \sum_{i=1}^{p} m_i \mathbf{v}_i = 0. \tag{1.3.15}$$

This minimum coincides with the convex envelope of the original Lagrangian with respect to its last argument (see (1.3.7)):

$$\min_{m_i, \mathbf{v}_i \in (1.3.15)} \sum_{i=1}^{p} m_i F(x, \mathbf{u}, \mathbf{v}_i) = \mathcal{C}F_{\mathbf{v}}(x, \mathbf{u}_0, \mathbf{v}). \tag{1.3.16}$$

By referring to the Carathéodory theorem (1.3.7) we conclude that it is enough to split the interval into

$$p = n + 1 \tag{1.3.17}$$

parts so that $\mathbf{v} = \mathbf{u}'$ takes $k+1$ values.

Note that the constraint (1.3.15) leaves the freedom to choose inner parameters m_i and $\mathbf{v}_i$ to minimize the function $\sum_{i=1}^{p} m_i F(u, \mathbf{v}_i)$ and thus to minimize the cost of the variational problem (see (1.3.16)).

Compare the costs $I_\varepsilon(\mathbf{u}_0)$ and $I_\varepsilon(\mathbf{u}_\varepsilon)$ of (1.3.11) corresponding to the smooth solution $\mathbf{u}_0$ and to the zigzag solution $(\mathbf{u}_\varepsilon)$. Using the definition of the convex envelope we obtain the inequality:

$$\begin{aligned} \frac{1}{\varepsilon}\left(I_\varepsilon(\mathbf{u}_0) - I_\varepsilon(\mathbf{u}_\varepsilon)\right) \;=\; & F(x_0, \mathbf{u}_0(x_0), \mathbf{u}_0'(x_0) \\ & -\mathcal{C}F(x_0, \mathbf{u}_\varepsilon(x_0), \mathbf{u}'(x_0)) + O(\varepsilon) \geq 0. \end{aligned}$$

We see that the zigzag solution $\mathbf{u}_\varepsilon$ corresponds to lower cost if $F > \mathcal{C}_z F$, that is, in the regions of nonconvexity of F.

Passing to the variational problem (1.3.11) in the whole interval $[0,\ 1]$ we perform the preceding extension in each interval of length ε. This extension replaces the Lagrangian $F(x, y, z)$ with the convex envelope $\mathcal{C}_z F(x, y, z)$ so that the relaxed problem becomes

$$I = \min_{\mathbf{u}} \int_0^1 \mathcal{C}_{\mathbf{u}'} F\left(x, \mathbf{u}(x), \mathbf{u}'(x)\right). \tag{1.3.18}$$

The curve $\mathbf{u}_\varepsilon$ strongly converges to the curve $\mathbf{u}_0$:

$$\|\mathbf{u}_\varepsilon - \mathbf{u}_0\|_{L_\infty[0,1]} \to 0, \quad \text{as } \varepsilon \to 0,$$

but its derivative converges to $\mathbf{u}_0'$ only weakly in L_p,

$$\int_0^1 \phi\left(\mathbf{u}_\varepsilon' - \mathbf{u}_0'\right) \to 0 \quad \forall\ \phi \in L_q(0,1), \quad \frac{1}{p} + \frac{1}{q} = 1.$$

For the definition and discussion of the weak convergence we refer the reader to books on analysis, such as (Shilov, 1996).

Remark 1.3.2 *The choice of the proper space L_p depends on the Lagrangian F because $F(.,.,\mathbf{u}')$ must be integrable.*

The cost of the reformulated (relaxed) problem (1.3.18) corresponds to the cost of the problem (1.3.11) on the minimizing sequence (1.3.12). Therefore, the cost of the relaxed problem is equal to the cost of the original problem (1.3.11). The extension of the Lagrangian that preserves the cost of the problem is called the *minimal extension*. The minimal extension enlarges the set of classical minimizers by including generalized curves in it.

Generally speaking, this extension leads to an attainable upper bound of the cost of an unstable problem because we cannot guarantee that the extension cannot be further improved. However, the Lagrangian of the relaxed problem is convex, which guarantees that its minimizers satisfy the Weierstrass test and is stable against fine-scale perturbations.

Minimal Extension, Based on the Weierstrass Test

We introduce an alternative method of relaxation that leads to the same results but does not require consideration of the structure of minimizing sequences.

Consider the class of Lagrangians $\mathcal{N}F(x,y,z)$ that are smaller than $F(x,y,z)$ and satisfy the Weierstrass test $\mathcal{W}(\mathcal{N}F(x,y,z)) \geq 0$:

$$\begin{cases} \mathcal{N}F(x,y,z) - F(x,y,z) \leq 0, \\ \mathcal{W}(\mathcal{N}F(x,y,z)) \geq 0 \end{cases} \quad \forall\ x,y,z.$$

Let us take the maximum on $\mathcal{N}F(x,y,z)$ and call it $\mathcal{S}F$. Clearly, $\mathcal{S}F$ corresponds to turning one of these inequalities into an equality:

$$\begin{array}{llll} \mathcal{S}F(x,y,z) = F(x,y,z), & \mathcal{W}(\mathcal{S}F(x,y,z)) \geq 0 & \text{if} & z \notin \mathcal{Z}_{\text{f}}, \\ \mathcal{S}F(x,y,z) \leq F(x,y,z), & \mathcal{W}(\mathcal{S}F(x,y,z)) = 0 & \text{if} & z \in \mathcal{Z}_{\text{f}}. \end{array}$$

This variational inequality describes the extension of the Lagrangian of an unstable variational problem. Notice that

1. The first equality holds in the region of convexity of F and the extension coincides with F in that region.

2. In the region where F is not convex, the Weierstrass test of the extended Lagrangian is satisfied as an equality; this equality serves to determine the extension.

These conditions imply that $\mathcal{S}F$ is convex everywhere. Also, $\mathcal{S}F$ is the maximum over all convex functions that do not exceed F. Again, $\mathcal{S}F$ is equal to the convex envelope of F:

$$\mathcal{S}F(x,y,z) = \mathcal{C}_z F(x,y,z).$$

The cost of the problem remains the same because the convex envelope corresponds to a minimizing sequence of the original problem.

Remark 1.3.3 *Note that the geometrical property of convexity never explicitly appears here. We simply satisfy the Weierstrass necessary condition everywhere. Hence, this relaxation procedure can be extended to more complicated multidimensional problems for which the Weierstrass condition and convexity do not coincide.*

Properties of the Relaxed Problem

- Recall that the derivative of the minimizer never takes values in the region $\mathcal{Z}_{\mathrm{f}}$ of nonconvexity of F. Therefore, a solution to a nonconvex problem stays the same if its Lagrangian $F(x, \mathbf{y}, \mathbf{z})$ is replaced by any Lagrangian $\mathcal{N}F(x, \mathbf{y}, \mathbf{z})$ that satisfies the restrictions

$$\begin{aligned} \mathcal{N}F(x, \mathbf{y}, \mathbf{z}) &= F(x, \mathbf{y}, \mathbf{z}) \quad \forall\, z \notin \mathcal{Z}_{\mathrm{f}}, \\ \mathcal{N}F(x, \mathbf{y}, \mathbf{z}) &> \mathcal{C}F(x, \mathbf{y}, \mathbf{z}) \;\forall\, z \in \mathcal{Z}_{\mathrm{f}}. \end{aligned}$$

  Indeed, the two Lagrangians $F(x, \mathbf{y}, \mathbf{z})$ and $\mathcal{N}F(x, \mathbf{y}, \mathbf{z})$ coincide in the region of convexity of F. Therefore, the solutions to the variational problem also coincide in this region. Neither Lagrangian satisfies the Weierstrass test in the forbidden region of nonconvexity. Therefore, no minimizer can distinguish between these two problems: It never takes values in Z_{f}. The behavior of the Lagrangian in the forbidden region is simply of no importance. In this interval, the Lagrangian cannot be computed from the minimizer.

- The infimum of the functional for the initial problem coincides with the minimum of the functional in the relaxed problem. The relaxed problem has a convex Lagrangian. The Weierstrass test is satisfied, and the minimal solution (if it exists) is stable against fine-scale perturbations. To be sure that the solution of the relaxed problem exists, one should also examine other sources of possible nonexistence (see, for example (Ioffe and Tihomirov, 1979)).

- The number of minimizers in the relaxed problem is increased. Instead of one n-dimensional vector minimizer $\mathbf{u}(x)$ in the original problem, they now include $n+1$ vector minimizers $\mathbf{v}_i(x)$ and $n+1$ minimizers $m_i(x)$ (see (1.3.17)) connected by two equalities (1.3.15) and the inequalities $m_i(x) \geq 0$. The relaxed problem is controlled by the larger number of independent parameters that are used to compute the relaxed Lagrangian $\mathcal{C}F(x, \mathbf{u}, \mathbf{u}')$.

In the forbidden region, the Euler equations degenerate. For example, suppose that u is a scalar; the convex envelope has the form

$$\mathcal{C}F = au' + b(x, u)$$

if it does not coincide with G. This representation implies that the first term in the left-hand side (1.2.11) of the Euler equation (1.2.12) vanishes

Average derivative	Pointwise derivatives	Optimal concentrations	Convex envelope $\mathcal{C}G(u,v)$
$v < -1$	$v_1^0 = v_2^0 = v$	$m_1^0 = 1,\ m_2^0 = 0$	$u^2 + (v-1)^2$
$\|v\| < 1$	$v_1^0 = 1, v_2^0 = -1$	$m_1^0 = m_2^0 = \frac{1}{2}$	u^2
$v > 1$	$v_1^0 = v_2^0 = v$	$m_1^0 = 0,\ m_2^0 = 1$	$u^2 + (v+1)^2$

TABLE 1.1. Characteristics of an optimal solution in Example 1.3.1.

identically: $\frac{d}{dx}\frac{\partial}{\partial u'}\mathcal{C}F \equiv 0$. The Euler equation degenerates into an algebraic equation $\frac{\partial}{\partial u}\mathcal{C}F = 0$.

In the general case, the order of the system of Euler equations decreases (for details, see (Gamkrelidze, 1962; Gabasov and Kirillova, 1973; Clements and Anderson, 1978)).

1.3.4 Examples: Solutions to Nonconvex Problems

Example 1.3.5 We revisit Example 1.3.1. Let us solve this problem by building the convex envelope of the Lagrangian $G(u,v)$:

$$\mathcal{C}_v G(u,v) = \min_{m_1,m_2} \min_{v_1,v_2} \left\{ u^2 + m_1(v_1-1)^2 + m_2(v_2+1)^2 \right\},$$
$$v = m_1 v_1 + m_2 v_2, \quad m_1 + m_2 = 1, \quad m_i \geq 0.$$

The form of the minimum depends on the value of $v = u'$. The convex envelope $\mathcal{C}G(u,v)$ coincides with either $G(u,v)$ if $v \notin [0,\ 1]$ or $\mathcal{C}G(u,v) = u^2$ if $v \in [0,\ 1]$; see Example 1.3.4. Optimal values $v_1^0,\ v_2^0,\ m_1^0\ m_2^0$ of the minimizers and the convex envelope $\mathcal{C}G$ are shown in Table 1.1. The relaxed form of the problem with zero boundary conditions

$$\min_u \int_0^1 \mathcal{C}G(u,\ u'), \quad u(0) = u(1) = 0,$$

has an obvious solution,

$$u(x) = u'(x) = 0,$$

that yields the minimal (zero) value of the functional. It corresponds to the constant optimal value m_{opt} of $m(x)$: $m_{\text{opt}}(x) = \frac{1}{2}\ \forall x \in [0,1]$.

The relaxed Lagrangian is minimized over four functions $u, m_1, v_1,\ v_2$ bounded by one equality, $u' = m_1 v_1 + (1-m_1)v_2$ and the inequalities $0 \leq m \leq 1$, while the original Lagrangian is minimized over one function u. In contrast to the initial problem, the relaxed one has a differentiable solution in terms of these four controls.

A Two-Well Lagrangian

We turn to a more advanced example of the relaxation of an ill-posed nonconvex variational problem. This example highlights more properties

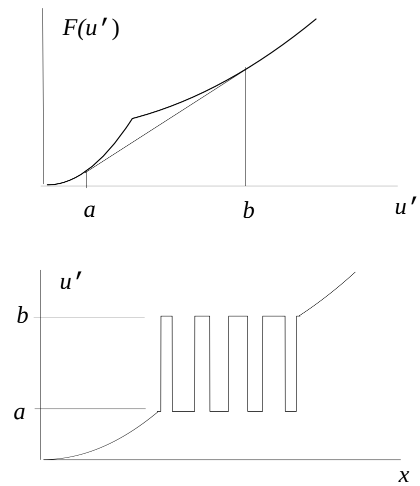

FIGURE 1.5. Convexification of the Lagrangian (top) and the minimizer (bottom); points a and b are equal to v_1 and v_2, respectively.

of relaxation and introduces piecewise quadratic Lagrangians that are the main tool in the investigation of optimal composites.

Example 1.3.6 Consider the minimization problem

$$\min_{u(x)} \int_0^z F_p(x, u, u'), \quad u(0) = 0, \; u'(z) = 0 \tag{1.3.19}$$

with a Lagrangian

$$F_p = (u - \alpha x^2)^2 + F_n(u'), \tag{1.3.20}$$

where

$$F_n(v) = \min\{a\,v^2, \; b\,v^2 + 1\}, \quad 0 < a < b, \; \alpha > 0.$$

Note that the second term F_n of the Lagrangian F_p is a nonconvex function of u'.

The first term $(u - \alpha x^2)^2$ of the Lagrangian forces the minimizer u and its derivative u' to increase with x, until u' at some point reaches the interval of nonconvexity of $F_n(u')$. The derivative u' must vary outside of the forbidden interval of nonconvexity of the function F_n at all times.. Formally, this problem is ill-posed because the Lagrangian is not convex with respect to u' (Figure 1.5); therefore, it needs relaxation.

To find the convex envelope $\mathcal{C}F$ we must transform $F_n(u')$ (in this example, the first term of F_p (see (1.3.20)) is independent of u' and it does not change after the convexification). The convex envelope $\mathcal{C}F_p$ is equal to

$$\mathcal{C}F_p = (u - \alpha x^2)^2 + \mathcal{C}F_n(u').$$

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

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

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

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

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

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

and we calculate the relaxed Lagrangian:

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

that linearly depends on $v = u'$ in the region of nonconvexity of F.

The relaxed problem has the form

$$\min_u \int \mathcal{C}F_p(x, u, u'),$$

where

$$\mathcal{C}F_p(x, u, u') = \begin{cases} (u - \alpha x^2)^2 + a(u')^2 & \text{if } \ |u'| \le v_1, \\ (u - \alpha x^2)^2 + 2u'\sqrt{\frac{ab}{a-b}} - \frac{b}{a-b} & \text{if } \ v_1 \le |u'| \le v_2, \\ (u - \alpha x^2)^2 + b(u')^2 + 1 & \text{if } \ |u'| \ge v_2. \end{cases}$$

Note that the variables $u,\ v$ in the relaxed problem are the averages of the original variables; they coincide with those variables everywhere when $\mathcal{C}F = F$. The Euler equation of the relaxed problem is

$$\begin{cases} au'' - (u - \alpha x^2) = 0 & \text{if } \ |u'| \le v_1, \\ (u - \alpha x^2) = 0 & \text{if } \ v_1 \le |u'| \le v_2, \\ bu'' - (u - \alpha x^2) = 0 & \text{if } \ |u'| \ge v_2. \end{cases}$$

The Euler equation is integrated with the boundary conditions shown in (1.3.19). Notice that the Euler equation degenerates into an algebraic equation in the interval of convexification. The solution u and the variable $\frac{\partial}{\partial u'}\mathcal{C}F$ of the relaxed problem are both continuous everywhere.

Integrating the Euler equations, we sequentially meet the three regimes when both the minimizer and its derivative monotonically increase with x (see Figure 1.5). If the length z of the interval of integration is chosen sufficiently large, one can be sure that the optimal solution contains all three regimes; otherwise, the solution may degenerate into a two-zone solution if $u'(x) \leq v_2 \ \forall x$ or into a one-zone solution if $u'(x) \leq v_1 \ \forall x$ (in the last case the relaxation is not needed; the solution is a classical one).

Let us describe minimizing sequences that form the solution to the relaxed problem. Recall that the actual optimal solution is a generalized curve in the region of nonconvexity; this curve consists of infinitely often alternating parts with the derivatives v_1 and v_2 and the relative fractions $m(x)$ and $(1 - m(x))$:

$$v = \langle u'(x) \rangle = m(x)v_1 + (1 - m(x))v_2, \quad u' \in [v_1, v_2], \tag{1.3.21}$$

where $\langle \ \rangle$ denotes the average, u is the solution to the original problem, and $\langle u \rangle$ is the solution to the homogenized (relaxed) problem.

The Euler equation degenerates in the second region into an algebraic one $\langle u \rangle = \alpha x^2$ because of the linear dependence of the Lagrangian on $\langle u \rangle'$ in this region. The first term of the Euler equation,

$$\frac{d}{dx}\frac{\partial F}{\partial \langle u \rangle'} \equiv 0 \quad \text{if} \ \ v_1 \leq | \langle u \rangle' | \leq v_2,$$

vanishes at the optimal solution.

The variable m of the generalized curve is nonzero in the second regime. This variable can be found by differentiation of the optimal solution:

$$(\langle u \rangle - \alpha x^2)' = 0 \quad \Longrightarrow \quad \langle u \rangle' = 2\alpha x.$$

This equality, together with (1.3.21), implies that

$$m = \begin{cases} 0 & \text{if} \ \ |u'| \leq v_1, \\ \frac{2\alpha}{v_1 - v_2}x - \frac{v_2}{v_1 - v_2} & \text{if} \ \ v_1 \leq |u'| \leq v_2, \\ 1 & \text{if} \ \ |u'| \geq v_2. \end{cases} \tag{1.3.22}$$

Variable m linearly increases within the second region (see Figure 1.5). Note that the derivative u' of the minimizing generalized curve at each point x lies on the boundaries v_1 or v_2 of the forbidden interval of nonconvexity of F; the average derivative varies only due to varying of the fraction $m(x)$ (see Figure 1.5).

1.3.5 Null-Lagrangians and Convexity

The convexity requirements of the Lagrangian F that follow from the Weierstrass test are in agreement with the concept of null-Lagrangians (see, for example (Strang, 1986)).

Definition 1.3.3 The Lagrangians $\phi(x, \mathbf{u}, \mathbf{u}')$ for which the Euler equation (1.2.12), (1.2.11) identically vanishes are called *Null-Lagrangians.*

It is easy to check that null-Lagrangians in one-dimensional variational problems are linear functions of $\mathbf{u}'$. Indeed, the Euler equation is a second-order differential equation with respect to $\mathbf{u}$:

$$\frac{d}{dx}\left(\frac{\partial}{\partial \mathbf{u}'}\phi\right) - \frac{\partial}{\partial \mathbf{u}}\phi = \frac{\partial^2 \phi}{\partial (\mathbf{u}')^2} \cdot \mathbf{u}'' + \frac{\partial^2 \phi}{\partial \mathbf{u}' \partial \mathbf{u}} \cdot \mathbf{u}' + \frac{\partial^2 \phi}{\partial \mathbf{u} \partial x} - \frac{\partial \phi}{\partial \mathbf{u}} \equiv 0.$$

The coefficient of $\mathbf{u}''$ is equal to $\frac{\partial^2 \phi}{\partial (\mathbf{u}')^2}$. If the Euler equation holds identically, this coefficient is zero, and therefore $\frac{\partial G}{\partial \mathbf{u}'}$ does not depend on $\mathbf{u}'$. Hence, ϕ linearly depends on $\mathbf{u}'$:

$$\begin{aligned} \phi(x, \mathbf{u}, \mathbf{u}') &= \mathbf{u}' \cdot A(\mathbf{u}, x) + B(\mathbf{u}, x); \\ A &= \tfrac{\partial^2 \phi}{\partial \mathbf{u}' \partial \mathbf{u}}, \quad B = \tfrac{\partial^2 \phi}{\partial \mathbf{u} \partial x} - \tfrac{\partial \phi}{\partial \mathbf{u}}. \end{aligned}$$

In addition, if the equality

$$\frac{\partial A}{\partial x} = \frac{\partial B}{\partial \mathbf{u}}$$

holds, then the Euler equation vanishes identically. In this case, ϕ is a null-Lagrangian.

Example 1.3.7 Function $\phi = u\, u'$ is the null-Lagrangian. We have

$$\frac{d}{dx}\left(\frac{\partial}{\partial u'}\phi\right) - \frac{\partial}{\partial u}\phi = u' - u' \equiv 0.$$

Consider a variational problem with the Lagrangian F,

$$\min_{\mathbf{u}} \int_0^1 F(x, \mathbf{u}, \mathbf{u}').$$

Adding a null-Lagrangian to the given Lagrangian does not affect the Euler equation of the problem. The family of problems

$$\min_{\mathbf{u}} \int_0^1 \left(F(x, \mathbf{u}, \mathbf{u}') + t\phi(x, \mathbf{u}, \mathbf{u}')\right),$$

where t is an arbitrary number, corresponds to the same Euler equation. Therefore, each solution to the Euler equation corresponds to a family of Lagrangians $F(x, \mathbf{u}, \mathbf{z}) + t\phi(x, \mathbf{u}, \mathbf{z})$, where t is an arbitrary real number. This says, in particular, that a Lagrangian cannot be uniquely defined by the solution to the Euler equation.

The stability of the minimizer against the Weierstrass variations should be a property of the Lagrangian that is independent of t. It should be a common property of the family of equivalent Lagrangians. On the other hand, if $F(x, \mathbf{u}, \mathbf{z})$ is convex with respect to $\mathbf{z}$, then $F(x, \mathbf{u}, \mathbf{z}) + t\phi(x, \mathbf{u}, \mathbf{z})$ is also convex. Indeed, $\phi(x, \mathbf{u}, \mathbf{z})$ is linear as a function of $\mathbf{z}$, and adding the term $t\phi(x, \mathbf{u}, \mathbf{z})$ does not affect the convexity of the sum. In other words, convexity is a characteristic property of the family. Accordingly, it serves as a test for the stability of an optimal solution.

1.3.6 Duality

Legendre Transform

A useful tool in variational problems is duality. Particularly, duality allows us to effectively compute the convex envelope of a Lagrangian. For detailed exposition, we refer to (Gelfand and Fomin, 1963; Rockafellar, 1967; Rockafellar, 1997; Ekeland and Temam, 1976; Fenchel, 1949; Ioffe and Tihomirov, 1979).

The classical version of the duality relations is based on the Legendre transform of the Lagrangian. Consider the Lagrangian $L(x, u, u')$ that is convex with respect to u'. Consider an extremal problem

$$\max_{u'} \{p\, u' - L(x, u, u'\} \tag{1.3.23}$$

that has a solution satisfying the following equation:

$$p = \frac{\partial L}{\partial u'}. \tag{1.3.24}$$

The variable p is called the *dual* or *conjugate* to the “prime” variable u; p is also called the *impulse*. Equation (1.3.24) is solvable for u', because $L(., ., u')$ is convex. We have

$$u' = \phi(p, u, x).$$

These relations allow us to construct the Hamiltonian H of the system.

Definition 1.3.4 The *Hamiltonian* is the following function of u, p, and x:

$$H(x, u, p) = p\, \phi(p, u, x) - L(x, u, \phi(p, u, x)).$$

The Euler equations and the dual relations yield to exceptionally symmetric representations, called *canonical equations*,

$$u' = -\frac{\partial H}{\partial p}, \quad p' = \frac{\partial H}{\partial u}.$$

Generally, u and p are n-dimensional vectors. The canonical relations are given by $2n$ first-order differential equations for two n-dimensional vectors u and p.

The dual form of the Lagrangian can be obtained from the Hamiltonian when the variable u is expressed as a function of p and p' and excluded from the Hamiltonian. The dual equations for the extremal can be obtained from the canonical system if it is reduced to a system of n second-order differential equations for p.

Example 1.3.8 Find a conjugate to the Lagrangian

$$F(u, u') = \frac{1}{2}\sigma(u')^2 + \frac{\gamma}{2}u^2.$$

The impulse p is

$$p = \frac{\partial F}{\partial u'} = \sigma u'.$$

Derivative u' is expressed through p as

$$u' = \frac{p}{\sigma}.$$

The Hamiltonian H is

$$H = \frac{1}{2}\frac{p^2}{\sigma} - \gamma u^2.$$

The canonical system is

$$u' = \frac{p}{\sigma}, \quad p' = \gamma u,$$

and the dual form F^* of the Lagrangian is obtained from the Hamiltonian using canonical equations to exclude u, as follows:

$$F^*(p, p') = \frac{1}{2}\left(\frac{p^2}{\sigma} - \frac{1}{\gamma}\left(p'\right)^2\right).$$

The Legendre transform is an involution: The variable dual to the variable $\mathbf{p}$ is equal to $\mathbf{u}$.

Conjugate

The natural generalization of the ideas of the Legendre transform to nonconvex and nondifferentiable Lagrangians yields to conjugate variables. They are obtained by the Young–Fenchel transform (Fenchel, 1949; Rockafellar, 1966; Ekeland and Temam, 1976).

Definition 1.3.5 Let us define $L^*(\mathbf{z}^*)$–the *conjugate* to the $L(z)$–by the relation

$$L^*(\mathbf{z}^*) = \max_{\mathbf{z}} \{\mathbf{z}^* \, \mathbf{z} - L(\mathbf{z})\}, \tag{1.3.25}$$

which implies that $\mathbf{z}^*$ is an analogue of p (compare with (1.3.23)).

Let us compute the conjugate to the Lagrangian $L(x, \mathbf{y}, \mathbf{z})$ with respect to $\mathbf{z}$, treating $x, \mathbf{y}$ as parameters. If L is a convex and differentiable function of $\mathbf{z}$, then (1.3.25) is satisfied if

$$z^* = \frac{\partial L(\mathbf{z})}{\partial \mathbf{z}},$$

which is similar to (1.3.24). This similarity suggests that the Legendre transform $\mathbf{p}$ and the Young–Fenchel transform $\mathbf{z}^*$ coincide if the Legendre transform is applicable.[3]

However, the Young–Fenchel transform is defined and finite for a larger class of functions, namely, for any Lagrangian that grows not slower than an affine function:

$$L(\mathbf{z}) \geq c_1 + c_2\|\mathbf{z}\| \quad \forall \mathbf{z},$$

where c_1 and $c_2 > 0$ are constants.

Example 1.3.9 Find a conjugate to the function

$$F(x) = |x|.$$

From (1.3.25) we have

$$F^*(x^*) = \begin{cases} 0 & \text{if} \quad |x^*| < 1, \\ \infty & \text{if} \quad |x^*| > 1. \end{cases}$$

The Use of the Young–Fenchel Transform. We can compute the conjugate to $F^*(\mathbf{z}^*)$, called the *second conjugate* F^{**} to F,

$$F^{**}(\mathbf{z}) = \max_{\mathbf{z}^*} \{\mathbf{z}^* \cdot \mathbf{z} - F^*(\mathbf{z}^*)\}.$$

We denote the argument of F^{**} by $\mathbf{z}$.

If $F(\mathbf{z})$ is convex, then the transform is an involution. If $F(\mathbf{z})$ is not convex, the second conjugate is the convex envelope of F (see (Rockafellar, 1997)):

$$F^{**} = \mathcal{C}F.$$

We relax a variational problem with a nonconvex Lagrangian $L(x, \mathbf{u}, \mathbf{u}')$ by replacing it with its second conjugate:

$$\mathcal{C}_{\mathbf{v}} L(x, \mathbf{u}, \mathbf{v}) = L^{**}(x, \mathbf{u}, \mathbf{v}) = \max_{\mathbf{v}^*} \{\mathbf{v}^* \cdot \mathbf{v} - L^*(x, \mathbf{u}, \mathbf{v}^*)\}.$$

Note that $x, \mathbf{u}$ are treated as constant parameters during this calculation.

[3] Later, we will also use the notation $\mathbf{z}^{\text{dual}}$ for the adjoint variable denoted here as $\mathbf{z}^*$.

1.4 Conclusion and Problems

We have observed the following:

- A one-dimensional variational problem has the fine-scale oscillatory minimizer if its Lagrangian $F(x, u, u')$ is a nonconvex function of its third argument.
- Homogenization leads to the relaxed form of the problem that has a classical solution and preserves the cost of the original problem.
- The relaxed problem is obtained by replacing the Lagrangian of the initial problem by its convex envelope. It can be computed as the second conjugate to F.
- The dependence of the Lagrangian on its third argument in the region of nonconvexity does not effect the relaxed problem.

To relax a one-dimensional variational problem we have used two ideas. First, we replaced the function with its convex envelope and got a stable extension of the problem. Second, we proved that the value of the integral of the convex envelope $\mathcal{C}F(\mathbf{v})$ of a given function is equal to the value of the integral of this function $F(\mathbf{v})$ if its argument $\mathbf{v}$ is a zigzag curve. We use the Carathéodory theorem, which tells that the number of subregions of constancy of the argument is less than or equal to $n+1$, where n is the dimension of the argument of the Lagrangian.

In principle, this construction is also valid for multidimensional variational problems unless the argument of the integral satisfies additional differential restrictions. However, these restrictions necessarily occur in multidimensional problems that deal with the minimization of Lagrangians that depend on gradients of some potentials or vectors of currents. The gradient of a function is not a free vector if the dimension of the space is greater than one; the field $\mathbf{e} = \nabla w$ is curlfree: $\nabla \times \mathbf{e} = 0$. Likewise, the current $\mathbf{j}$ is divergencefree: $\nabla \cdot \mathbf{j} = 0$. These differential restrictions express integrability conditions (the equality of mixed second derivatives) for potentials; they are typical for multidimensional variational problems and they do not have a one-dimensional analogue. Generally, the multidimensional problem cannot be relaxed by convexification of its Lagrangian. In this case, convexity of the Lagrangian $F(\mathbf{x}, w, \nabla w)$ with respect to the last argument is replaced by the more delicate property of *quasiconvexity*, which will be discussed in Chapter 6. Relaxation of multidimensional problems requires replacing the Lagrangian by its *quasiconvex envelope.*

Problems

1. Formulate the Weierstrass test for the extremal problem

$$\min_u \int_0^1 F(x, u, u', u'')$$

that depends on the second derivative u''.

2. Find the relaxed formulation of the problem

$$\min_{u_1,u_2} \int_0^1 \left(u_1^2 + u_2^2 + F(u_1', u_2')\right),$$
$$u_1(0) = u_2(0) = 0, \quad u_1(1) = a, \quad u_2(1) = b,$$

where $F(v_1, v_2)$ is defined by (1.3.9). Formulate the Euler equations for the relaxed problems and find minimizing sequences.

3. Find the relaxed formulation of the problem

$$\min_u \int_0^1 \left(u^2 + \min\left\{|u' - 1|,\ |u' + 1| + 0.5\right\}\right),$$
$$u(0) = 0, \quad u(1) = a.$$

Formulate the Euler equation for the relaxed problems and find minimizing sequences.

4. Find the conjugate and second conjugate to the function

$$F(x) = \min\left\{x^2,\ 1 + ax^2\right\}, \quad 0 < a < 1.$$

Show that the second conjugate coincides with the convex envelope $\mathcal{C}F$ of F.

5. Let $x(t) > 0$, $y(t)$ be two scalar variables and $f(x, y) = x\,y^2$. Demonstrate that

$$f(\langle x\rangle, \langle y\rangle) \geq \langle y\rangle^2 \langle \frac{1}{x}\rangle^{-1}.$$

When is the equality sign achieved in this relation?

Hint: Examine the convexity of a function of two scalar arguments,

$$g(y, z) = \frac{y^2}{z}, \quad z > 0.$$

This is page 34
Printer: Opaque this