Relaxation method in the calculus of variations

Ivan Yashchuk

1 Introduction

Question of existence of minimizers of non-lsc functionals.

The fundamental problem of the calculus of variations is

$$\inf \left\{ I(u) = \int_{\Omega} f(x, u(x), Du(x)) \mathrm{d}x \, : \, u \in u_0 + W_0^{1,p}\left(\Omega; \mathbb{R}^N\right) \right\}, \tag{P}$$

where

$\Omega \subset \mathbb{R}^n$ is a bounded open set, with Lipschitz boundary $\partial\Omega$;

$u : \Omega \to \mathbb{R}^N$,

$$u = u(x) = u\left(x_1, \ldots, x_n\right) = \left(u^1(x), \ldots, u^N(x)\right)$$

$Du \in \mathbb{R}^{N\times n}$ denotes its Jacobian matrix

$f : \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}$ is continuous, $f = f(x, u, \xi)$

$1 \leqslant p \leqslant \infty$ and $W^{1,p}\left(\Omega; \mathbb{R}^N\right)$ denotes the usual space of Sobolev functions, where

$$u^i, \frac{\partial u^i}{\partial x_j} \in L^p(\Omega), \quad i = 1, \ldots, N, j = 1, \ldots, n;$$

$u_0 \in W^{1,p}\left(\Omega; \mathbb{R}^N\right)$ is a given function;

$u \in u_0 + W_0^{1,p}\left(\Omega; \mathbb{R}^N\right)$, meaning that $u \in W^{1,p}\left(\Omega; \mathbb{R}^N\right)$ and $u = u_0$ on $\partial\Omega$ in the Sobolev sense.

In order to simplify the presentation, consider the case where there is no dependence on lower-order terms, i.e., $f(x, u, \xi) = f(\xi)$.

When dealing with nonconvex problems, the first step is the relaxation theorem, established by L.C. Young, Mac Shane, Ekeland and others. This consists in replacing the problem (P) by the so-called relaxed problem

$$\inf \left\{ \overline{I}(u) := \sup\{\mathcal{H}(u) : \mathcal{H} \leqslant I \text{ and } \mathcal{H} \text{ is weakly lsc}\} \, : \, u \in u_0 + W_0^{1,p}(\Omega) \right\} \tag{QP}$$

Therefore the direct methods, which do not apply to (P), apply to (QP). It can be shown that

$$\inf(\mathrm{P}) = \inf(\mathrm{QP})$$

and that minimizers of (P) are necessarily minimizers of (QP), the converse being false.

The second step in proving the existence of minimizers for (P) is to see if among all solutions of (QP) , if any, at least one of them is also a solution of (P). This amounts in finding $\overline{u} \in u_0 + W_0^{1,p}(\Omega)$ so that

$$\int_\Omega Cf(D\overline{u}(x))\mathrm{d}x = \inf(\mathrm{P}) = \inf(\mathrm{QP})$$

and at the same time in solving the first-order differential equation

$$Cf(D\overline{u}(x)) = f(D\overline{u}(x)) \quad \text{ a.e. } x \in \Omega.$$

In this work only the first step is dicussed, the relaxation theorem in the vector valued case.

For relaxation, there are essentially two options: one is replacing the integrand, the second (can be called relaxation-extension) is replacing the original problem with the generalized minimization problem with the help of Young measures.

First, if we only care about the infimal value of I, we can compute the relaxation $\overline{I}$ of I, which by definition is the largest lower semicontinuous functional below I. However, the minimizer of $\overline{I}$ may not say much about the minimizing sequence of our original I since all oscillations (and concentrations in some cases) have been "averaged out".

Second, we can focus on the minimizing sequences themselves and try to find a generalized limit object to a minimizing sequence that encapsulates "interesting" information. The natural candidates for such limit objects are (gradient) Young measures. Young measure theory allows one to replace a minimization problem over a Sobolev space by a generalized minimization problem over (gradient) Young measures. This generalized minimization problem always has a solution. This approach is described well in [2].

2 Relaxation of functionals

2.1 The different notions of convexity

Definition 1. (i) A function $f : \mathbb{R}^{N\times n} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is said to be *rank one convex* if

$$f(\lambda\xi + (1-\lambda)\eta) \leqslant \lambda f(\xi) + (1-\lambda) f(\eta)$$

for every $\lambda \in [0,1], \xi, \eta \in \mathbb{R}^{N\times n}$ with rank $(\xi - \eta\} \leqslant 1$.

(ii) A Borel measurable and locally integrable function $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ is said to be *quasiconvex* if

$$f(\xi) \leqslant \fint_D f(\xi + D\varphi(x))\mathrm{d}x$$

for every bounded domain $D \subset \mathbb{R}^n$, for every $\xi \in \mathbb{R}^{N\times n}$ and for every $\varphi \in W_0^{1,\infty}\left(D; \mathbb{R}^N\right)$

(iii) A Borel measurable and locally integrable function $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ is said to be quasiaffine if f and $-f$ are quasiconvex.

(iv) A function $f : \mathbb{R}^{N\times n} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is said to be polyconvex if there exists $g : \mathbb{R}^{r(n,N)} \to \mathbb{R}$ convex, such that

$$f(\xi) = g(T(\xi)),$$

where $T : \mathbb{R}^{N\times n} \to \mathbb{R}^{r(n,N)}$ is such that

$$T(\xi) = \left(\xi, \operatorname{adj}_2 \xi, \ldots, \operatorname{adj}_{n\wedge N} \xi\right).$$

$\operatorname{adj}_s\xi$ stands for the matrix of all $s \times s$ minors of the matrix $\xi \in \mathbb{R}^{N\times n}, 2 \leqslant s \leqslant n \wedge N = \min\{n, N\}$

These notions are related as f convex $\Rightarrow$ f polyconvex $\Rightarrow$ f quasiconvex $\Rightarrow$ f rank one convex. For the scalar case, all these notions reduce to the notion of convexity.

When $u : \mathbb{R} \to \mathbb{R}^N$ or $u : \mathbb{R}^n \to \mathbb{R}$, Qf is simply the convexication of f.

Note that in the definition of quasiconvexity if the inequality holds for a given domain $D \subset \mathbb{R}^n$, then it holds for every such domain D.

If the function $f : \mathbb{R}^{N\times n} \to \mathbb{R}$, i.e., f takes only finite values, is convex or polyconvex or quasiconvex or rank one convex, then it is continuous and even locally Lipschitz.

When the function f depends on lower-order terms, i.e., $f : \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}$ with $f = f(x, u, \xi)$, all the above notions are understood only with respect to the variable ξ, all the other variables being kept fixed.

The important concept from the point of view of minimization in the calculus of variations is the notion of quasiconvexity. This condition is equivalent to the fact that the functional I is *weakly lower semicontinuous* in $W^{1,p}\left(\Omega; \mathbb{R}^N\right)$ meaning that

$$I(u) \leqslant \liminf_{v\to\infty} I\left(u_v\right)$$

for every sequence $u_v \to u$ in $W^{1,p}$ (Morrey's theorem, 1952)

2.2 Envelopes

Define

$$\begin{aligned} Cf &= \sup\{g \leqslant f : g \text{ convex}\}, \\ Pf &= \sup\{g \leqslant f : g \text{ polyconvex}\}, \\ Qf &= \sup\{g \leqslant f : g \text{ quasiconvex}\}, \\ Rf &= \sup\{g \leqslant f : g \text{ rank one convex}\}, \end{aligned}$$

to be the *convex, polyconvex, quasiconvex, rank one convex envelope* of f. They are related as

$$Cf \leqslant Pf \leqslant Qf \leqslant Rf \leqslant f$$

Several representation formulas exist for computing these envelopes, we need quasiconvex envelope case here.

Theorem 1. (Dacorogna formula). If $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ is locally bounded and Borel measurable then, for every $\xi \in \mathbb{R}^{N\times n}$,

$$Qf(\xi) := \inf\left\{ ⨍_\Omega f(\xi + D\varphi(x))\mathrm{d}x : \varphi \in W_0^{1,\infty}\left(\Omega; \mathbb{R}^N\right) \right\}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain. The infimum is independent of the choice of the domain.

The Alibert-Dacorogna-Marcellini example. Here $N = n = 2$ and

$$f(\xi) = |\xi|^2 \left(|\xi|^2 - 2\gamma \det \xi\right),$$

where $|\xi|$ stands for the Euclidean norm of the matrix and $\gamma \geqslant 0$. Then

$$\begin{aligned}
&f \text{ is convex } \iff \gamma \leqslant \gamma_c = \tfrac{2}{3}\sqrt{2} \\
&f \text{ is polyconvex } \iff \gamma \leqslant \gamma_p = 1 \\
&f \text{ is quasiconvex } \iff \gamma \leqslant \gamma_q, \quad \text{where } \gamma_q > 1 \\
&f \text{ is rank one convex } \iff \gamma \leqslant \gamma_r = \tfrac{2}{\sqrt{3}}
\end{aligned}$$

2.3 Relaxation theorem

Theorem 2. Let $\Omega \subset \mathbb{R}^n$ be a bounded open set. Let $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ be Borel measurable satisfying, for $1 \leqslant p < \infty$,

$$0 \leqslant f(\xi) \leqslant \alpha_1 (1 + |\xi|^p) \quad \text{for every } \xi \in \mathbb{R}^{N\times n} \tag{1}$$

where $\alpha_1 > 0$ is a constant and for $p = \infty$ it is assumed that f is locally bounded. Let

$$Qf = \sup\{g \leqslant f : g \text{ quasiconvex}\}$$

be the quasiconvex envelope of f. Then

$$\inf(\mathrm{P}) = \inf(\mathrm{QP}).$$

More precisely, for every $u \in W^{1,p}\left(\Omega; \mathbb{R}^N\right)$, there exists a sequence $\{u^\upsilon\}_{\upsilon=1}^{\infty} \subset u_0 + W_0^{1,p}\left(\Omega; \mathbb{R}^N\right)$ such that

$$\int_\Omega f\left(Du^\nu(x)\right) \mathrm{d}x \to \int_\Omega Qf(Du(x))\mathrm{d}x \quad \text{as } \upsilon \to \infty.$$

Remark. The theorem remains also valid if the function f depends on lower-order terms, i.e., $f = f(x, u, \xi)$. The quasiconvex envelope is then to be understood as the quasiconvex envelope only with respect to the variable ξ, the other variables (x, u) being kept fixed.

Proof. Step 1. We start with an approximation of the given function u. Let $\varepsilon > 0$ be arbitrary, we can then find disjoint open sets $\Omega_1, \ldots, \Omega_k \subset \Omega, \xi_1, \ldots, \xi_k \in \mathbb{R}^{N\times n}, \gamma$ independent of ε and $\upsilon \in u + W_0^{1,p}\left(\Omega; \mathbb{R}^N\right)$ such that

$$\begin{cases} \operatorname{meas}\left[\Omega - \bigcup_{i=1}^k \Omega_i\right] \leqslant \varepsilon \\ \|u\|_{W^{1,p}}, \|\upsilon\|_{W^{1,p}} \leqslant \gamma, \quad \|u-\upsilon\|_{W^{1,1}} \leqslant \varepsilon \\ D\upsilon(x) = \xi_i \quad \text{if } x \in \Omega_i \end{cases} \tag{2}$$

By taking ε smaller if necessary we can also assume, using the continuity of Qf and the growth condition on f, that

$$\int_\Omega |Qf(Du(x)) - Qf(D\upsilon(x))|\mathrm{d}x \leqslant \varepsilon \tag{3}$$

$$0 \leqslant \int_{\Omega - \bigcup_{i=1}^k \Omega_i} [f(D\upsilon(x)) - Qf(D\upsilon(x))]\mathrm{d}x \leqslant \varepsilon. \tag{4}$$

Indeed let us discuss the case $1 \leqslant p < \infty$, the case $p = \infty$ being easier. Any quasiconvex function is locally Lipschitz continuous (Theorem 5.3 in Dacorogna book [1]). Since it also satisfies 1, then there exists $\beta > 0$ (see Proposition 2.32 in Dacorogna book [1]) such that

$$|Qf(Du) - Qf(Dv)| \leqslant \beta \left(1 + |Du|^{p-1} + |Dv|^{p-1}\right) |Du - Dv|.$$

Using the Hölder inequality we obtain

$$\begin{aligned}&\int_\Omega |Qf(Du) - Qf(Dv)| \mathrm{d}x \\ &\quad \leqslant \beta \left[\int_\Omega \left[\left(1 + |Du|^{p-1} + |Dv|^{p-1}\right)\right]^{p/(p-1)}\right]^{p/(p-1)} \left[\int_\Omega |Du - Dv|^p\right]^{1/p}\end{aligned}$$

and (3) follows from (2). The inequality (4) follows from (2) and a classical property of the integrals (since $f(Dv), Qf(Dv) \in L^1$).

Step 2. Now use Dacorogna Formula on every Ω_i to find $\varphi_i \in W_0^{1,\infty}\left(\Omega_i; \mathbb{R}^N\right)$

$$\fint_{\Omega_i} f\left(\xi_i + D\varphi_i(x)\right) \mathrm{d}x \geqslant Qf\left(\xi_i\right) \geqslant -\varepsilon + \fint_{\Omega_i} f\left(\xi_i + D\varphi_i(x)\right) \mathrm{d}x$$

Setting

$$w(x) = \begin{cases} v(x) + \varphi_i(x) & \text{if } x \in \Omega_i, i = 1, \dots, k \\ v(x) & \text{if } x \in \Omega - \bigcup_{i=1}^k \Omega_i \end{cases}$$

we get that $w \in u + W_0^{1,p}\left(\Omega; \mathbb{R}^N\right)$ and (using (4))

$$0 \leqslant \int_{\cup_{i=1}^k \Omega_i} [f(Dw(x)) - Qf(Dv(x))] \mathrm{d}x \leqslant \varepsilon \text{ meas } \left[\bigcup_{i=1}^k \Omega_i\right],$$

$$0 \leqslant \int_{\Omega - \cup_{i=1}^k \Omega_i} [f(Dw(x)) - Qf(Dv(x))] \mathrm{d}x = \int_{\Omega - \cup_{i-1}^k \Omega_i} [f(Dv(x)) - Qf(Dv(x))] \mathrm{d}x \leqslant \varepsilon.$$

In other words, combining these inequalities, we have proved that

$$0 \leqslant \int_\Omega [f(Dw(x)) - Qf(Dv(x))] \mathrm{d}x \leqslant \varepsilon(1 + \text{meas } \Omega)$$

Invoking (3), we find

$$\left|\int_\Omega [f(Dw(x)) - Qf(Du(x))] \mathrm{d}x\right| \leqslant \varepsilon(2 + \text{meas}\Omega)$$

Setting $\varepsilon = 1/v$ with $v \in \mathbb{N}$ and $u^v = w$, we have indeed obtained the theorem.

□

In the case $N = n = 1$, this result has been proved by L.C. Young and then generalized by others to the scalar case, $N = 1$ or $n = 1$, notably by Berliochi and Lasry, Ekeland, Ioffe and Tihomirov and Marcellini and Sbordone.

2.4 Rank one convexity

For a given f, is it in need of relaxation? Is $Qf = f$ or not?

A very powerful test is provided by the "layering": if f is lsc ($Qf = f$) then f must be rank one convex.

Sverak (1992) gave an example of a rank one convex function that is not quasiconvex. Therefore rank one convexity is not equaivalent to quasiconvexity.

References

[1] Bernard Dacorogna. *Direct Methods in the Calculus of Variations (Applied Mathematical Sciences)*. Springer, 2007. URL: https://doi.org/10.1007/978-0-387-55249-1.

[2] Filip Rindler. *Calculus of Variations*. Springer International Publishing, 2018. DOI: 10.1007/978-3-319-77637-8. URL: https://doi.org/10.1007/978-3-319-77637-8.