An example of non-convex minimization and an application to Newton's problem of the body of least resistance

T. Lachand-Robert M. A. Peletier

11th July 1997

Abstract

We study the minima of the functional $\int_\Omega f(\nabla u)$. The function f is not convex, the set Ω is a domain in $\mathbb{R}^2$ and the minimum is sought over all convex functions on Ω with values in a given bounded interval. We prove that u is almost everywhere 'on the boundary of convexity', in the sense that there exists no open set on which u is both strictly convex and differentiable. In particular, wherever the Gaussian curvature is defined and continuous, it is zero.

An important application of this result is the problem of the body of least resistance as formulated by Newton (where $f(p) = 1/(1+|p|^2)$ and Ω is a ball), implying that the minimizer is not radially symmetric. This generalizes a result in [3].

Résumé

On examine les minimums d'une fonctionnelle de la forme $\int_\Omega f(\nabla u)$ où f n'est pas convexe et Ω est un domaine borné de $\mathbb{R}^2$, l'ensemble des fonctions admissibles u étant restreint aux fonctions convexes à valeurs dans un intervalle fixé. On démontre que ces minimaux sont presque partout à la limite de la convexité, en ce sens qu'il n'existe pas d'ouvert où ils sont à la fois strictement convexes et différentiables.

The second author wishes to acknowledge support from EPSRC grant no. GR/L17177. Part of this work was carried out during a visit of the second author to Université Pierre et Marie Curie under the contract of the European Union 921 CHRX CT 94.

En particulier, aux points où leurs graphes possèdent une courbure gaussienne continue, elle est nulle.

Ce résultat s'applique notamment au problème de la résistance minimale de Newton (où $f(p) = 1/(1+|p|^2)$ et Ω est une boule). Il implique que le minimum n'est pas à symétrie radiale, généralisant ainsi le résultat de [3].

1 Introduction

1.1 Newton's problem

The problem of the body of minimal resistance was introduced by Sir Isaac Newton in *Principia Mathematica* [7, 5] and his treatment has generally become known as one of the first applications of variational calculus. The problem can be stated mathematically in different ways. One of them is as follows: let Ω be a given domain (an open connected set) in $\mathbb{R}^2$, and minimize

$$\int_\Omega \frac{dx}{1+|\nabla u(x)|^2} \tag{1}$$

(where $|p|^2 = p_1^2 + p_2^2$ is the Euclidean norm) over the set C_M of concave functions u defined on Ω, bounded above and below by constants 0 and $M > 0$. Here the graph of u is the shape of the body whose resistance is estimated by (1) under some simple physical assumptions. We refer the reader to [4] for a more detailed presentation of this problem, and some alternative ways to state it. It is also proved in this paper that the minimum is attained.

Newton computed a minimizing solution when Ω is a circular disc, assuming radial symmetry of the minimizer. Indeed, since the problem is invariant under rotation, one might reasonably expect that the minimizer has the same invariance. However, it was recently proved [3] that the solution calculated by Newton only minimizes (1) in the set of radially symmetric functions; there exists a negative second variation of the functional (1) in some 'non-radial' direction at this point, hence it cannot be the minimum. An important consequence is that minimizers of Newton's problem on a ball are not unique. (The main distinguishing feature is that the functional defined by (1) is not convex; indeed, the function $p \in \mathbb{R}^2 \mapsto (1+|p|^2)^{-1}$ has a Hessian matrix with a negative eigenvalue at every point.)

Unfortunately, this argument gives no information on the shape of the minima other than the lack of radial symmetry. It has been conjectured by H. Berestycki that the minimizers could be 'affine by parts'. This conjecture arises from the following result:

Lemma 1 *Assume that u is of class C^2 in $\omega \subset \Omega$ and satisfies $\det d^2u > 0$ in ω. Then u is not a minimizer of (1).*

Proof:

The proof of this Lemma is similar to the one given in [3]; since it is not essential to this paper, we will just give a sketch of it.

The crucial observation is that u is 'interior' to the set C_M, in the sense that for any function $\phi \in C_c^2(\omega)$, $u \pm \varepsilon\phi$ are both in C_M if $\varepsilon > 0$ is small enough. Therefore, if we assume that u is a minimum, then it satisfies the Euler equation

$$df(\nabla u) = 0 \qquad \text{in } \omega \tag{2}$$

where $f(p) := (1 + |p|^2)^{-1}$. A standard second variation argument gives

$$\int_\omega d^2 f(\nabla u)\nabla\phi \cdot \nabla\phi \geq 0 \tag{3}$$

for all $\phi \in C_c^2(\omega)$. The idea is to consider a function ϕ whose gradient is mainly oriented in the direction of concavity of $d^2 f(\nabla u)$. Indeed, if ω is a sufficiently small set, we can assume that the matrices $\{d^2 f(\nabla u)(x)\}_{x\in\omega}$ have a common negative direction. Supposing this direction to be the first coordinate axis, we choose $\phi(x_1, x_2) := \eta(x)\sin(nx_1)$ where η is a fixed function in $C_c^2(\omega)$ and n a sufficiently large number. It is now easy to verify that there is a contradiction with (3). □

This lemma strongly indicates that in some sense a minimizer u should satisfy $\det d^2u \equiv 0$. Unfortunately, the C^2 requirement is very strong, and nothing indicates that it is actually satisfied by a minimizer of Newton's problem. Even the radial solution is not C^2, in fact not even C^1 (it has a flat circular zone where it is maximal, and the gradient is not continuous on its boundary). Moreover, 'most' convex functions are smooth (i.e. C^1) and strictly convex, but 'almost none' of them are C^2, in the sense of Baire category (see for instance [10] and the references cited there). Since there is no regularizing effect in this sort of problems (see [6]), there is little hope that the minimizer is C^2.

Hence it is not clear what generalization of $\det d^2u \equiv 0$ should hold for minimizers. We should note that the condition $\det d^2u \equiv 0$, even if it makes sense in some open subset $\omega \subset \Omega$, does not imply that u is affine. It only indicates that the graph u has vanishing Gauss curvature; hence u could also be (locally) cylindrical or conical.

In this paper we investigate this question and prove that minimizers cannot be strictly convex and differentiable in the same open set. As we remarked above this behaviour is very unusual for a convex function, since 'most' convex functions are of class C^1 and strictly convex. Note that this result contains the statement of Lemma 1.

1.2 The general problem

We consider the problem

$$\inf_{u\in\mathcal{C}} \int_\Omega f(\nabla u)\,dx \tag{4}$$

where $\Omega \subset \mathbb{R}^2$ is an open bounded domain, $f : \mathbb{R}^2 \to \mathbb{R}_+$ is a given nonnegative smooth function, and

$$\mathcal{C} := \{u : \Omega \to [0,1];\ u \text{ convex}\}\,. \tag{5}$$

This corresponds to the problem of least resistance as stated before by changing u to $(M-u)/M$ and setting $f(p) = 1/(M^2+|p|^2)$.

Since the set $\mathcal{C}$ is compact in the $H^1_{\rm loc}(\Omega)$ topology [4], the minimum is always attained, since f is bounded from below.

If f is convex or concave, the problem has already been studied in [6] with a general constraint having the form (7). It is proved there that if $\underline{u} = \overline{u}$ on $\partial\Omega$, the minimizer is either $\underline{u}$ or $\overline{u}$.

It was already explained in [6] that this sort of problem cannot be studied through the usual methods of the calculus of variations, due to the convexity constraint on u. In particular there is no 'regularizing effect' even if f is convex. Also the Euler equation does not apply (see below).

In this paper we are interested in the case where f is not convex but may not be concave either: the Hessian matrix d^2f has at least one negative eigenvalue. If this is true for every p in some subset Q of $\mathbb{R}^2$ then we call f *nowhere convex* in Q. As remarked before, the function $p \mapsto 1/(1+|p|^2)$ considered by Newton is nowhere convex in $\mathbb{R}^2$.

1.3 Statement of main results

Let us recall that a convex function u is almost everywhere differentiable in its domain; moreover, it is actually C^1 on the set $D(\nabla u)$ where it is differentiable (see for instance [9], Theorem 25.5).

If $A \subset \Omega$, we will note $\nabla u(A)$ the image set of A under the map $x \mapsto \nabla u(x)$ (for those $x \in A$ where it is defined).

Our main theorem is as follows:

Theorem 1 *Let u be a minimizer of Problem (4). Suppose that u is differentiable on an open subset $\Omega_1 \subset \Omega$ and that f is nowhere convex on $\nabla u(\Omega_1)$. Then u is not strictly convex in Ω_1.*

Note that by repeated application this theorem implies that the region of non-strict-convexity is dense in Ω_1.

The proof of this statement will follow in Section 5. Let us first give some important consequences of this theorem.

First of all let us observe that the radial minimizer calculated by Newton (which is differentiable and strictly convex on some regions inside the ball) can not be a minimizer of (4), and consequently the minimizer of (4) is not radially symmetric. As mentioned before, this was already proved in [3] by a very different argument.

Second, here is an alternative statement of Theorem 1:

Corollary 1 *Under the conditions of the theorem, if $\phi \in C^1(\Omega)$ is strictly convex in some open subset of $D(\nabla u)$, then $u - \phi$ is not convex.*

(This follows from the theorem by observing that if ϕ is strictly convex in ω open, and $u - \phi$ is convex, then u is strictly convex in ω.) Hence, u is an 'almost extremal' point of $\mathcal{C}$, considered as a convex subset of the set of functions $H^1_{\mathrm{loc}}(\Omega)$.

Finally, the proof of Theorem 1 is based on the fact that a minimizer that is of class C^1 and strictly convex does not satisfy the Euler equation associated with (4),

$$\operatorname{div} df(\nabla u) = 0. \tag{6}$$

In fact, on carefully reading the proof of Lemma 3 the reader can verify that the result remains true without the condition of regularity, under some assumptions on the local variation of ∇u. We conjecture that this is in fact

the general case: no minimizer satisfies (6), even in the sense of distributions and restricted to some subdomain.

Theorem 1 can also be generalized in some ways. For instance, if u is a minimizer of (4), then it is easy to prove that the sets $N_0 := \{x \in \overline{\Omega}; u(x) = 0\}$ and $\{x \in \overline{\Omega}; u(x) = 1\}$ are non-empty; the last one is included in the boundary of Ω, due to convexity. In the following, we do not examine the behaviour of u near N_0, nor the shape of this set, since u is not strictly convex in its interior. Therefore one can easily generalize the constraint '$u(x) \in [0,1]$' to

$$\underline{u}(x) \leq u(x) \leq \overline{u}(x), \tag{7}$$

where $\underline{u}$, $\overline{u}$ are any functions (provided there exists at least one convex function satisfying (7)). In that case, N_0 is simply the set where equality is attained in one side of (7), and the theorem still holds if $\underline{u}$, $\overline{u}$ are affine functions, or if $\Omega_1 \cap N_0 = \emptyset$ for instance.

More generally, the reader should note that in all the following the condition $u \in [0,1]$ is not taken into account. This is permitted by the fact that we consider some restrictions of u to smaller sets ω compactly inbedded in Ω_1, and therefore we have $u \in [\varepsilon, 1-\varepsilon]$ for some $\varepsilon > 0$. We do not recall this in each instance.

2 On convex functions

Let us recall some well-known properties of convex functions on $\mathbb{R}^2$, and give some definitions.

We recall that if u and v are convex, then so is $\max(u,v)$; hence this is an internal operation in $\mathcal{C}$. We will note aff $[p,s]$ the affine function $x \mapsto p \cdot x - s$, where $p \in \mathbb{R}^2$ and $s \in \mathbb{R}$ are given and $\cdot$ is the conventional inner product in $\mathbb{R}^2$.

Let $u \in \mathcal{C}$ be given, and $x_0 \in \Omega$ be any point where u is differentiable. Then there exists a unique affine function $\pi[u; x_0]$ satisfying

$$\pi[u;x_0](x_0) = u(x_0); \qquad \pi[u;x_0](x) \leq u(x), \quad \forall x;$$

and $\nabla \pi[u; x_0] \equiv \nabla u(x_0)$.

Conversely, if $p \in \mathbb{R}^2$ is given, there exists a unique number $u^*(p)$ such that $u \geq$ aff $[p, u^*(p)]$, with equality at least in one point. Indeed we have

$$u^*(p) = \sup_{x \in \Omega} [p \cdot x - u(x)], \tag{8}$$

which is sometimes called the *conjugate* of u. We also have:

$$u(x) = \max_{p \in \nabla u(\Omega)} \text{aff}\,[p, u^*(p)](x) \tag{9}$$

since $D(\nabla u)$ is dense in Ω.

An *exposed point* of a convex body C is a point $X \in \partial C$ such that every supporting hyperplane through X only intersects C in X (see [9], p. 162). An exposed point has the property that a small parallel displacement of the hyperplane leads to a small intersection with C. In the context of convex functions $u : \Omega \to \mathbb{R}$ we shall call $x \in \Omega$ an exposed point if $X = (x, u(x))$ is an exposed point of the graph of u. Note that for any convex open subset $\omega \subset \Omega$,

$$\text{every } x \in \omega \text{ is exposed} \iff u \text{ is strictly convex in } \omega. \tag{10}$$

Let ω be a subdomain of Ω. We shall say that a function $u \in \mathcal{C}$ is *developable* on ω if at any point $x \in \omega$ where u is differentiable the tangent plane intersects the boundary set $\{(x, u(x)) : x \in \partial\omega\}$. This is equivalent to the assertion that the Gaussian curvature, where defined, is zero; other ways of characterizing are

- the surface defined by the graph of u is ruled, i.e. the union of line segments;
- the surface can be 'folded' *(developed)* from a flat surface.

If u is a given convex function defined on $\overline{\omega}$, there exists a unique maximal developable convex function $\hat{u}_\omega$ defined on $\overline{\omega}$, and satisfying $\hat{u}_\omega \leq u$ in $\overline{\omega}$. If u is differentiable in a neighbourhood of $\partial\omega$ then we can define it by setting

$$\hat{u}_\omega(x) = \max_{p \in \nabla u(\partial\omega)} \text{aff}\,[p, u^*(p)](x), \qquad \forall x \in \overline{\omega}, \tag{11}$$

which will be sufficient for our applications (the more general case could be obtained by an approximation argument, for instance).

We will also use some information about generalized second-order derivatives of convex functions. In general the second derivative of a convex function on Ω is a Radon measure ([1], Th. 6.3.2), i.e. a distribution μ such that

$$|< \mu, \zeta >| \leq C \left\| \zeta \right\|_{L^\infty(\Omega)} \qquad \text{for all } \zeta \in C_0^0(\Omega).$$

For differentiable convex functions it is possible to say more.

Lemma 2 *Let $v : \overline{\Omega} \subset \mathbb{R}^2 \to \mathbb{R}$ be convex and of class C^1 on $\overline{\Omega}$. Then $d^2v \in L^1(\Omega; \mathbb{R}^{2\times 2})$ and*

$$\left\|d^2 v\right\|_{L^1(\Omega)} \leq 4\,\mathrm{diam}(\Omega) \operatorname*{osc}_{\partial\Omega} \nabla v.$$

(Here $\operatorname*{osc}_{\partial\Omega} \nabla v := \sup_{x_1, x_2 \in \partial\Omega} |\nabla v(x_1) - \nabla v(x_2)|$.)

Proof:

First suppose that v is smooth. For any $\zeta \in C(\overline{\Omega})$, $\zeta \geq 0$,

$$0 \leq \int_\Omega \zeta \partial_{xx} v \leq \|\zeta\|_{L^\infty(\Omega)} \int_\Omega \partial_{xx} v \leq \|\zeta\|_{L^\infty(\Omega)}\,\mathrm{diam}(\Omega)\ \mathrm{osc}\,\partial_x v, \tag{12}$$

and a similar estimate holds for $\partial_{yy} v$. By the convexity condition $\partial_{xx} v \partial_{yy} v - (\partial_{xy} v)^2 \geq 0$ we have

$$|\partial_{xy} v| \leq \sqrt{\partial_{xx} v \partial_{yy} v}$$

and therefore

$$\begin{aligned} \int_\Omega \zeta\,|\partial_{xy} v| &\leq \left(\int_\Omega \zeta \partial_{xx} v\right)^{\frac{1}{2}} \left(\int_\Omega \zeta \partial_{yy} v\right)^{\frac{1}{2}} \\ &\leq \|\zeta\|_{L^\infty(\Omega)}\,\mathrm{diam}\,\Omega\ \mathrm{osc}\,\nabla v. \end{aligned} \tag{13}$$

Inequalities (12) and (13) prove the Lemma for smooth functions.

For a general convex C^1 function v we choose a sequence of convex functions $v_n \in C^2(\overline{\Omega})$ that converges to v in $C^1(\overline{\Omega})$. By applying the previous argument to this sequence we find that $d^2 v_n$ is bounded in L^1 and we can therefore assume that $v_n \rightharpoonup v$ in $W^{2,1}(\Omega)$. Consequently

$$\left\|d^2 v\right\|_{L^1(\Omega)} \leq \liminf_{n\to\infty} \left\|d^2 v_n\right\|_{L^1(\Omega)} \leq 4\,\mathrm{diam}(\Omega) \liminf_{n\to\infty} \operatorname*{osc}_{\partial\Omega} \nabla v_n,$$

and we conclude by the convergence of v_n in $C^1(\overline{\Omega})$. □

3 Another minimization problem

In this section we consider a different minimization problem, which we believe has an interest of its own. The results here will be used as auxiliary lemmas in the proof of the main theorem.

It is well known that if $\omega \subset \mathbb{R}^N$ is an open set, and M a given symmetric matrix, then the infimum

$$\inf_{\phi \in H_0^1(\omega)} \frac{\int_\omega M\nabla\phi \cdot \nabla\phi}{\int_\omega |\nabla\phi|^2}$$

is not attained (except if $M \in \mathbb{R}\,\mathrm{Id}$) and equals the first eigenvalue $\lambda_1(M)$ of M. This can be proved by considering the sequence $\phi_n(x) := \eta(x) \times \sin(nx\xi_0)$ where $\eta \in C_0^1(\omega)$ is fixed and $\xi_0 \in \mathbb{R}^N \setminus \{0\}$ satisfies $M\xi_0 = \lambda_1(M)\xi_0$; it is now easy to check out that $\int_\omega M\nabla\phi_n \cdot \nabla\phi_n \Big/ \int_\omega |\nabla\phi_n|^2$ converges to $\lambda_1(M)$ as $n \to \infty$.

We are interested here in the similar minimization problem, under the additional constraint that $\phi \in H_0^1(\omega)$ be *convex*. Since the set of convex functions is much smaller than the whole of H_0^1, and since the sequence ϕ_n obviously does not satisfy this condition, it is quite surprising that we can prove a similar result and obtain the *same* infimum. However, it is necessary here to assume the set ω to be convex, otherwise there is no non-zero convex functions in $H_0^1(\omega)$. We also have to assume that ω has a C^1 boundary: it is not clear whether the same result holds with a general convex set ω; we strongly suspect that *it does not.*

Theorem 2 *Let $\omega \subset \mathbb{R}^2$ be convex with boundary of class C^1, and let $M \in \mathbb{R}^{2\times 2}$ be a given symmetric matrix. Then*

$$\inf_{\substack{\phi \in H_0^1(\omega) \\ \phi \text{ convex}}} \frac{\int_\omega M\nabla\phi \cdot \nabla\phi}{\int_\omega |\nabla\phi|^2} = \lambda_1(M)$$

where $\lambda_1(M)$ is the smallest eigenvalue of M.

Proof:

Choosing an appropriate set of coordinates $(x, y) \in \mathbb{R}^2$, we can assume that M is the diagonal matrix $\mathrm{diag}\,(\lambda_1, \lambda_2)$; moreover

$$M = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \lambda_1\, I + (\lambda_2 - \lambda_1) \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

implies

$$\frac{\int M\nabla\phi \cdot \nabla\phi}{\int |\nabla\phi|^2} = \lambda_1 + (\lambda_2 - \lambda_1)\frac{\int (\partial_y\phi)^2}{\int |\nabla\phi|^2}.$$

Hence it is sufficient to prove that

$$\inf_{\substack{\phi\in H^1_0(\omega)\\ \phi \text{ convex}}} \frac{\int(\partial_y\phi)^2}{\int|\nabla\phi|^2} = 0$$

Since $\partial\omega$ is of class C^1, it can be parameterized in the form $s \in [-a,a] \mapsto (x(s),y(s))$, where $a>0$ is half the length of $\partial\omega$, and x, y are C^1 functions whose derivatives $\dot{x},\dot{y}$ satisfy $\dot{x}^2+\dot{y}^2 \equiv 1$.

Let $\varepsilon > 0$ be a given number, $(x_\varepsilon, y_\varepsilon)$ a point (to be fixed in a while) in ω. We consider the largest convex function ϕ_ε defined in $\overline{\omega}$ and satisfying $\phi_\varepsilon(x_\varepsilon,y_\varepsilon) = -1$ and $\phi_\varepsilon = 0$ on $\partial\omega$. Its epigraph is a (generalized) cone in $\mathbb{R}^3$ with vertex $(x_\varepsilon, y_\varepsilon, -1)$. This implies that

$$\forall t\in[0,1],\ \forall s\in[-a,a], \qquad \phi_\varepsilon(tx(s)+(1-t)x_\varepsilon, ty(s)+(1-t)y_\varepsilon) = t-1.$$

Hence we get by differentiation

$$\begin{cases} (x(s)-x_\varepsilon)\partial_x\phi_\varepsilon + (y(s)-y_\varepsilon)\partial_y\phi_\varepsilon = 1 \\ \dot{x}(s)\partial_x\phi_\varepsilon + \dot{y}(s)\partial_y\phi_\varepsilon = 0 \end{cases}$$

That gives

$$\partial_x\phi_\varepsilon = \frac{\dot{y}}{w_\varepsilon(s)}, \qquad \partial_y\phi_\varepsilon = -\frac{\dot{x}}{w_\varepsilon(s)}$$

where $w_\varepsilon(s) := \dot{y}(x-x_\varepsilon) - \dot{x}(y-y_\varepsilon)$. We have $w_\varepsilon(s) \neq 0$ for all s since $|w_\varepsilon(s)|$ is the distance from the interior point $(x_\varepsilon,y_\varepsilon)\in\omega$ to the tangent of $\partial\omega$ at $(x(s),y(s))$. Since the Jacobian determinant in the change of variable

$$(s,t) \mapsto (tx(s)+(1-t)x_\varepsilon, ty(s)+(1-t)y_\varepsilon)$$

is $t\,|w_\varepsilon(s)|$ we get

$$\int_\omega (\partial_y\phi_\varepsilon)^2 = \int_{-a}^{a}\int_0^1 \frac{\dot{x}^2}{w_\varepsilon^2(s)}\, t\,|w_\varepsilon(s)|\, dt\, ds = \frac{1}{2}\int_{-a}^{a}\frac{\dot{x}^2}{|w_\varepsilon(s)|}\, ds,$$

and similarly

$$\int_\omega |\nabla\phi_\varepsilon|^2 = \frac{1}{2}\int_{-a}^{a}\frac{1}{|w_\varepsilon(s)|}\, ds.$$

Since $\partial\omega$ is C^1, there exists at least one point and at most a segment where the unit normal is directed in the positive y-direction. More precisely,

we can find $s_1 \leq s_2$ such that $\dot{x}(s) = 0$ and $\dot{y}(s) = 1$ for all $s \in [s_1, s_2]$. (If ω is strictly convex, then $s_1 = s_2$.) We can assume $0 \in [s_1, s_2]$, and we define $x_0 := x(0)$, $y_0 := y(0)$.

We now choose the point $(x_\varepsilon, y_\varepsilon)$ on the normal to (x_0, y_0), at distance ε, that is we set

$$x_\varepsilon := x_0 - \varepsilon, \qquad y_\varepsilon = y_0$$

We also have $w_\varepsilon(s) = \varepsilon \dot{y}(s) + w_0(s)$ with $w_0(s) := \dot{y}x - \dot{x}y$ does not depend on ε. Note that both functions $\dot{y}$ and w_0 are continuous. Since $|w_0(s)|$ is the distance from (x_0, y_0) to the tangent of $\partial\omega$ at $(x(s), y(s))$, this function vanishes exactly on the interval $[s_1, s_2]$ and is at most of order s near $s = 0$. Moreover $\dot{y}(0) = 1$; so we have

$$\lim_{\varepsilon \to 0} \int_{-a}^{a} \frac{ds}{|w_\varepsilon(s)|} = \int_{-a}^{a} \frac{ds}{|w_0(s)|} = +\infty \tag{14}$$

This implies that for all $\varepsilon > 0$, we can find $\delta(\varepsilon) > 0$ such that $\lim_{\varepsilon \to 0} \delta(\varepsilon) = 0$ and

$$\int_{-\delta(\varepsilon)+s_1}^{\delta(\varepsilon)+s_2} \frac{ds}{|w_\varepsilon(s)|} \Bigg/ \int_{-a}^{a} \frac{ds}{|w_\varepsilon(s)|} \longrightarrow 1 \tag{15}$$

as $\varepsilon \to 0$.

Denoting these integrals by I_δ and I for short, we now observe that

$$\begin{aligned} \int_{-a}^{a} \frac{\dot{x}^2}{|w_\varepsilon|} &= \int_{-\delta+s_1}^{\delta+s_2} \frac{\dot{x}^2}{|w_\varepsilon|} + \int_{[-a,-\delta+s_1]\cup[\delta+s_2,+a]} \frac{\dot{x}^2}{|w_\varepsilon|} && (16) \\ &\leq I_\delta \sup_{-\delta+s_1 < s < \delta+s_2} \dot{x}^2(s) + (I - I_\delta) \left\| \dot{x}^2 \right\|_{L^\infty(-1,1)} && (17) \end{aligned}$$

which implies

$$I^{-1} \int_{-a}^{a} \frac{\dot{x}^2}{|w_\varepsilon|} \longrightarrow 0 \tag{18}$$

as $\varepsilon \to 0$ using (15) and the fact that the continuous function $\dot{x}^2(s)$ vanishes on $[s_1, s_2]$.

Writing again (18) with the original function ϕ_ε, we get

$$\int_\omega (\partial_y \phi_\varepsilon)^2 \Bigg/ \int_\omega |\nabla \phi_\varepsilon|^2 \longrightarrow 0$$

and therefore (3) is proved. □

Corollary 2 *Let $\omega \subset \mathbb{R}^2$ be convex with boundary of class C^1, and $M : \overline{\omega} \to \mathbb{R}^{2\times 2}$ be a continuous map of symmetric matrices, satisfying*

$$\exists \lambda \in \mathbb{R},\ \exists \xi_0 \in S^1,\ \forall x \in \overline{\omega}, \qquad M(x)\xi_0 \cdot \xi_0 < \lambda$$

(where S^1 is the unit ball of $\mathbb{R}^2$). Then we have

$$\inf_{\substack{\phi \in H^1_0(\omega) \\ \phi \text{ convex}}} \frac{\int_\omega M(x)\nabla\phi \cdot \nabla\phi}{\int_\omega |\nabla\phi|^2} \leq \lambda$$

The corollary can be proved in the same way than the theorem. It is just necessary to consider the point $(x_0, y_0) \in \partial\omega$ where the normal vector is exactly equal to ξ_0. We omit the details.

4 An instability lemma

The following result will be crucial in the proof of Theorem 1. It conveys the property that a C^1 minimizer can not be a stationary point and strictly convex at the same time.

Lemma 3 *Let u minimize F and suppose that u is strictly convex and of class C^1 in $\Omega_1 \subset \Omega$. Then there exists an affine function $\theta : \mathbb{R}^2 \to \mathbb{R}$ such that*

$$\omega = \{x \in \Omega : u(x) < \theta(x)\} \tag{19}$$

is non-empty and included in Ω_1, and such that

$$\int_\omega df(\nabla u)\nabla(u - \theta) < 0.$$

Proof:

Define Θ as the set of all affine functions θ such that the set ω defined by (19) is non-empty and included in Ω_1; for a given θ we denote the corresponding set ω_θ. Note that Θ is not empty since Ω_1 consists of exposed points by (10). Several times we shall use the Taylor expansion

$$\begin{aligned} f(p) &= f(p_0) + df(p_0)(p - p_0) + \frac{1}{2}d^2 f(p_0)(p - p_0)\cdot(p - p_0) + \\ &\quad + g(p_0, p)(p - p_0)\cdot(p - p_0)\cdot(p - p_0), \end{aligned} \tag{20}$$

where g is bounded on compact subsets of $\mathbb{R}^2 \times \mathbb{R}^2$.

The proof proceeds in three steps.

Step 1. For every $\theta \in \Theta$, $\int_{\omega_\theta} df(\nabla u)\nabla(u-\theta) \le 0$.

Suppose that $\int_{\omega_\theta} df(\nabla u)\nabla(u-\theta) > 0$ for some θ. Define on ω_θ the functions $\phi = u - \theta$ and $\phi_\varepsilon = (1-\varepsilon)\phi$. The function

$$u_\varepsilon = \begin{cases} u & \text{on } \Omega \setminus \omega_\theta \\ \theta + \phi_\varepsilon & \text{on } \omega_\theta \end{cases}$$

is admissible if $0 \le \varepsilon \le 1$, and since on ω_θ we have $u_\varepsilon = u - \varepsilon\phi$, and using (20),

$$\begin{aligned} F(u_\varepsilon) - F(u) &= -\varepsilon \int_{\omega_\theta} df(\nabla u)\nabla\phi + \varepsilon^2 \int_{\omega_\theta} d^2 f(\nabla u)\nabla\phi\cdot\nabla\phi \\ &\quad - \varepsilon^3 \int_{\omega_\theta} g(\nabla u, \nabla u_\varepsilon)\nabla\phi\cdot\nabla\phi\cdot\nabla\phi. \end{aligned}$$

Clearly this contradicts the minimality of u when ε is small.

Step 2. We claim that if

$$\int_{\omega_\theta} df(\nabla u)\nabla(u-\theta) = 0, \qquad \forall \theta \in \Theta, \tag{21}$$

then $\operatorname{div} df(\nabla u) = 0$ a.e. in Ω_1. (Note that since u is of class C^1 in Ω_1, $\operatorname{div} df(\nabla u) \in L^1(\Omega_1)$ by Lemma 2.)

Indeed, pick a $\theta \in \Theta$; for small enough $\varepsilon > 0$ the function $\theta_\varepsilon(x) = \theta(x) - \varepsilon$ also belongs to Θ. Set $\phi = u - \theta$ and $\phi_\varepsilon = u - \theta_\varepsilon$. Then from (21),

$$\begin{aligned} 0 &= \int_{\omega_\theta} df(\nabla u)\nabla\phi - \int_{\omega_{\theta_\varepsilon}} df(\nabla u)\nabla\phi_\varepsilon \\ &= -\int_{\omega_\theta} \phi \operatorname{div} df(\nabla u) + \int_{\omega_{\theta_\varepsilon}} \phi_\varepsilon \operatorname{div} df(\nabla u) \\ &= \int_{\omega_{\theta_\varepsilon}} (\phi_\varepsilon - \phi)\operatorname{div} df(\nabla u) - \int_{\omega_\theta \setminus \omega_{\theta_\varepsilon}} \phi \operatorname{div} df(\nabla u) \\ &= -\varepsilon \int_{\omega_{\theta_\varepsilon}} \operatorname{div} df(\nabla u) + o(\varepsilon) \end{aligned}$$

as ε tends to zero, since $\|\phi\|_{L^\infty} < \varepsilon$ on $\omega_\theta \setminus \omega_{\theta_\varepsilon}$. Consequently we get $\int_{\omega_\theta} \operatorname{div} df(\nabla u) = 0$ for all $\theta \in \Theta$.

As we remarked before, by strict convexity all points in Ω_1 are exposed points. This implies that we can write any open set $A \subset \Omega_1$ as a disjoint union $A = E \cup \bigcup_{i \in I} \omega_{\theta_i}$ where E is a set of measure zero. Therefore we have $\int_A \operatorname{div} df(\nabla u) = 0$ for all open sets $A \subset \Omega_1$, which implies $\operatorname{div} df(\nabla u) = 0$ a.e. in Ω_1.

Step 3. $\operatorname{div} df(\nabla u) = 0$ a.e. in Ω_1 contradicts the minimality of u.

Note that this implies that the hypothesis '$\int_{\omega_\theta} df(\nabla u)\nabla(u - \theta) = 0$ for all $\theta \in \Theta$' leads to a contradiction and that therefore the lemma is proved.

Pick $\theta \in \Theta$ such that ω_θ is small enough for $M(x) = d^2 f(\nabla u(x))$ to satisfy the hypotheses of Corollary 2. Note that the C^1 regularity of $\partial\omega_\theta$ follows from the regularity of u and the implicit function theorem. We can then find a convex $h_0 \in C_0^1(\omega_\theta)$ such that

$$\int_{\omega_\theta} d^2 f(\nabla u)\nabla h_0 \cdot \nabla h_0 < 0. \tag{22}$$

Define h_ε for $\varepsilon > 0$ by

$$u + h_\varepsilon = (u + \varepsilon h_0) \vee \hat{u}_{\omega_\theta} \qquad \text{on } \omega_\theta.$$

Then the function

$$u_\varepsilon = \begin{cases} u & \text{on } \Omega \setminus \omega_\theta \\ u + h_\varepsilon & \text{on } \omega_\theta \end{cases}$$

is convex on Ω and therefore admissible. Since $\operatorname{div} df(\nabla u) = 0$ we have

$$\begin{aligned} F(u_\varepsilon) - F(u) &= \frac{1}{2}\int_{\omega_\theta} d^2 f(\nabla u)\nabla h_\varepsilon \cdot \nabla h_\varepsilon \\ &\quad + \int_{\omega_\theta} g(\nabla u, \nabla u_\varepsilon)\nabla h_\varepsilon \cdot \nabla h_\varepsilon \cdot \nabla h_\varepsilon. \end{aligned} \tag{23}$$

Since ϕ is strictly convex, the measure of the set $\{x \in \omega_\theta : h_\varepsilon(x) > \varepsilon h_0(x)\}$ vanishes as $\varepsilon \to 0$ and consequently $\varepsilon^{-1}\nabla h_\varepsilon \to \nabla h_0$ in $L^p(\omega_\theta; \mathbb{R}^2)$ for all $1 \le p < \infty$. Therefore if ε is small enough,

$$\int_{\omega_\theta} d^2 f(\nabla u)\nabla h_\varepsilon \cdot \nabla h_\varepsilon < \frac{\varepsilon^2}{2}\int_{\omega_\theta} d^2 f(\nabla u)\nabla h_0 \cdot \nabla h_0,$$

and on the other hand,

$$\left| \int_{\omega_\theta} g(\nabla u, \nabla u_\varepsilon) \nabla h_\varepsilon \cdot \nabla h_\varepsilon \cdot \nabla h_\varepsilon \right| \leq \|g\|_{L^\infty} \varepsilon^3 \, \|\nabla h_0\|^3_{L^3(\omega_\theta)} .$$

Inserting these inequalities in (23), and using (22), yields a contradiction with the minimality of u. □

5 Proof of the main theorem

We now give the proof of Theorem 1.

Suppose that u is strictly convex and of class C^1 in Ω_1, and let θ be the affine function given by Lemma 3. We write simply ω for the corresponding set ω_θ. Let us define $\phi := u - \theta$ in $\overline{\omega}$, and deduce a convex function $\hat{\phi}_\omega$ defined in $\overline{\omega}$ through the rule given in Section 2 with a relation similar to (11).

If ζ is any function in the class

$$\mathcal{C}_0 = \{\zeta \in C^0(\overline{\omega}) : \zeta \text{ is convex and } \hat{\phi}_\omega \leq \zeta \leq 0\},$$

then the function

$$v := \begin{cases} \zeta + \theta & \text{in } \overline{\omega} \\ u & \text{in } \Omega \setminus \omega \end{cases}$$

is convex and admissible. We have $v = u$ if and only if $\zeta = \phi$.

Let us note $p_0 := \nabla\theta$, and rewrite f near p_0 with a Taylor expansion (20); we will also write

$$M(\nabla\zeta) := \frac{1}{2} d^2 f(p_0) + g(p_0, p_0 + \nabla\zeta)\nabla\zeta \quad \text{and} \quad \widetilde{f}(\nabla\zeta) := M(\nabla\zeta)\nabla\zeta \cdot \nabla\zeta$$

for short. Using the fact that $\int_\omega \nabla(v - \theta) = 0$ we can write

$$\begin{aligned} F(v) - F(\theta) &= \int_\omega f(\nabla v) - f(\nabla\theta) \\ &= \int_\omega M(\nabla\zeta)\nabla\zeta \cdot \nabla\zeta \\ &= \int_\omega \widetilde{f}(\nabla\zeta). \end{aligned}$$

Therefore the function ϕ minimizes

$$\widetilde{F}(\zeta) := \int_\omega \widetilde{f}(\nabla\zeta)$$

within the class $\mathcal{C}_0$. Observe that $\widetilde{f}$ has properties similar to f, except that $\widetilde{f}(0) = 0$, $d\widetilde{f}(0) = 0$; in particular, $\widetilde{f}$ is nowhere convex on the set $\nabla\phi(\omega)$.

Define for $0 < \varepsilon \leq 1$ the functions

$$\phi_\varepsilon = (1+\varepsilon)\phi \qquad \text{and} \qquad \psi_\varepsilon = \phi_\varepsilon \vee \hat{\phi}_\omega$$

and set $h_\varepsilon = \psi_\varepsilon - \phi_\varepsilon$ and $\omega_\varepsilon = \{x \in \omega : h_\varepsilon(x) > 0\}$. Note that $\psi_\varepsilon \in \mathcal{C}$. We make two important remarks here: first,

$$\begin{aligned} \|h_\varepsilon\|_{L^\infty(\omega_\varepsilon)} &= \|\psi_\varepsilon - \phi_\varepsilon\|_{L^\infty(\omega_\varepsilon)} = \|\hat{\phi}_\omega - \phi_\varepsilon\|_{L^\infty(\omega_\varepsilon)} \\ &\leq \varepsilon \|\phi\|_{L^\infty(\omega_\varepsilon)}. \end{aligned} \tag{24}$$

Second, the Lebesgue measure of ω_ε tends to zero as $\varepsilon \to 0$: since by the strict convexity $\phi > \hat{\phi}_\omega$ on ω, for every $x \in \omega$ we have $h_\varepsilon(x) = 0$ if ε is small enough.

We have

$$\begin{aligned} \widetilde{F}(\phi_\varepsilon) - \widetilde{F}(\phi) &= \int_\omega \widetilde{f}\Big((1+\varepsilon)\nabla\phi\Big) - \widetilde{f}(\nabla\phi) \\ &= \varepsilon \int_\omega d\widetilde{f}(\nabla\phi)\nabla\phi + O(\varepsilon^2). \end{aligned} \tag{25}$$

We claim that we also have, as $\varepsilon \to 0$,

$$\widetilde{F}(\psi_\varepsilon) - \widetilde{F}(\phi_\varepsilon) \in o(\varepsilon). \tag{26}$$

Indeed, using a simple Taylor expansion near ϕ_ε, we have

$$\left| \widetilde{F}(\psi_\varepsilon) - \widetilde{F}(\phi_\varepsilon) - \int_{\omega_\varepsilon} df(\nabla\phi_\varepsilon)\nabla h_\varepsilon \right| \leq c \int_{\omega_\varepsilon} |\nabla h_\varepsilon|^2$$

(here and in the following, c denotes various constants related to the derivatives of f on ω).

We first note that in ω_ε, if ε is small enough,

$$\left| \operatorname{div} df(\nabla\phi_\varepsilon) \right| < c \left| d^2\phi \right|$$

and, since $h_\varepsilon = \hat{\phi}_\omega - \phi_\varepsilon$ on ω_ε,

$$-\Delta h_\varepsilon = \Delta\phi_\varepsilon - \Delta\hat{\phi}_\omega \leq c\left|d^2\phi\right|.$$

Here we use the convexity of $\hat{\phi}_\omega$ to conclude that $\Delta\hat{\phi}_\omega \geq 0$. Note that by Lemma 2 the term $|d^2\phi|$ belongs to $L^1(\omega)$.

Hence the vanishing of ω_ε together with (24) implies that, as $\varepsilon \to 0$,

$$\begin{aligned}\int_{\omega_\varepsilon} |\nabla h_\varepsilon|^2 &= -\int_{\omega_\varepsilon} h_\varepsilon \Delta h_\varepsilon \\ &\leq c\varepsilon \left\|\phi\right\|_{L^\infty(\omega_\varepsilon)} \left\|d^2\phi\right\|_{L^1(\omega_\varepsilon;\mathbb{R}^{2\times 2})} \\ &\in o(\varepsilon)\end{aligned}$$

and similarly

$$\begin{aligned}\left|\int_{\omega_\varepsilon} df(\nabla\phi_\varepsilon)\nabla h_\varepsilon\right| &= \left|\int_{\omega_\varepsilon} h_\varepsilon \operatorname{div} df(\nabla\phi_\varepsilon)\right| \\ &\leq c\varepsilon \left\|\phi\right\|_{L^\infty(\omega_\varepsilon)} \left\|d^2\phi\right\|_{L^1(\omega_\varepsilon;\mathbb{R}^{2\times 2})} \\ &\in o(\varepsilon).\end{aligned}$$

All these inequalities together imply (26).

Since we chose θ as the function given by Lemma 3 we have

$$\int_\omega d\widetilde{f}(\nabla\phi)\nabla\phi < 0.$$

Combining this inequality with (25) and (26) we find that by choosing ε small enough we obtain

$$\widetilde{F}(\psi_\varepsilon) - \widetilde{F}(\phi) \leq \frac{\varepsilon}{2}\int_\omega d\widetilde{f}(\nabla\phi)\nabla\phi < 0,$$

which contradicts the assumption that ϕ is minimal. This concludes the proof of the theorem. □

References

[1] L. C. Evans & R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Applied Mathematics, CRC Press (1992).

[2] Y. Brenier, *Polar Factorization and Monotone Rearrangement of Vector-valued Functions*, Comm. Pure Appl. Math. vol. XLIV (1991), pp. 375–417.

[3] F. Brock, V. Ferone, B. Kawohl, *A Symmetry Problem in the Calculus of Variations*, Calc. Var. Partial Differential Equations, **4** (1996), pp. 593–599.

[4] G. Buttazzo, V. Ferone, B. Kawohl, *Minimum Problems over Sets of Concave Functions and Related Questions*, Math. Nachrichten, **173** (1993), pp. 71–89.

[5] H. H. Goldstine, *A History of the Calculus of Variations from the 17th through the 19th Century*, Springer-Verlag, Heidelberg (1980).

[6] T. Lachand-Robert, M. A. Peletier, *Extremal Points of a Functional on the Set of Convex Functions*, to appear.

[7] I. Newton, *Philosophiae Naturalis Principia Mathematica* (1686).

[8] A. V. Pogorelov, *Extrinsic Geometry of Convex Surfaces*, Transl. Math. Monographs, Am. Math. Soc., Providence, Rhode Island (1973).

[9] R. T. Rockafellar, *Convex Analysis*, Princeton University Press (1970).

[10] T. Zamfirescu, *Baire Category in Convexity*, Atti Sem. Mat. Fis. Univ. Modena **39** (1991), pp. 139–164. *See also: Nearly All Convex Bodies are Smooth and Strictly Convex*, Monatsh. Math., **103** (1987), pp. 57–62.

Laboratoire d'Analyse Numérique, *Tour 55-65*, Université Paris–VI,
4 place Jussieu, 75252 Paris Cedex 05.

University of Bath, Department of Mathematics, Claverton Down,
Bath BA2 7AY (UK)