Nonconvex Lagrangians: Oscillating minimizers

March 24, 2019

Contents

Figure 1: Convexity

1 Convexity: Scalar function

1.1 Convex function of a scalar argument

Convex function A real-valued continuous function $f(x)$ of a scalar argument $x \in [a, b]$ is convex on an interval if for any two points x_1 and x_2 in $[a, b]$ and any m, $0 \le m \le 1$, the inequality holds

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

which states that the graph of $f(x)$ lies below the secant line between x_1 and x_2.

Convexity in a point Another useful form of that inequality defines convexity in a point $x \in R$. If we rename the arguments in (1) as:

$$x = m\, x_1 + (1-m)x_2, \quad v_1 = x_1 - x, \quad v_2 = x_2 - x$$

inequality (1) will define the convexity of $f(x),\ x \in R$, in the point x:

$$f(x) \le m_1 f(x+v_1) + m_2 f(x+v_2), \quad \forall m_1, m_2, v_1, v_2, \quad \text{such that}$$
$$m_1 \ge 0,\ m_2 \ge 0, \quad m_1 + m_2 = 1, \quad m_1 v_1 + m_2 v_2 = 0, \tag{2}$$

Here v_1 and v_2 are any perturbations of the argument x with zero mean value, see (2).

Example 1.1 Function $f(x) = \left(x^2 - 1\right)^2$ is convex in all points outside the interval $x \notin [-1, 1]$ and is not convex inside this interval.

Jensen's inequality, integral form The definition of convexity (2) in the point x is extended to the Jensen inequality for a function $f(x)$, $x \in R$

$$f(A) \le \sum_{i=1}^{n} m_i f(A + v_i), \quad \forall m_i, v_i, \quad i = 1, \dots n, \quad \text{such that} \tag{3}$$

$$m_i \ge 0, \quad \sum_{i=1}^{n} m_i = 1, \quad \sum_{i=1}^{n} m_i v_i = 0 \tag{4}$$

We define convexity of an integrable function $f(A)$ where $A \in [a, b]$ at the point $x = A$. Instead of the sequence $v_1, \dots, v_n$ we consider a perturbation as function $v(x)$. Then Jensen inequality is naturally extended to the statement: An integrable function $f(x)$ is convex, if the inequality holds

$$f(A) \le \frac{1}{b-a} \int_a^b f(A + v(x))dx \quad \forall v(x) \quad \text{such that} \int_a^b v(x)dx = 0 \tag{5}$$

Here, function $v(x)$ has zero mean value. The relation (5) says that any perturbation $v(x)$ with zero mean value does not decrease the value of the integral.

Perturbation of a strictly convex function Function $f(x)$ is strictly convex, if the equality

$$f(A) = \frac{1}{b-a}\int_a^b f(A+v(x))dx \tag{6}$$

implies that $v(x) = 0$. In other words, any nonzero perturbation with zero mean value increases the value of the integral.

Example 1.2 Function $f(x) = x^2$ is strictly convex because

$$\begin{aligned}
&\frac{1}{b-a}\int_a^b (A+v(x))^2 dx \\
&= A^2 + \frac{2A}{b-a}\int_a^b v(x)\,dx \quad + \frac{1}{b-a}\int_a^b v^2(x)dx > A^2 \quad \text{if} \quad \int_a^b v^2(x)dx \neq 0
\end{aligned}$$

Notice, that the second term in the right-hand side in the above sum is zero because the mean value of v is zero and the third term is positive if $v(x) \neq 0$.

Example 1.3 Affine function $f(x) = c\,x + d$ is convex but not strictly convex because

$$\begin{aligned}
&\frac{1}{b-a}\int_a^b (c(A+v(x)) + d)\,dx \\
&= cA + d + \frac{c}{b-a}\int_a^b v(x)\,dx \quad = cA + d = f(A)
\end{aligned}$$

Example 1.4 Function $f(x) = |x|$ is convex everywhere but it is not strictly convex if $x \neq 0$. At the point $x = 0$, it is strictly convex.

1.2 Application to variational problems

If a Lagrangian $F(x, u, u')$ is a convex function with respect to derivative u', the optimal trajectory $u(x)$ does not exibit high-frequency oscillations. Let $J(u)$ be an integral

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

where $F(.,.,.)$ and $u(x)$ be continuous functions. Let $\psi(x)$ be a 1- periodic function $\psi(x)$ with zero mean:

$$\int_0^T \psi(x)\,dx = 0, \quad \psi(x) = \psi(x+1), \quad x \in R, \quad |\psi(x)| \leq 1$$

then function $\epsilon\psi\left(\frac{x}{\epsilon}\right)$ oscillates with the frequency $\frac{1}{\epsilon}$ and have magnitude of the order of ϵ.

Let us compute the high-frequency variation

$$u_\epsilon(x) = u(x) + \epsilon\psi\left(\frac{x}{\epsilon}\right)$$

of the solution. The variation of the solution is of the order of ϵ, and the perturbation of the derivative is of the order of one,

$$|u_\epsilon(x) - u(x)| = O(\epsilon) \quad |u'_\epsilon(x) - u'(x)| = \psi'\left(\frac{x}{\epsilon}\right)$$

The integral J at the varied function, rounded up to the terms of $O(\epsilon)$, becomes the function of varied derivative u':

$$J(u_\epsilon) = \int_a^b F\left(x, u, u' + \psi'\left(\frac{x}{\epsilon}\right)\right) dx + O(\epsilon)$$

Since $\psi(x)$ is periodic, the integral

$$\int_x^{x+n\epsilon} \psi'\left(\frac{x}{\epsilon}\right) = 0, \quad n = 1, \ldots, n, \ldots$$

and

$$\int_a^b \psi'\left(\frac{x}{\epsilon}\right) = O(\epsilon).$$

Jensen inequality (5) shows that if $F(.,.,u')$ is convex as a function of u', any such variation ψ does not decrease the cost of the problem

$$J(u+\psi) \geq J(u) \quad \text{if } \psi \neq 0$$

If $F(.,.,u')$ is strictly convex, any perturbation ψ increases the cost

$$J(u+\psi) > J(u) \quad \text{if } \psi \neq 0$$

Variational problems with convex Lagrangians are stable to high-frequency variations.

If Lagrangian is linear with respect to u' in an interval $u' \in [a,b]$ and the varied trajectory also belongs this interval, $u'_\epsilon \in [a,b]$, the Lagrangian is invariant to the variation.

1.3 Convex envelope

Assume that a differentiable function $f(x)$ grows superlinearly and is bounded from below

$$\lim_{|x|\to\infty} \frac{f(x)}{|x|} = \infty, \quad \exists c: \ f(x) \geq c \quad \forall x \in R$$

Assume also that $f(x)$ is nonconvex and Jensen inequality (5) is not valid. There exist perturbations that make the integral in the right-hand side of (5)

Figure 2: Convex envelope

smaller than $f(A)$. A natural question arises: Find a perturbation $v(x)$ with zero mean that delivers minimum of this integral

$$\mathcal{C}f(A) = \min_{v(x)} \frac{1}{b-a} \int_a^b f(A + v(x))dx \quad \text{subject to} \ : \ \int_a^b v(x)dx = 0 \tag{7}$$

This minimum $\mathcal{C}f(A)$ is called the *convex envelope* of $f(A)$.

It is geometrically clear that optimal perturbation $v(x)$ is piece-wise constant and takes no more than two values. Convex envelope $\mathcal{C}f(x)$ of $f(x)$ either coincides with $f(x)$ or is a linear function on an interval of non-convexity; it is supported by two boundary points v_1, v_2 of this interval that have the same derivative. Because the component of the convex envelope being a linear function is supported by two points, optimal $v(x)$ takes no more than two values. The convex envelope is defined as

$$\mathcal{C}f(x) = \min_{v_1,v_2,t} \left(m_1 f(x + v_1) + m_2 f(x + v_2)\right), \tag{8}$$

$$m_1 \geq 0, \quad m_2 \geq 0, \quad m_1 + m_2 = 1, \quad m_1 v_1 + m_2 v_2 = 0 \tag{9}$$

In the points x where $f(x)$ is convex, $v_1 = v_2 = 0$ and the convex envelope $\mathcal{C}f(x)$ coincides with the function itself, $\mathcal{C}f(x) = f(x)$.

Properties of convex envelopes

• The derivative $\frac{d}{dx}\mathcal{C}f(x)$ of $\mathcal{C}f(x)$ monotonically increases; it coincides with $f'(x)$ in the intervals where $f(x)$ is convex and is constant in the intervals of non-convexity of $f(x)$.

• The second derivative $\frac{d^2}{dx^2}\mathcal{C}f(x)$ is nonnegative; it is equal to zero in the interval of non-convexity.

• One can show that $\mathcal{C}f(x)$ is the maximal convex function that is smaller than or equal to $f(x)$ in each point x.

1.4 Examples

Example 1.5 Function

$$f(x) = \left(x^2 - 1\right)^2, \quad x \in R$$

is not convex in the interval $(-1, 1)$, and is convex outside of this interval. Convex envelope of function $f(x)$ is

$$\mathcal{C}f(x) = \begin{cases} f(x) & |x| \geq 1 \\ 0 & |x| < 1 \end{cases}$$

the supporting points are $x_{1,2} = \pm 1$. In these points, the function and its derivative coincide with the convex envelope and its derivative, respectively, $\mathcal{C}f(x)|_{x=\pm 1} =$

$f(x)|_{x=\pm 1}$ and $\mathcal{C}f'(x)|_{x=\pm 1} = f'(x)|_{x=\pm 1}$. The derivatives of $f(x)$ and $\mathcal{C}f(x)$ are shown at Figure 34

Example 1.6 Consider the nonconvex function $F(v)$ called a *two-well* function

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

F is the minimum of two convex functions (wells).

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

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

The next example deals with more general case:

Example 1.7 Consider a two-well function

$$F(v) = \min\{W_1(v), W_2(v)\}, \quad W_1 = av^2, \; W_2 = b\,v^2 + 1, \tag{10}$$

where parameters are arranged as $0 < a < b$.

Compute convex envelope $\mathcal{C}F(v)$. It coincides with either the graph of the original function or with an affine function $l(v) = A\,v + B$ that touches the original graph in two points. This affine function can be found as the common tangent $l(v)$ to both convex branches (wells) of $F(v)$.

Recall that equation of the tangent line to a convex curve $g(v)$ is

$$l(v) - g(v_s) = g'(v_s)(v - v_s),$$

where v_s is the supporting point where the tangent touches the graph of $g(v)$.

Let v_1 and v_2 be the supporting points or the points where $\mathcal{C}F(v)$ touches $F(v)$. Compute the values of the common tangent $l(v)$ in the supporting points:

$$l'(v_1) = \left.\frac{dW_1}{dv}\right|_{v=v_1} = 2a\,v_1, \quad l'(v_2) = \left.\frac{dW_2}{dv}\right|_{v=v_2} = 2b\,v_2.$$

where the supporting points v_1 and v_2 belong to the corresponding wells. The equation $l'(v_1) = l'(v_2)$ gives one relation between v_1 and v_2

$$av_1 = bv_2 \tag{11}$$

From (10), we write tangent lines to each well:

$$\begin{cases} l(v) = a\,v_1^2 + 2a\,v_1(v - v_1), & v_1 \in W_1 \\ l(v) = (b\,v_2^2 + 1) + 2b\,v_2(v - v_2), & v_2 \in W_2 \end{cases} \tag{12}$$

Setting $v = 0$, we obtain the the second relation:

$$a\,v_1^2 = b\,v_2^2 - 1 \tag{13}$$

and solve (11), (13) for the coordinates of the supporting points v_1 and v_2:

$$v_1 = \sqrt{\frac{b}{a(a-b)}}, \quad v_2 = \sqrt{\frac{a}{b(a-b)}}, \tag{14}$$

Using (12) and (14) we compute linear component of the envelope and the convex envelope itself:

$$\mathcal{C}F(v) = \begin{cases} av^2 & \text{if } \ |v| < v_1, \\ 2v\sqrt{\frac{ab}{a-b}} - \frac{b}{a-b} & \text{if } \ v \in [v_1, \ v_2], \\ 1 + bv^2 & \text{if } \ |v| < v_2 \end{cases} \tag{15}$$

$\mathcal{C}F$ linearly depends on v in the interval of non-convexity of F and coincides with F outside of this interval.

2 Relaxation of Nonconvex Problems with scalar minimizer

If Lagrangian $F(x, u, u')$ that is nonconvex with respect to the third argument u' at some part of the stationary trajectory $u(x)$, the Weierstrass test fails, and the stationary trajectory cannot be optimal. Here we investigate variational problems with such Lagrangians.

Consider the Lagrangian as a function of three real arguments $F(x, u, z)$ that is

(a) a nonconvex function of its third argument;

(b) is bounded from below (say, by zero),

$$F(x, u, z) \geq 0 \quad \forall x, u, z; \tag{16}$$

(c) grows superlinearly:

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

Then the infimum I_0

$$I_0 = \inf_u J(u), \quad J(u) = \int_a^b F(x, u, u')dx$$

is nonnegative, $I_0 \geq 0$ and the minimizer has a bounded derivative.

We can construct a minimizing sequence $\{u^s\}$ such that $I(u^s) \to I_0$. The minimizing sequence $\{u^s\}$ consists of continuous functions with bounded derivatives; see [**?**].

The differentiable minimizer (if it exists) is a solution to the Euler equation. Besides, it satisfies an additional independent inequality, the Weierstrass test, see Chapter . The inequality states that the Lagrangian at $F(., ., z)$ as a function of the third argument $z = u'$ is convex everywhere at the optimal trajectory

Figure 3: Graph of nonconvex $G(v)$

$u(x)$. However, Euler equation does not leave freedom to choose the derivative of the solution. If $F(.,.,z)$ is not convex Weierstrass condition cannot be satisfied.

The derivative u' of a minimizer cannot belong to intervals of non-convexity of F (we call them "forbidden intervals").Recall that the Weierstrass test is based on adding a local perturbation to a stationary minimizer. The perturbation is an infinitesimal zig-zag of the trajectory. If such perturbations decrease the cost of the problem, the minimizer fails the Weierstrass test. The test also hints of the type of minimizers for the problems with non-convex Lagrangians. We will demonstrate that a minimizing sequence tends to a "generalized function" that consists of infinitesimal zigzags. The limiting generalized function $u(x)$ has a dense set of points of discontinuity of the derivative. This exotic limit can be effectively described by passing to the averages by the so-called homogenization procedure; we end up at the relaxed formulation of the problem. Here we give a brief description of this procedure mainly by working on several examples.

2.1 A non-convex problem

Consider a simple variational problem that yields to the generalized solution

$$I_0 = \inf_u J(u), \quad J(u) = \int_0^1 F(u, u')dx, \quad u(0) = u(1) = 0 \tag{17}$$

with Lagrangian $F(u, u') = u^2 + G(u')$ where

$$G(v) = \begin{cases} (v-1)^2, & \text{if } v \geq \frac{1}{2} \\ \frac{1}{2} - v^2 & \text{if } -\frac{1}{2} \leq v \leq \frac{1}{2} \\ (v+1)^2 & \text{if } v \leq -\frac{1}{2} \end{cases} . \tag{18}$$

The graph of the function $G(v)$ is presented in Figure 3. The Lagrangian $F = u^2 + G(u')$ penalizes the trajectory $u(x)$ for having the magnitude $|u'(x)|$ of the derivative different from one and also penalizes the deviation of the trajectory $u(x)$ from zero. These contradictory requirements cannot be resolved in the class of classical trajectories. Indeed, a differentiable minimizer satisfies the Euler equation

$$\begin{array}{ll} u'' - u = 0 & \text{if} \quad |u'| \geq \frac{1}{2} \\ u'' + u = 0 & \text{if} \quad |u'| \leq \frac{1}{2}. \end{array} \tag{19}$$

The Lagrangian $F(u, u')$ is nonconvex in the interval $u' \in (-1,\ 1)$ (see **??**). The Weierstrass test that requires convexity of $G(v)$ supplements the Euler equation (19) with the inequality

$$u' \notin (-1,\ 1) \quad \text{at the optimal trajectory.} \tag{20}$$

Euler equation does not show how to choose the trajectory that satisfies (20) avoiding the forbidden interval.

Remark 2.1 Notice that the second regime in (19) is never optimal because it is realized inside of the forbidden interval $u' \in (-1,\ 1)$. Moreover, the form of Lagrangian in the whole interval of non-convexity can be arbitrarily changed as long as it remains non-convex there; such transformation does not affect the minimizer.

Minimizing sequence We construct a minimizing sequence for problem (17) without a reference to Euler equation. The infimum of (17) is nonnegative, $\inf_u I(u) \geq 0$. Therefore, any sequence u^s such that

$$\lim_{s\to\infty} I(u^s) = 0 \tag{21}$$

is a minimizing sequence.

(i) Consider a set $\tilde{U}$ of functions $\tilde{u}^s(x)$ that belong to the boundary of the *forbidden interval* (20) of nonconvexity of $G(u')$; the derivative $\tilde{u}'(x)$ of these function is equal to ± 1:

$$\tilde{U} = \{\tilde{u}'(x) : \ \tilde{u}'(x) = \pm 1, \quad \forall x\}$$

The functions $\tilde{u}^s(x)$ make the second term (18) in the Lagrangian vanish,

$$G(\tilde{u}') = \min\{(\tilde{u}'-1)^2,\ (\tilde{u}'+1)^2\} = 0, \quad \forall \tilde{u}' \in \tilde{U}$$

and the problem becomes

$$I(\tilde{u}^s, (\tilde{u}^s)') = \int_0^1 (\tilde{u}^s)^2 dx. \tag{22}$$

(ii) Next, construct the minimizing sequence: The first term of it is a triangle

$$\tilde{u}^{(1)}(x) = \begin{cases} x, & x \in [0, \frac{1}{2}] \\ 1-x, & x \in [\frac{1}{2}, 1] \end{cases}.$$

We compute the cost of the problem J and the range r of $u(x)$

$$J(\tilde{u}^{(1)}) = 2\int_0^{\frac{1}{2}} x^2 dx = \frac{2}{3}\frac{1}{2^3} = \frac{1}{12},$$

and

$$\tilde{u}^{(1)}(x) \in r_1, \quad r_1 = \left[0, \frac{1}{2}\right]$$

The second term $\tilde{u}^{(2)}(x)$ consists of two sequential triangles in twice smaller scale

$$\tilde{u}^{(2)}(x) = \begin{cases} \tilde{u}^{(1)}(\frac{x}{2}), & x \in [0, \frac{1}{2}] \\ \tilde{u}^{(1)}(\frac{x}{2} - \frac{1}{2}), & x \in [\frac{1}{2}, 1] \end{cases}$$

We compute

$$J(\tilde{u}^{(2)}) = 4\int_0^{\frac{1}{4}} x^2 dx = \frac{1}{3\,4^2} = \frac{1}{4} J(\tilde{u}^{(1)}), \quad \tilde{u}^{(2)}(x) \in r_1, \quad r_2 = \left[0, \frac{1}{4}\right]$$

Figure 4: f02.3

Continuing this procedure, we construct $\tilde{u}^{(n)}(x)$ as the sequence of piecewise linear functions that forms a chain of n triangles along the x-axis. We compute

$$J(\tilde{u}^{(n)}) = \frac{1}{3(2n)^2}, \quad \tilde{u}^{(n)}(x) \in r_n, \quad r_n = \left[0, \frac{1}{2n}\right] \quad \forall x \tag{23}$$

The term $\tilde{u}^s$ oscillates near zero if the derivative $(\tilde{u}^s)'$ changes its sign on small intervals of equal length. Cost $J(\tilde{u}^s)$ depends on the number of switching points and tends to zero when the number of these points increases, see (23). 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 $s \to \infty$.

f2.30.4

Note that 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. Tamely, the derivative must keep the absolute value equal to one, but the function itself must be arbitrarily close to zero:

$$|(\tilde{u}^s)'| = 1 \quad \forall x \in [0,1], \quad \max_{x\in[0,1]} |\tilde{u}^s| \to 0 \quad \text{as } s \to \infty. \tag{24}$$

The limit has zero L_2 norm and unit L_2 norm of the derivative

$$\lim_{s\to\infty} \int_0^1 (\tilde{u}^s) dx = 0, \quad \int_0^1 (\tilde{u'}^s) dx = 1,$$

but the norm in $C_1[0,\ 1]$.

$$\lim_{s\to\infty} \int_0^1 (\tilde{u}^s)^2 dx = 0, \quad \lim_{s\to\infty} \int_0^1 (\tilde{u'}^s)^2 dx = 1, \quad \lim_{s\to\infty} \int_0^1 \tilde{u'}^s dx = 0$$

Notice that $\tilde{u'}^s$ in this solution takes one of two values $u'(x) = \pm 1$; these values are the supporting points of the convex envelope $\mathcal{C}G(u')$ of $G(u')$, or the points where the envelope touches $G(u')$:

$$\mathcal{C}G(u') = G(u'), \quad 0 = \frac{d}{du'}\mathcal{C}G(u') = \frac{d}{du'}G(u')$$

2.2 Relaxed problem

The same type of oscillating solution occurs in a more general problem with non-convex Lagrangian. Consider Lagrangian $F(x, u, u')$ where $u(x)$ is a scalar function. Assume that it is bounded from below, is of superlinear growth with respect to u', and that it is non-convex function of u' in a finite interval $[\rho_1(u, x), \rho_2(u, x)]$. The Weierstrass test is not satisfied for u' in the forbidden interval, therefore

$$u' \notin [\rho_1, \rho_2]$$

When u' reaches an endpoint of this interval, it must instantly jump to the other end of it. At the jump point, the Lagrangian satisfies Weierstrass-Erdmann condition

$$\left.\frac{\partial F}{\partial u'}\right|_{u'=\rho_1} = \left.\frac{\partial L}{\partial u'}\right|_{u'=\rho_2} \tag{25}$$

After u' jumps from ρ_1 to ρ_2, it may jump back, and again. When the jumps occur infinitely fast, the minimizer becomes a generalized curve.

Relaxation We describe a generalized zig-zag type solution to a variational problem. Consider a small interval $x = [x_0, x_0 + \epsilon]$ and a function $u_\epsilon(x)$. with the derivative $u'_\epsilon(x)$ that takes two values

$$u'_\epsilon(x) = \begin{cases} \rho_1, \ x \in [x_0, x_0 + m\epsilon) \\ \rho_2, \ x \in [x_0 + m\epsilon, x_0 + \epsilon) \end{cases}. \tag{26}$$

Here, $m \in [0, 1]$ is a length (measure) of the fraction of the interval where $u'_\epsilon(x) = \rho_1$. The average (homogenized) value $v_h(x_0)$ of the derivative of $u(x)$ at the interval $[x_0, x_0 + \epsilon]$ is

$$u'_h(x_0) = \frac{1}{\epsilon}\int_{x_0}^{x_0+\epsilon} u'_\epsilon(x)dx = m\rho_1 + (1-m)\rho_2, \quad u'_\epsilon(x) \in [\rho_1, \rho_2]; \tag{27}$$

it depends of ρ_1, ρ_2 and the fraction $m \in [0, 1]$.

The function $u_\epsilon(x)$ is a zig-zag curve. The value $u_\epsilon(x)$ in the point $x_0 + \epsilon$ is

$$u_\epsilon(x_0 + \epsilon) = u_\epsilon(x_0) + \int_{x_0}^{x_0+\epsilon} u'_\epsilon(x)dx = \epsilon v_h(x_0) = o(\epsilon)$$

which gives (rounding to the terms of the order of ϵ)

$$\frac{u_\epsilon(x_0 + \epsilon) - u_\epsilon(x_0)}{\epsilon} = v_h(x).$$

To obtain a homogenized description of the solution, we approximate the zig-zag function $u_\epsilon(x)$ with the smooth function $u_h(x)$ with the constant derivative $v_h(x)$, $u'_h(x) = v(x)$. Function $u_h(x)$ coincides with $u_\epsilon(x)$ in the points x_0 and $x_0 + \epsilon$ and is linear in between:

$$u_h(x) = u_\epsilon(x_0) + (x - x_0)v_h \quad \forall x \in [x_0, x_0 + \epsilon];$$

$u_h(x)$ replaces the triangular zig-zag with the linear function that connects the points of its base. The difference $|u_h(x) - u_\epsilon(x_0)|$ is of the order ϵ.

Repeating this smoothing procedure for a curve $u(x)$ with ϵ-periodic zig-zags, we approximate it with a smooth piece-wise linear function $u_h(x)$ that tends to a differentiable limit when $\epsilon \to 0$. The L_2-norm of the difference between the minimizer and its approximation tends to zero when $\epsilon \to 0$.

$$\int_a^b (u_\epsilon(x) - u_h(x))^2 \, dx = O(\epsilon) \to 0, \quad \text{if } \epsilon \to 0$$

The point-wise derivative of $u_\epsilon(x)$ (26) takes the values ρ_1 and ρ_2; the cost J of the variational problem with the Lagrangian $F(x,u,u')$ in an interval $[x = x_0, x_0+\epsilon]$ is

$$J(x_0,\epsilon) = \int_{x_0}^{x_0+\epsilon} \hat{F} dx = \epsilon \left[\int_{x_0}^{x_0+m\epsilon} F(x,u,\rho_1) + \int_{x_0+m\epsilon}^{x_0+\epsilon} F(x,u(x),\rho_2) + O(\epsilon) \right]$$

When $\epsilon \to 0$, the cost tends to the convex combination of $F(x,u,\rho_1)$ and $F(x,u(x),\rho_2)$:

$$J(x_0,\epsilon) = \epsilon \left[mF(x,u,\rho_1) + (1-m)F(x,u(x),\rho_2) + O(\epsilon) \right]$$

Optimizing $J(x_0,\epsilon)$ with respect of ρ_1, ρ_2, and m, we find the best value I_ϵ of the cost at the zigzag minimizers:

$$\begin{aligned} I_\epsilon &\geq \min_{m\in[0,1]} \min_{\rho_1,\rho_2\in\mathcal{M}} \epsilon \mathcal{R}F(x,u,\rho_1,\rho_2,m) \quad \text{where} \\ \mathcal{R}F(x,u,\rho_1,\rho_2,m) &= [mF(x,u,\rho_1) + (1-m)F(x,u,\rho_2)] \\ \mathcal{M} &= \{m\rho_1 + (1-m)\rho_2 = u'\} \end{aligned}$$

e observe, that the relaxed cost $\mathcal{R}F$ is the convex envelope $\mathcal{C}_v F$ of $F(x,u,v)$ with respect to its third argument $v = u'$.

$$\mathcal{R}F = \mathcal{C}_v F \tag{28}$$

In the region of convexity, $m = 1$, the supporting points coincides $\rho_1 = \rho_2 = u'$, and $\mathcal{R}F = F$.

This procedure defines the *relaxed* problem for the smoothen minimizer $u = \lim_{\epsilon\to 0} u_h$.

$$I_R = \min_u \int_a^b \mathcal{C}F(x,u,u')dx \tag{29}$$

$$\mathcal{C}F(x,u,u') = \min_{m\in[0,1]} \min_{\rho_1,\rho_2\in\mathcal{M}} \mathcal{R}F(x,u,\rho_1,\rho_2,m) \tag{30}$$

$$\mathcal{R}F(x,u,\rho_1,\rho_2,m) = [mF(x,u,\rho_1) + (1-m)F(x,u,\rho_2)] \tag{31}$$

$$\mathcal{M} = \{m\rho_1 + (1-m)\rho_2 = u'\} \tag{32}$$

The obtained relaxed formulation has the following features

1. The convex envelope of the Lagrangian is not larger than the Lagrangian, therefore the solution u of the relaxed problem corresponds to a lower bound of the cost of the initial problem, $I \geq I_R$

2. The cost of the relaxed problem is realized on a special sequence, (zig-zag minimizers), therefore it is not smaller that the infimum of the problem cost over all minimizing sequences, $I_R \geq I$.

   Properties 1 and 2 mean that $I_R = I$, the cost of the relaxed problem is equal to the cost of the original problem.

3. The Lagrangian of the relaxed problem is the convex envelope of F and is obviously convex. Therefore, the Weierstrass condition is satisfied, and the solution is stable against fine-scale perturbations, unlike the Lagrangian of the original problem.

   Such transformation is called the *minimal extension* of the ill-posed problem.

Remark 2.2 The derivative of a minimizer never takes values in the region of nonconvexity of $F(,,u')$. Therefore, a solution to a nonconvex problem stays the same if its Lagrangian $F(x,u,z)$ is replaced by any Lagrangian $\mathcal{N}F(x,u,z)$ that (i) coincides with F in the region of convexity and (ii) it nonconvex in in the region of non-convexity. The behavior of the Lagrangian in the nonconvex region is simply of no importance. The convex envelope $\mathcal{C}F$ is the lower bound of such Lagrangians, it is on the boundary of convexity (is convex but not strongly convex) in the region of non-convexity of F.

Properties of the Relaxed Problem

1. The supporting points of F where F touches touching the convex envelope $\mathcal{C}F$, have the common value of the derivative. Therefore, Weierstrass-Erdmann conditions (25) are safisfied

2. Minimizing sig-zag sequence is parametrized: it depends on three parameters ρ_1, ρ_2, and m. They define the derivative

$$u_h'(x) = m(x)\rho_1(x) + (1 - m(x))\rho_2(x) \tag{33}$$

   of the relaxed minimizer, see (27). Thus, the number of minimizers in the relaxed problem increases. Instead of one minimizer $u(x)$ in the original problem, the minimizer is controlled by three "slow-varied" functions $\rho_1(x), \rho_2(x)$ and $m(x)$

3. In the forbidden region, Euler equation degenerates. If the convex envelope does not coincide with G, it is linear with respect to u'

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

   This representation implies that the Euler equation degenerates into an algebraic equation $\frac{\partial}{\partial u} b(x,u) = 0$.

   Optimal fraction m does not enters this representation. It can be found by diffferentiation of the Euler equation along the optimal trajectory:

$$\frac{d}{dx}\frac{\partial}{\partial u} b(x,u) = \frac{\partial^2 b(x,u)}{\partial x \partial u} + \frac{\partial^2 b(x,u)}{\partial u^2} u' = 0$$

   Together with the representation of u' (33) it defines optimal $m(x)$

$$m = \frac{1}{\rho_1 - \rho_2}\left[\left(\frac{\partial^2 b(x,u)}{\partial x \partial u}\right)\left(\frac{\partial^2 b(x,u)}{\partial u^2}\right)^{-1} - \rho_2\right] \tag{34}$$

   in the region of non-convexity.

Figure 5: f02.4

2.3 Examples: Solutions to Nonconvex Problems

A Two-Well Lagrangian We turn to a more advanced example of the relaxation of a nonconvex variational problem. This example highlights more properties of relaxation and introduces piecewise quadratic Lagrangians.

Example 2.1 Consider the minimization problem

$$\min_{u(x)} \int_0^q F(x,u,u'), \quad u(0)=0,\ u'(q)=0 \tag{35}$$

with a Lagrangian

$$F = (u-\gamma x^2)^2 + G(u'), \tag{36}$$

where

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

The first term $(u-\gamma x^2)^2$ 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 $G(v)$, to pass this interval, and increase further. The term G is a nonconvex function of $v=u'$. The derivative u' stays outside of the forbidden interval of nonconvexity of the function G. The convex envelope $\mathcal{C}G(v)$ of $G(v)$ is (see Example 34)

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

where

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

f2.4 0.4

The relaxed problem has the form

$$\min_u \int \mathcal{C}F(x,u,u')dx, \tag{37}$$

where

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

Recall that the variables u, u' 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

$$\begin{cases} au'' - (u - \gamma x^2) = 0 & \text{if } \ |u'| \leq v_1, \\ (u - \gamma x^2) = 0 & \text{if } \ v_1 \leq |u'| \leq v_2, \\ bu'' - (u - \gamma x^2) = 0 & \text{if } \ |u'| \geq v_2. \end{cases} \tag{38}$$

where v_1 and v_2 are defined in (12). The boundary conditions are shown in (35).

Notice that the Euler equation degenerates into an algebraic equation in the interval where convex envelope of F does not coincide with F.

Integrating the Euler equations, we sequentially meet all three regimes when both the minimizer and its derivative monotonically increase with x (see **??**). If the length z of the interval of integration is sufficiently large, one sees all three regimes.

Minimizing sequence 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 $1-m(x)$, respectively:

$$u'(x) = m(x)v_1 + (1-m(x))v_2, \quad u' \in [v_1, v_2], \tag{39}$$

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

$$\frac{d}{dx}\frac{\partial F}{\partial \langle u \rangle'} \equiv 0 \quad \text{if} \quad v_1 \le |\langle u \rangle'| \le v_2, \tag{40}$$

vanishes.

Obtaining optimal fraction m The variable $m(x)$ of the generalized solution is found by differentiation of the optimal solution:

$$(u(x) - \gamma x^2)' = 0 \quad \text{or} \quad u'(x) = 2\gamma x. \tag{41}$$

Using definition (39) of the average derivative, we find

$$u'(x) = m(x)v_1 + [1-m(x)]v_2 = 2\gamma x.$$

(recall that the boundaries v_1 and v_2 of the forbidden interval are constant in the interval of nonconvexity). Solving the equality for m, we obtain

$$m = \begin{cases} 0 & \text{if} \quad |u'| = v_1, \\ \frac{2\alpha}{v_1 - v_2}x - \frac{v_2}{v_1 - v_2} & \text{if} \quad v_1 \le |u'| \le v_2, \\ 1 & \text{if} \quad |u'| = v_2. \end{cases} \tag{42}$$

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

Remark 2.3 Notice that in the region of non-convexity, the supporting points v_1 and v_2 are constant, and the fraction (measure) m varies with x. In the regions of convexity, the fraction m degenerate into zero of one, but the supporting point becomes derivative of the classical minimizer u; it varies with x.