Noname manuscript No.
(will be inserted by the editor)

Local-Global Minimum Property in Unconstrained Minimization Problems

Pál Burai

Communicated by Paul I. Barton

the date of receipt and acceptance should be inserted later

Abstract The main goal of this paper is to prove some new results and extend some earlier ones about functions, which possess the so called local-global minimum property. In the last section, we show an application of these in the theory of calculus of variations.

Keywords Nonlinear optimization. Nonconvex optimization. First order sufficient condition. Generalized convexity. Local-global minimum property.

Mathematics Subject Classification (2000) 90C26, 49K

1 Introduction

With regard to a general nonlinear optimization problem, it is typical that the first order condition is only necessary. It becomes sufficient, if the problem is convex in the usual sense or in a generalized sense. When having a non-convex case a second order condition is needed, but even this ensures (in general) only the sufficiency of being a local optimum instead of being a global one. Therefore, it is useful to investigate such classes of functions, which possess the so-called local-global minimum property. Or,

TU Berlin, Department of Mathematics, Strae des 17. Juni 136, 10623 Berlin
E-mail: burai@math.tu-berlin.de

in other words, such functions should be characterized whose every local minimizer is also global. Of course, this type of analysis is not new. A couple of authors have already dealt with it (see, e.g., [1–7]).

Here, we prove some new results in this field and also extend some earlier ones.

In the second section, we deal with functions, which have lower semicontinuous lower-level-set function. This leads to a possible characterization of local-global minimum property. This fact was first observed by Avriel and Zang (see [6]). In the third part, we study connected and quasi-connected functions, which are natural generalizations of convexity and quasi-convexity. These concepts were introduced by Ortega and Rheinboldt in [5].

In the fourth part, we discuss the directional derivative with respect to a path and prove a variational inequality. Finally, with the help of an application from the theory of calculus of variations, we verify how useful these concepts are.

2 Functions with Lower Semicontinuous Lower-Level-Set Function

Here, we use the notations of the book [8]. Let X and Y be metric spaces. A set valued map $F\colon X \rightrightarrows Y$ is *lower semicontinuous at* $x \in \operatorname{Dom} F$ iff for any $y \in F(x)$ and for any sequence of elements $x_n \in \operatorname{Dom} F$ converging to x, there exists a sequence of elements $y_n \in F(x_n)$ converging to y, where $\operatorname{Dom} F := \{x \in X | F(x) \neq \emptyset\}$.

Lemma 2.1 *Let $f\colon X \to \mathbb{R}$ be a function, and let $\bar{\alpha} \in \operatorname{Dom} L_f$ such that $\bar{\alpha} > \inf_X f$. Then L_f is lower semicontinuous at $\bar{\alpha}$ if and only if there exist a strictly monotone increasing sequence α_n tending to $\bar{\alpha}$ and a sequence x_n tending to $\bar{x}$, where $f(\bar{x}) = \bar{\alpha}$ and $x_n \in L_f(\alpha_n)$.*

Proof The sufficiency part is a trivial consequence of the definition. So, let $\bar{\alpha} \in \operatorname{Dom} L_f$ be arbitrary such that $\bar{\alpha} > \inf_X f$.

We can assume without any loss of generality that $f(\bar{x}) = \bar{\alpha}$. Indeed, if $f(\bar{x}) < \bar{\alpha}$ and α_n is an arbitrary sequence which tends to $\bar{\alpha}$, then for sufficiently large n we can choose $x_n \equiv \bar{x}$.

Let now $\bar{x}$ be such an element of $L_f(\bar{\alpha})$ that $f(\bar{x}) = \bar{\alpha}$, and let β_n be an arbitrary sequence tending to $\bar{\alpha}$.

We construct a new sequence x'_n from the sequence x_n.

$$x'_n := \begin{cases} \bar{x}, \text{ if } \beta_n \geq \bar{\alpha}; \\ x_k, \text{ if } \beta_n < \bar{\alpha} \text{ , where } \alpha_k \leq \beta_n < \alpha_{k+1}. \end{cases} \tag{1}$$

It is easy to see that this new sequence tends to $\bar{x}$, and $x'_n \in L_f(\beta_n)$ for all n.

Because β_n was arbitrarily chosen, L_f is lower semicontinuous at $\bar{\alpha}$. □

The next theorem is a modified version of [7, Theorem 2.3]. The authors got a similar result when X and Y are subsets of $\mathbb{R}^n$.

Theorem 2.1 *Let X, Y be metric spaces, $S\colon X \to Y$ be a homeomorphism, and let $g\colon Y \to \mathbb{R}$ be a function so that L_g is lower semicontinuous. Then the function $f\colon X \to \mathbb{R}$, $f(x) := g(Sx)$ has a lower semicontinuous lower-level-set function L_f. Moreover, every local minimizer of f is global.*

Proof Let $\bar{\alpha}$ be an arbitrary element of $\operatorname{Dom} L_f$. Then it is also an element of $\operatorname{Dom} L_g$ and vice versa. Choose an $\bar{x} \in L_f(\bar{\alpha})$ and a sequence $\{\alpha_n\}_{n\in\mathbb{N}} \subset \operatorname{Dom} L_f$ tending to $\bar{\alpha}$. In view of Lemma 2.1, we can assume without any loss of generality that

$$f(\bar{x}) = \bar{\alpha} \qquad \text{and} \qquad \alpha_n \uparrow \bar{\alpha}. \tag{2}$$

Let us define the following two non-empty sets and a sequence:

$$Y_n := \{y \in Y | g(y) \leq \alpha_n\}, \qquad X_n := S^{-1}(Y_n), \qquad r_n := \inf_{x\in X_n} d(x,\bar{x}). \tag{3}$$

Choose an arbitrary positive ε. Now, we can construct a sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$ such that

$$d(x_n, \bar{x}) < r_n + \frac{\varepsilon}{2^n}. \tag{4}$$

Such a sequence exists because of the definition of r_n and the positivity of $\frac{\varepsilon}{2^n}$. S is a homeomorphism, so for every $n \in \mathbb{N}$ there is a $y_n \in Y_n$ such that $x_n = S^{-1}y_n$.

Using (2), we have that $X_n \subset X_{n+1}$. From this, we get that $\{r_n\}_{n\in\mathbb{N}}$ is monotone decreasing. On account of (3), it is also non-negative, so it has a non-negative limit r. If it is positive, then for all $x \in \cup_{n\in\mathbb{N}} X_n$,

$$d(x, \bar{x}) \geq r > 0. \tag{5}$$

On the other hand, $\cup_{n\in\mathbb{N}} X_n = \cup_{n\in\mathbb{N}} S^{-1}(Y_n)$, so there exists a sequence $\{y_n\}_{n\in\mathbb{N}} \subset Y_n$, which tends to $\bar{y} := S(\bar{x})$, since L_g is lower semicontinuous. Because of the continuity of S,

$$X_n \ni x_n := S^{-1}(y_n) \to S^{-1}(\bar{y}) := \bar{x},$$

which contradicts (5). This means that $r = 0$. Using (4), the distance $d(x_n, \bar{x})$ tends to zero as n tends to infinity. It follows that L_f is lower semicontinuous at $\bar{\alpha}$, but $\bar{\alpha}$ was arbitrarily chosen, so L_f is lower semicontinuous.

The last part follows from [4, Proposition 5.1]. □

It is well known that convex functions (in this case X must be such a space that this concept becomes meaningful) have convex lower level sets, hence they have connected lower level sets. Consequently, the corresponding lower level set function of a convex function is lower semicontinuous. This is the reason why every local minimizer of a convex function is global. The same trail of thoughts works similarly with quasi convex functions. The next corollary uses this and the previous theorem.

Corollary 2.1 *Let X, Y be normed spaces, let $S\colon X \to Y$ be a homeomorphism, and let $g\colon Y \to \mathbb{R}$ be a convex function. Then $L_{g\circ S}$ is lower semicontinuous, that is, every local minimizer of $g \circ S$ is global.*

If g is strictly convex, then $g \circ S$ has at most one global minimizer.

Proof We prove that $g \circ S$ has connected lower level sets. Let α be an arbitrary element of $\operatorname{Dom} L_g = L_{g\circ S}$. Define the following set:

$$Y_\alpha := \{y \in Y | g(y) \leq \alpha\}. \tag{6}$$

Since this set is convex, it is connected also. Let the preimage of Y_α under S be X_α. A continuous mapping preserves connectedness, so $X_\alpha = S^{-1}(Y_\alpha)$ is connected.

Assume that g is strictly convex. Then we have two distinct cases.

In the first one is, $\operatorname{Dom} L_{g\circ S}$ is not bounded from below. In this case, there is no minimizer at all.

In the second one, $\operatorname{Dom} L_{g\circ S}$ is bounded from below. Denote by $\bar{\alpha} \in \mathbb{R}$ the infimum. If there is no $y \in Y$ such that $g(y) = \bar{\alpha}$, then there is no minimizer again, because S is a homeomorphism.

At last, assume that there exist $x_1, x_2 \in X$ such that $g(Sx_1) = g(Sx_2) = \bar{\alpha}$. As g is strictly convex, it has at most one global minimizer so, $Sx_1 = Sx_2$. Since S is one-to-one, we have $x_1 = x_2$. □

Examples 2.1 – *Let us consider the so called Rosenbrock function (see* [9]*). $f\colon \mathbb{R}^2 \to \mathbb{R}^2,\ f(x,y) := 100(y-x^2)^2 + (1-x)^2$. This is a popular test function in nonlinear optimization. As is well-known, it is not convex. Define the following map $S\colon \mathbb{R}^2 \to \mathbb{R}^2,\ S(x,y) := (10(y-x^2), 1-x)$. A straightforward calculation shows that S is a homeomorphism. On the other hand, $f = g \circ S$, where $g\colon \mathbb{R}^2 \to \mathbb{R}^2,\ g(x,y) := x^2 + y^2$ is a convex function. We can apply now the previous corollary or Theorem* 2.1. *This entails that the Rosenbrock function has* *a lower semicontinuous lower-level-set function. Therefore, its every local minimizer is also global.*

– *If* $f\colon \mathbb{R}^n \to \mathbb{R}^n,\ f(x_1,\dots,x_n) := (x_2 - x_1^2)^2 + \cdots + (x_n - x_{n-1}^2)^2 + (1-x_1)^2$, *then an n dimensional version of the previous example can be derived in a very similar way.*

3 Quasi-Connected and Connected Functions

Let X be a topological space, and let $D \subset X$ be a nonempty set. A function $f\colon D \to \mathbb{R}$ is called *quasi-connected (connected) on D*, iff for all $\bar{x}, x \in D$ there exists a continuous function (*a path joining* $\bar{x}$ *and* x) $\gamma\colon [0,1] \to D$ such that γ and f fulfill the following conditions:

(i) $\gamma(t) \in D$ for all $t \in [0,1]$;

(ii) $\gamma(0) = \bar{x}$ and $\gamma(1) = x$;

(iii) $f(\gamma(t)) \le \max\{f(\bar{x}), f(x)\}$ $\big(f(\gamma(t)) \le (1-t)f(\bar{x}) + tf(x)\big)$ for all $t \in [0,1]$.

Remarks 3.1 – *In the definition of* γ *the order of the points* $\bar{x}, x$ *is very important. A more correct notation would be* $\gamma_{\bar{x},x}$, *but this is too troublesome. Therefore, we will use the simpler one when there is no ambiguity.*

– *It is quite straightforward that every connected function is also quasi-connected.*

– *The above concepts are a generalization of quasi-convexity (convexity), just take* $\gamma(t) = (1-t)\bar{x} + tx$, *which has a very important role in optimization theory (see, e.g.,* [10] *and the references therein).*

The function is called *strictly quasi-connected (strictly connected) on D* iff, whenever $\bar{x} \neq x$, γ may be chosen so that a strict inequality holds in (iii).

All the previously mentioned concepts were introduced in [5] when $X = \mathbb{R}^n$. These are preserved if we compose our quasi-connected (strictly quasi-connected, connected, strictly connected) function with a homeomorphism.

Theorem 3.1 *Let X, Y be metric spaces or topological spaces, let $S\colon X \to Y$ be a homeomorphism, and let $f\colon Y \to \mathbb{R}$ be a quasi-connected (strictly quasi-connected, connected, strictly connected) function. Then $f \circ S$ is also quasi-connected (strictly quasi-connected, connected, strictly connected).*

Proof Let $\bar{x}, x$ be arbitrary elements of X, and let $S\bar{x} =: \bar{y}, Sx =: y \in Y$. Then there exists a path $\gamma\colon [0,1] \to Y$ such that (i), (ii) and (iii) are fulfilled. Let us define a new path $\tilde{\gamma}\colon [0,1] \to X$, $\tilde{\gamma} := S^{-1} \circ \gamma$. Clearly, this new function is continuous, $\tilde{\gamma}(0) = S^{-1}(\gamma(0)) = S^{-1}\bar{y} = \bar{x}$, and $\tilde{\gamma}(1) = S^{-1}(\gamma(1)) = S^{-1}y = x$. Moreover, $f(S(\tilde{\gamma}(t))) = f(\gamma(t)) \le \max\{f(\bar{y}), f(y)\} = \max\{f(S\bar{x}), f(Sx)\}$. These show that $f \circ S$ is quasi-connected.

The remaining part of the proof runs in a very similar way, so we omit it. □

It is a well known fact that every lower level set of a convex or a quasi-convex function is a convex set, which is connected. The next corollary is a simple, but important consequence of the previous theorem.

Corollary 3.1 *Let X, Y be locally convex topological vector spaces, let $S\colon X \to Y$ be a homeomorphism, and let $f\colon Y \to \mathbb{R}$ be a quasi-convex (convex) function. Then $f \circ S$ is quasi-connected (connected).*

Proof Let $\gamma(t) = (1-t)x_0 + tx_1$ in the previous proof. □

In [1] the authors proved that connectedness (strict connectedness, strict quasi-connectedness) of a function implies that every local minimizer of this function is also global. However, it was proved only in the case when $X = \mathbb{R}^n$. If X is a topological space, then the proof runs in a pretty similar way.

Theorem 3.2 *Let X be a topological space, $D \subset X$, and let $f\colon D \to \mathbb{R}$ be a function. If f is connected(strictly connected, strictly quasi-connected) on D, then every local minimizer of f is also global.*

Examples 3.1 – *Let X be a reflexive Banach space, let $S\colon X^* \to X$ be a continuous (with respect to the norm topology both in X and in X^*), strongly monotone operator, and let $g\colon X \to \mathbb{R}$ be a convex function. Strong monotonicity implies that S is a bijection. The continuity of S^{-1} follows from the reflexivity of the space and from the continuity of S. According to Theorem* 3.1*, f is connected. On the other hand, Theorem* 3.2 *entails that every local solution of the hereunder optimization problem is global.*

$$\min_{x^* \in X^*} f(x^*) := g(Sx^*).$$

– *Let us minimize the following objective function subject to an ODE.*

$$\begin{aligned} &\min \|x\|^2_{L^2(0,1)} := \int_0^1 x^2(t)dt \\ \textit{s.t.} \quad &-x''(t) + h(x(t)) = u(t), \quad t \in]0,1[\\ &x(0) = x(1) = 0, \end{aligned} \tag{7}$$

where the nonlinear term $h\colon \mathbb{R} \to \mathbb{R}$ is a bounded, continuous, monotone function and $u \in L^2(0,1)$ are given.

Firstly, we rephrase this problem. For this, we need the weak formulation of the above boundary value problem.

We denote with $V = H^1_0(0,1)$ the Hilbert space of such $L^2(0,1)$ functions, whose first weak derivative is also in $L^2(0,1)$, and fulfill the previously given boundary condition. This makes sense in this case, because the space is one dimensional, so all the elements of V are continuous.

A function $x \in V$ is called a weak solution *of the above boundary value problem iff*

$$\int_0^1 x'\varphi' + \int_0^1 h(x)\varphi = \int_0^1 u\varphi \qquad \textit{for all } \varphi \in V.$$

Let us define the operator $A\colon V \to V^*$ and the functional $F\colon V \to \mathbb{R}$ in the following way.

$$\langle Ax, v\rangle := \int_0^1 x'v' + \int_0^1 h(x)v, \qquad v \in V,$$
$$F(v) := \int_0^1 uv, \qquad v \in V.$$

The operator A is continuous and strongly monotone, moreover, $F \in V^*$ (for the details see [11]*). This means that we have the same conclusion as in the previous example.*

4 Directional Differentiation with respect to a Path

Let X be a real normed space, $D \subset X$, and let $f\colon D \to \mathbb{R}$ be a function. Assume that $\bar{x}, x \in D$ are given and $\gamma\colon [0,1] \to D$ is a path joining them. That is to say, γ is continuous, $\gamma(t) \in D$, $t \in [0,1]$, and $\gamma(0) = \bar{x}$, $\gamma(1) = x$. We say that f *is directionally differentiable with respect to* γ *at* $\bar{x}$ iff the following limit exists.

$$\lim_{t \downarrow 0} \frac{f(\gamma(t)) - f(\bar{x})}{t} =: f'(\bar{x}, \gamma). \tag{8}$$

Proposition 4.1 *Assume that f is Fréchet differentiable at $\bar{x}$, and directionally differentiable with respect to a path γ, which is differentiable at* 0*, i.e., the limit* $\gamma'(0) := \lim_{t\downarrow 0} \frac{\gamma(t)-\gamma(\bar{x})}{t}$ *exists. Then*

$$f'(\bar{x}, \gamma) = f'(\bar{x})\gamma'(0). \tag{9}$$

Proof It is an easy consequence of the chain rule for Fréchet differentiability. □

A function f is called *regularly quasi-connected (regularly connected)* iff for every $\bar{x}, x \in X$, γ can be chosen such a way that it is differentiable at 0.

Proposition 4.2 *If $f: D \to \mathbb{R}$ is connected and differentiable with respect to the corresponding path, then*

$$f'(\bar{x},\gamma) \leq f(x) - f(\bar{x}) \quad \textit{for all } \bar{x}, x \in D.$$

If f is Fréchet differentiable and regularly connected, then

$$f'(\bar{x})\gamma'(0) \leq f(x) - f(\bar{x}) \quad \textit{for all } \bar{x}, x \in D.$$

Proof Connectivity implies that

$$f(\gamma(t)) \leq (1-t)f(\bar{x}) + tf(x),$$

that is,

$$\frac{f(\gamma(t)) - f(\bar{x})}{t} \leq f(x) - f(\bar{x}), \qquad t \in]0,1].$$

Taking the limit $t \downarrow 0$, we get the first part.

The second one follows from the first part and the previous proposition. □

The next theorem gives a necessary condition for the local optimality in the form of a variational inequality. This condition also becomes sufficient whenever the function is connected.

Theorem 4.1 *Assume that f has a local minimum at $\bar{x}$. If $x \neq \bar{x}$ is such a point that there exists a path γ joining $\bar{x}$ with x, and f is directionally differentiable with respect to γ, then*

$$f'(\bar{x},\gamma) \geq 0. \tag{10}$$

If for all $x \neq \bar{x}$ there is a corresponding γ such that (10) *is fulfilled, and f is connected, then $\bar{x}$ is a global minimizer of f.*

Remarks 4.1 – *In the second part of the theorem the assumption implies the directional differentiability of f with respect to the corresponding path.*

– *It is very important to note that f must be directionally differentiable with respect to the same paths, which ensure its connectivity.*

Proof Let $\bar{x}$ be a local minimizer of f and γ such a path, which is mentioned in the first part of the theorem. Using the continuity of γ, there exists $\bar{t} \in]0,1[$ such that

$$f(\gamma(t)) \geq f(\bar{x})$$

for all $t \in [0,\bar{t}]$. Clearly, $\bar{t}$ depends on x here. From this we get

$$\frac{f(\gamma(t)) - f(\bar{x})}{t} \geq 0$$

for all $t \in]0,\bar{t}]$. Taking the limit $t \downarrow 0$ of the right-hand side, we have the first part of the theorem.

For the second part, let f be connected such that (10) is fulfilled for all $x \neq \bar{x}$; then from the definition of connectivity, we get for an arbitrary x

$$\frac{f(\gamma(t)) - f(\bar{x})}{t} \leq f(x) - f(\bar{x})$$

for every $t \in]0,1]$. Taking the limit $t \downarrow 0$ of the right-hand side and using (10), we come to the following conclusion

$$0 \leq f'(\bar{x},\gamma) \leq f(x) - f(\bar{x}).$$

Since x was arbitrary, the proof is ready. □

It is worthwhile to mention that the connectedness of f is not completely used in the second part of the proof.

Examples 4.1 – *Let us consider the following function $f\colon \mathbb{R} \to \mathbb{R},\ f(x) = \sqrt{|x|}$. Then f is not directionally differentiable at zero in the usual sense. Moreover, it* *is not directionally differentiable at zero in the sense of Clarke either. Indeed,*

$$f^{\circ}(\bar{x};h) = \limsup_{x\to 0,\ t\downarrow 0} \frac{f(x+th)-f(x)}{t} = \limsup_{x\to 0,\ t\downarrow 0} \frac{\sqrt{|x+th|}-\sqrt{|x|}}{t} = \infty,$$

where $f^{\circ}(\bar{x};h)$ *denotes the Clarke directional derivative at* $\bar{x}=0$ *in the direction* $h\neq 0$ *(for more information about the Clarke derivative see, e.g.,* [12]*).*

Let $x\in\mathbb{R}\setminus\{0\}$*, and define the corresponding path in the following way,* $\gamma_x(t) := xt^2$*. Then* γ_x *is continuous,* $\gamma_x(0)=0$ *and* $\gamma_x(1)=x$*. So, this path joins* 0 *and* x*. Furthermore,* f *is directionally differentiable with respect to* γ_x *for all* $x\in\mathbb{R}\setminus\{0\}$*. Indeed,*

$$\lim_{t\downarrow 0} \frac{f(\gamma_x(t))-f(0)}{x} = \lim_{t\downarrow 0} \frac{\sqrt{|xt^2|}-\sqrt{0}}{t} = \sqrt{|x|} \geq 0.$$

The function f *is also quasi connected with respect to* $\gamma_{\bar{x},x}(t) = t^2x+(1-t)^2\bar{x}$*. Using Theorem* 4.1 *we have that* 0 *is a global minimizer of* f*.*

- *Let us consider the following Rosenbrock-type function* $f\colon \mathbb{R}^2\to\mathbb{R}$*,* $f(x,y) := (y-x^2)^2+(1-x)^2$*, and define the family*

$$\gamma_{(\bar{x},\bar{y}),(x,y)}(t) := ((1-t)\bar{x}+tx,(1-t)^2\bar{y}+2(1-t)tx\bar{x}+t^2y).$$

Then f *is convex with respect to this family on* $\mathbb{R}^2$ *and* $f'(\bar{x},\gamma_{(1,1),(x,y)})=0$*. This means that we can apply Theorem* 4.1*. Therefore,* $(1,1)$ *is a global minimizer of* f*.*

5 An Application

Let $\Phi\colon [a,b]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ be a given continuous function. Consider the following problem.

$$\inf_{x\in X_{ad}} J_{\Phi}(x) := \int_a^b \Phi(s,x(s),x'(s))ds, \tag{11}$$

where $C^1([a,b])$ denotes the class of all continuously differentiable functions on $[a,b]$, and the set $X_{ad} = \{x \in C^1([a,b]) \mid x(a) = x(b) = 0\}$ is called *admissible set*, or *the set of admissible functions.*

This is the simplest problem from calculus of variations. The most important first order necessary condition (under appropriate smoothness condition on Φ) is the Euler-Lagrange equation

$$D_2\Phi - \frac{\partial}{\partial t}\big(D_3\Phi\big) = 0, \tag{12}$$

where D_i denotes the derivative with respect to the ith variable. If $\bar{x}$ is a solution of (11), then it is also a solution of (12). The reverse is not true in general, but if, e.g., Φ is convex with respect to its last two variables, then every solution of (12) also becomes a solution of (11) (see, e.g., [13] or [14]).

We give a generalization of this result. For this, we need the following assumption.

(A) Let Φ be regularly connected with respect to its last two variables in such a way that, for every $x \in X_{ad}$, the derivative at 0 of the corresponding one-parameter family of paths $\gamma_s\colon [0,1] \to X_{ad},\ s \in [a,b]$ has the form $\gamma_s'(0) = (\alpha(s) - x(s), \alpha'(s) - x'(s))$, where $\alpha \in X_{ad}$. In other words, the derivative of the path at zero can be represented as a difference of two admissible functions and their derivatives.

Remark 5.1 *The previous assumption requires that the set* $\{(x,x')|x \in X_{ad}\}$ *is connected with respect to the corresponding* γ_s*.*

Theorem 5.1 *Assume that* Φ *is continuously differentiable and fulfills assumption (A). Then every local minimizer of* (11) *is global.*

Proof Let us define the following function $\Phi_s(x,x') := \Phi(s,x(s)x'(s))$ for every $s \in [a,b]$. Using the assumptions of the theorem and Proposition 4.1, we have

$$\langle \nabla\Phi_s(\bar{x},\bar{x}'), \gamma_s'(0)\rangle = \Phi_s'((\bar{x},\bar{x}'),\gamma_s) \leq \Phi_s(x,x') - \Phi(\bar{x},\bar{x}') \tag{13}$$

for all $s \in [a,b]$. On account of the assumption (A), there exists an $\alpha \in X_{ad}$ such that $\gamma_s'(0) = (\alpha(s) - \bar{x}(s), \alpha'(s) - \bar{x}'(s))$. Using this and (13)- after integrating both sides of the inequality on the interval $[a,b]$ and applying the theorem of integration by parts- we get

$$\int_a^b \left(D_2\Phi(s,\bar{x}(s),\bar{x}'(s)) - \frac{\partial}{\partial s}\big(D_3\Phi(s,\bar{x}(s),\bar{x}'(s))\big) \right) \big(\alpha(s) - \bar{x}(s)\big) ds + \\ \big[D_3\Phi(s,\bar{x}(s),\bar{x}'(s))\big(\alpha(s) - \bar{x}(s)\big)\big]_a^b \leq \\ \int_a^b \Phi(s,x(s),x'(s))ds - \int_a^b \Phi(s,\bar{x}(s),\bar{x}'(s))ds = J_\Phi(x) - J_\Phi(\bar{x}).$$

The left-hand side of the inequality is equal to zero if $\bar{x}$ is a local minimizer. Indeed, the integral on the left-hand side is equal to zero by virtue of (12). The second term is equal to zero, because $\alpha, \bar{x} \in X_{ad}$. The function $x \in X_{ad}$ was arbitrarily chosen, so

$$0 \leq J_\Phi(x) - J_\Phi(\bar{x})$$

for every $x \in X_{ad}$, that is, $\bar{x}$ is a global minimizer. □

Example 5.1 (The simplest case of the Almansi-Wirtinger-Friedrichs-Poincaré inequality family)

The Wirtinger inequality states that

$$\int_0^1 x^2 ds \leq \frac{1}{\pi^2} \int_0^1 x'^2 ds \tag{14}$$

for all $x \in X_{ad} = \{x \in C^1([0,1]) \mid x(0) = x(1) = 0\}$*, and the constant* $\frac{1}{\pi^2}$ *can not be improved (see* [15] *or* [14]*). Probably the first developer was Almansi in 1906 (see* [16]*), but in the literature it is called Wirtinger inequality, or Friedrichs-Poincaré inequality, because it is a special case of Friedrichs inequality, and closely related to Poincaré inequality. These last two are very important tools in the Sobolev space theory and also in the theory of PDE (see, e.g.,* [17] *or* [18]*). In* [15] *inequality* (14) *was proved with the help of Plancherel identity and Hilbert space theory. On the* *other hand, in* [14] *the proof is an application and also an illustration of the use of the so called Fields theories. We give here a new proof of this theorem with the aid of Theorem* 5.1.

Following the line of [14]*, we have to prove that* $\Phi_\lambda(s,x,x') = \frac{x'^2-\lambda^2x^2}{2}$ *fulfills the assumptions of Theorem* 5.1*, where* λ *is a non-negative parameter. This function is not convex with respect to its last two variables, so the classical result, which states that in this case, every local minimizer is also a global one, cannot be applied.*

In the light of the previous paragraph, let us consider the following one-parameter family of problems:

$$\inf_{x\in X_{ad}} J_{\Phi_\lambda}(x) = \int_0^1 \frac{x'^2-\lambda^2x^2}{2}ds. \tag{15}$$

If $\lambda > \pi$*, then the infimum is* $-\infty$ *(see* [14]*), so we can assume that* $0 \le \lambda \le \pi$*. The Euler-Lagrange equation is*

$$x'' + \lambda^2 x = 0.$$

Obviously, $\bar{x} \equiv 0$ *is a solution of this equation. Actually this is the only solution if* $0 \le \lambda < \pi$ *and* $x(s) = c\sin\pi s,\ c \in \mathbb{R}$ *if* $\lambda = \pi$*. In all these cases the infimum of* (15) *is* 0 *(see* [14]*).*

Assume that $0 \le \lambda < \pi$ *and* $p \in X_{ad}$ *is an arbitrarily fixed polynomial different from zero. This means that the degree of* p *is at least two, that is, its derivative* p' *is not zero on* $]0,\varepsilon[$ *with a suitable positive* ε*. Now, we define a "good" path to* p*. For this, let* $K := \sup_{]0,\varepsilon]} \frac{p^2}{p'^2}$*, where* $K < \infty$*. We get this from the fact that the degree of* p *is higher than the degree of* p'*. For the same reason,* ε *can be chosen in such a way that* $\pi^2 \le \frac{1}{K}$*. Let us define now the path*

$$\gamma_s(t) := (\beta(t)p(s), \beta(t)p'(s)), \quad \textit{where} \quad \beta(t) = \begin{cases} \sqrt{t}, & \textit{if } t \in [\varepsilon,1]; \\ \frac{1}{\sqrt{\varepsilon}}t, & \textit{if } t \in [0,\varepsilon[. \end{cases}$$

At first, we have to prove that

$$\Phi_s(\gamma_s(t)) \leq (1-t)\Phi_s(0,0) + t\Phi_s(p,p') \qquad \text{for all } (s,t) \in [0,1]^2, \tag{16}$$

where 0 *denotes here the constant zero function, which is a polynomial element of* X_{ad}*. This inequality is equivalent to the following one.*

$$\lambda^2(t-\beta^2(t))p^2(s) \leq (t-\beta^2(t))p'^2(s) \qquad \text{for all } (s,t) \in [0,1]^2. \tag{17}$$

If $\lambda = 0$*, or* $t \in [\varepsilon, 1] \cup \{0\}$*, then the above-mentioned inequality is trivially true. As* $p \in X_{ad}$*, it is also true when* $s = 0$*. Taking into account of these and the assumption that* $p' \neq 0$ *on* $]0, \varepsilon[$*, we have to prove that*

$$\lambda^2 \frac{p^2(s)}{p'^2(s)} \leq \frac{t - \frac{1}{\varepsilon}t^2}{t - \frac{1}{\varepsilon}t^2} = 1.$$

However, $\lambda < \pi,\ \pi^2 \leq \frac{1}{K}$*, where* $K := \sup_{]0,\varepsilon]} \frac{p^2}{p'^2}$*, that is, the left-hand side of the previous inequality is less than or equal to* 1*, so* (16) *is fulfilled.*

We check now whether $\gamma_s'(0)$ *has the required form.*

$$\frac{\gamma_s(t) - \gamma_s(0)}{t} = \frac{1}{\sqrt{\varepsilon}}(p(s), p'(s)), \qquad t \in]0, \varepsilon[,\ s \in [0,1].$$

This means that $\gamma_s'(0) = \frac{1}{\sqrt{\varepsilon}}(p(s), p'(s))$*, which implies the suitable form of* $\gamma_s'(0)$ *with* $\alpha(s) = \frac{1}{\sqrt{\varepsilon}}p(s)$*.*

All in all, we have that

$$\int_0^1 \frac{p'^2 - \lambda^2 p^2}{2} ds = J_{\Phi_\lambda}(p) \geq J_{\Phi_\lambda}(0) = 0$$

for all polynomial elements of X_{ad} and for all $0 \leq \lambda < \pi$. Taking the limit $\lambda \to \pi$, we can derive from this that

$$\int_0^1 p^2 ds \leq \frac{1}{\pi^2} \int_0^1 p' ds \tag{18}$$

for all polynomial elements of X_{ad}.

Let $x \in X_{ad}$ be an arbitrary function now, then let us define $p_n := B_n(x)$, where B_n denotes the nth Bernstein operator

$$B_n(x(s)) := \sum_{k=0}^{n} x\left(\frac{k}{n}\right) B_{k,n}(s), \quad \text{where } B_{k,n}(s) := \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}.$$

Evidently, p_n is a polynomial and $p(0) = p(1) = 0$, hence $p_n \in X_{ad}$. On the other hand, using the continuous differentiability of x and the theorem of Lorentz (see [19] *or* [20]*), (p_n, p_n') tends to (x, x') uniformly on $[0,1]$. Because of the these, we can take the limit in* (18)*, and we have*

$$\int_0^1 x^2 ds \leq \frac{1}{\pi^2} \int_0^1 x' ds$$

for all $x \in X_{ad}$.

6 Conclusions

Some possible characterizations of the local-gobal minimum property of functions were presented in the second and in the third sections. Roughly speaking, these results say that the objective function has the local-global minimum property, if either its lower -level-set function is nice, or it is convex in some generalized sense. Both these properties are preserved under an appropriate modification (homeomorphism) of the domain (Theorems 2.1 and 3.1).

If a function is convex in some generalized sense, then the the corresponding first order necessary condition becomes sufficient (Theorem 4.1). For this, we need a new directional derivative, which fits better to the introduced generalized convexity. This is included in the fourth section.

The use of the new variational inequality (Theorem 4.1) is confirmed by Theorem 5.1 and by the example in the fifth section.

All in all, the goal of this work is to show, how we can avoid cumbersome second order conditions, with the help of generalized convexity. Some possible ways for this were presented, and supported by examples from both finite and infinite dimensional optimization problems.

Acknowledgements

This research has been supported by the Hungarian Scientific Research Fund OTKA "Mobility" call HUMAN-MB8A-84581 and Zoltán Magyary Postdoctoral Scholarship A2-MZPD-12-0253.

References

1. Avriel, M., Zang, I.: Generalized arcwise-connected functions and characterizations of local-global minimum properties. J. Optim. Theory Appl. 32(4), 407–425 (1980)
2. Bhatia, D., Mehra, A.,: Optimality conditions and duality involving arcwise connected and generalized arcwise connected functions. J. Optim. Theory Appl. 100(1), 181–194 (1999)
3. Horst, B.: A note on functions whose local minima are global. J. Optim. Theory Appl. 36(3), 457–463 (1982)
4. Luc, D.T., Volle, M.: Levels sets infimal convolution and level addition. J. Optim. Theory Appl. 94(3), 695–714 (1997)

5. Ortega, J.M., Rheinboldt, W.C.: Iterative solution of nonlinear equations in several variables, volume 30 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2000)
6. Avriel, M., Zang, I.: On functions whose local minima are global. J. Optimization Theory Appl. 16, 183–190 (1975)
7. Avriel, M., Choo, E.U., Zang, I.: A note on functions whose local minima are global. J. Optimization Theory Appl. 18(4), 555–559 (1976)
8. Aubin, J.P., Frankowska, H.: Set-valued analysis, volume 2 of Systems and Control: Foundations and Applications. Birkhäuser Boston Inc., Boston, MA (1990)
9. Rosenbrock, H.H. : An automatic method for finding the greatest or least value of a function. Comput. J. 3, 175–184 (1960/1961)
10. Jahn, J.: Introduction to the theory of nonlinear optimization. Springer-Verlag, Berlin (1996)
11. Drábek, P., Milota, J.: Methods of nonlinear analysis. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel (2007)
12. Clarke, F.H.: Optimization and nonsmooth analysis. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley and Sons Inc., New York (1983)
13. Gelfand, I.M., Fomin, S.V.: Calculus of variations. Revised English edition translated and edited by Richard A. Silverman. Prentice-Hall Inc., Englewood Cliffs, N.J. (1963)
14. Dacorogna, B.: Introduction to the calculus of variations. Imperial College Press, London (2009)
15. Dym, H., McKean, H.P.: Fourier series and integrals. Academic Press, New York (1972)
16. Almansi, E.: Sopra una delle esperienze del plateau. Annali di Matematica Pura ed Applicata (1898 - 1922) 12, 1–17 (1906)

17. Evans, L.C.: Partial differential equations. volume 19 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (1998)
18. Tröltzsch, F.: Optimal control of partial differential equations, volume 112 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (2010)
19. Lorentz, G.: Zur theorie der polynome von S. Bernstein. Rec. math., Moscou 2(44), 543–556 (1937)
20. DeVore, R.A., Lorentz, G.G.: Constructive approximation, volume 303 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin (1993)