SIAM J. OPTIM.
Vol. 29, No. 2, pp. 994–1016

VARIATIONAL PRINCIPLES IN ANALYSIS AND EXISTENCE OF MINIMIZERS FOR FUNCTIONS ON METRIC SPACES*

ARAM V. ARUTYUNOV† AND SERGEY E. ZHUKOVSKIY‡

Abstract. Functions defined on metric spaces are studied. For these functions, a generalized Caristi-like condition is introduced. It is shown that this condition is sufficient for a bounded below, lower semicontinuous function to attain its minimum. Criteria for a generalized Caristi-like condition to hold are derived. Generalizations of the Ekeland and Bishop–Phelps variational principles are obtained and compared with their prototypes.

Key words. Ekeland variational principle, Bishop–Phelps variational principle, existence of minima of functions on metric spaces

AMS subject classifications. 58E30, 49K27

DOI. 10.1137/18M1164287

1. Introduction.

During the 1970s two theorems that played an important role in the further development of nonlinear analysis—the Bishop–Phelps variational principle (BPVP) (see [25] and section 2.2 of [18]) and the Ekeland variational principle (EVP) (see, for example, [9])—were obtained. It turned out that these principles were equivalent (see, for example, [17]). Recall the EVP (the BPVP is formulated and discussed in sections 3 and 4).

EVP. *Let X be a complete metric space and let $U : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function, that is bounded below by a given $\gamma \in \mathbb{R}$. Then for all $x_0 \in X$, for all $\varepsilon > 0$ such that $U(x_0) \le \gamma + \varepsilon$, for all $\lambda > 0$ there exists a point $\bar{x} \in X$ such that*

$$U(\bar{x}) \le U(x_0), \quad \rho(\bar{x}, x_0) \le \lambda, \tag{1.1}$$

$$U(x) + \frac{\varepsilon}{\lambda}\rho(x, \bar{x}) > U(\bar{x}) \quad \forall\, x \ne \bar{x}. \tag{1.2}$$

It is important to note that this principle is a variational formulation of the notion of metric completeness, i.e., if every bounded from below continuous function defined on a metric space satisfies the EVP then this space is complete (see [26]).

Later the Borwein–Preiss variational principle was obtained (see [13]), which allows one to consider perturbations that are of a more complicated form but are more smooth in a certain sense. The Borwein–Preiss variational principle and the EVP do not follow from each other, since in the EVP perturbations are not smooth and the Borwein–Preiss variational principle does not guarantee that the perturbed function

*Received by the editors January 8, 2018; accepted for publication (in revised form) December 26, 2018; published electronically April 3, 2019.
http://www.siam.org/journals/siopt/29-2/M116428.html
Funding: This research was supported by the "RUDN University Program 5-100," the Russian Foundation for Basic Research (projects 17-51-12064, 18-01-00106, 18-01-00590, and 19-01-00080), and by the Volkswagen Foundation.
†Peoples' Friendship University of Russia (RUDN University), Moscow, 117198, Russia; V. A. Trapeznikov Institute of Control Sciences of RAS, Moscow, 117997, Russia; and Institute for Information Transmission Problems of RAS, Moscow, 127051, Russia (arutyunov@cs.msu.ru).
‡Department of Higher Mathematics, Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, 141701, Russia and V. A. Trapeznikov Institute of Control Sciences of RAS, Moscow, 117997, Russia (S-E-Zhuk@yandex.ru).

has a strict minimum. Later, an analogue of the variational principle of Borwein and Price was obtained and used in [1] to study the problems of optimal control theory. In [5], a smooth variational principle was obtained that was applied to derive necessary second-order conditions for minimizing sequences. We also refer the reader to the papers [19, 20, 21] and the book [14, Chapter 2] for further developments in this area.

The EVP is widely used in optimization theory. For example, using the EVP for a wide class of optimal control problems, including nonsmooth problems, necessary optimality conditions in the form of Euler–Lagrange equations, the Weierstrass condition, and the Pontryagin maximum principle were obtained (see, for example, [23]). Let us demonstrate the idea of the EVP application by way of the simple example of the Lagrange multiplier rule proof (see [16]).

Let X be a Banach space, $f_0 : X \to \mathbb{R}$ be a smooth function, $F : X \to \mathbb{R}^k$ be a smooth map, and $x_0 \in X$ be a point of local minimum of the problem

$$f_0(x) \to \min, \quad F(x) = 0.$$

For an arbitrary $\varepsilon > 0$, applying the EVP to the function

$$\varphi_\varepsilon(x) := \max\{f_0(x) - f_0(x_0) + \varepsilon, 0\} + |F(x)|, \quad x \in X,$$

with $\lambda = \sqrt{\varepsilon}$, we obtain that there exists a point of minimum $x_\varepsilon \in X$ of the function $\varphi_\varepsilon(x) + \sqrt{\varepsilon}\|x - x_\varepsilon\|$ in the variable x such that $x_\varepsilon \to x_0$ as $\varepsilon \to 0$. Then, calculating the subdifferentials of the functions φ_ε and $|F|$, we obtain that there exist $\lambda_{0,\varepsilon} \geq 0$ and $\eta_\varepsilon \in \mathbb{R}^k$ such that $\lambda_{0,\varepsilon} + |\eta_\varepsilon| \geq 1$ and $\lambda_{0,\varepsilon} f_0'(x_\varepsilon) + F'(x_\varepsilon)^*\eta_\varepsilon \in \sqrt{\varepsilon}B$. Here B stands for the unit ball in X^*. Normalizing $(\lambda_{0,\varepsilon}, \eta_\varepsilon)$ and passing to the limit as $\varepsilon \to 0$, we obtain the Lagrange multiplier rule for the initial problem.

The EVP is applied in the investigation of many other problems of variational analysis (see, for example, [12, 11, 24, 15, 4]). In [3], the EVP was used to derive sufficient conditions for a lower semicontinuous function to attain its minimal value without assuming the compactness of the domain. We recall this as follows.

Let (X, ρ) be a complete metric space, and let $U : X \to \mathbb{R} \cup \{+\infty\}$ be a function that is lower semicontinuous and bounded below by a certain $\gamma \in \mathbb{R}$. Denote by dom$U$ the domain of U, i.e.,

$$\mathrm{dom}U = \{x \in X : U(x) < +\infty\}.$$

THEOREM 1.1 (see [3, Theorem 3]). *Let the function U satisfy the Caristi-like condition (CLC) with constants $m \in \mathbb{R} \cup \{+\infty\}$ and $k > 0$, i.e.,*

$$\forall\, x \in X : \quad \gamma < U(x) \leq m \quad \exists\, x' \neq x : \quad U(x') + k\rho(x, x') \leq U(x).$$

Then for an arbitrary $x_0 \in X$ such that $U(x_0) \leq m$, there exists a point $\bar{x} \in X$ at which the function U attains its minimum and, moreover, $U(\bar{x}) = \gamma$ and

$$\rho_X(x_0, \bar{x}) \leq \frac{U(x_0) - \gamma}{k}.$$

In [27], it was proved that the CLC is sufficient for minimum existence, however the estimate of the distance from a given point to a point of minimum was not obtained. In [28], a generalization of the results from [27] was derived. In [3], Theorem 1.1 was applied to derive sufficient conditions for the existence of fixed points and coincidence points. For more references devoted to this issue, see [7, 8, 10].

Theorem 1.1, the EVP, and the BPVP are equivalent to each other (see [17]).

Note that the given conditions about the existence of a minimum from Theorem 1.1 have a significant drawback. Namely, let X be a Banach space. Then the assumptions of Theorem 1.1 do not hold if the function U attains a strict minimum at a certain point and at this point the function U is differentiable. In fact, let $\bar{x}$ be a point of strict minimum of the function U and let $U(\bar{x}) = \gamma = 0$. Then, if the assumptions of Theorem 1.1 hold, then for all points x sufficiently close to $\bar{x}$, the inequality $U(x) \geq \|\bar{x} - x\|/k$ holds. Thus, the equality $U(\bar{x}) = 0$ implies that U is not differentiable at $\bar{x}$.

In sections 2 and 3, we introduce the generalization of Theorem 1.1, which is applicable to smooth functions as well. In section 4, we study conditions of applicability of the generalization of Theorem 1.1. In section 5, the generalizations of EPVP and BFVP are deduced, and in section 6, it is proved that the obtained estimates are qualitatively better than the corresponding estimates in the EVP and the BPVP. In section 7, we consider certain applications and generalizations of the obtained results.

2. Generalized Caristi-like condition and existence of a minimum of a function. Let (X, ρ) be a metric space, and let $U : X \to \mathbb{R} \cup \{+\infty\}$ be a given function. Everywhere below, we assume that (X, ρ) is complete, and U is lower semicontinuous and bounded from below by a given $\gamma \in \mathbb{R}$, i.e., $U(x) \geq \gamma$ for all $x \in X$.

Let us introduce a special set of scalar nonnegative functions. Let $\mathcal{K}$ stand for the set of all functions $k : (0, +\infty) \to (0, +\infty)$ that satisfy the following conditions: k is increasing (i.e., if $\tau_2 > \tau_1$, then $k(\tau_2) \geq k(\tau_1)$) and the function $1/k$ is summable in a neighborhood of zero. Here and below we use the following notation: $1/k$ stands for the function $\tau \mapsto \frac{1}{k(\tau)}$ and k^{-1} stands for the inverse function (if it exists). We will use the same notation for other functions as well.

For $k \in \mathcal{K}$ set

$$D_k(u) = \int_0^u \frac{d\tau}{k(\tau)}, \quad u \in \mathbb{R}_+.$$

Due to the properties of the function k the function $D = D_k$ is absolutely continuous, strictly increasing, concave and $D(0) = 0$. Denote the set of all functions D satisfying these properties by $\mathcal{D}$.

DEFINITION 2.1. *Let $m \in \mathbb{R} \cup \{+\infty\}$. If there exists a function $k \in \mathcal{K}$ such that*

$$\forall x \in X : \quad \gamma < U(x) \leq m \quad \exists x' \neq x : \quad U(x') + k(U(x) - \gamma)\rho(x', x) \leq U(x), \tag{2.1}$$

then we say that the function U satisfies the generalized Caristi-like condition (generalized CLC) with the constant m and the function k. If $m = +\infty$, then we say that U satisfies the generalized CLC with the function k.

If the function U satisfies the generalized CLC with a constant m and a constant function $k(\tau) \equiv k = \text{const.}$, then it satisfies the CLC with m and k.

Let us present a generalization of Theorem 1.1 that is applicable to the functions that satisfy the generalized CLC.

THEOREM 2.2. *Let the function U satisfy the generalized CLC* (2.1) *with a constant $m \in \mathbb{R} \cup \{+\infty\}$ and a function $k \in \mathcal{K}$. Then, for every $x_0 \in \mathrm{dom}U$ such that $U(x_0) \leq m$, there exists a point $\bar{x} \in X$ at which the function U attains its minimum and, moreover, $U(\bar{x}) = \gamma$ and*

$$\rho_X(x_0, \bar{x}) \leq D_k(U(x_0) - \gamma) = \int_0^{U(x_0)-\gamma} \frac{dt}{k(t)}.$$

Proof. Without loss of generality, assume that $\gamma = 0$. Set $D = D_k$. It is obvious that the function $D(U(\cdot))$ is lower semicontinuous and $\inf_{x\in X} D(U(x)) = 0$. Let us show that the function $D(U(\cdot))$ satisfies the CLC with the constants $D(m)$ and $k = 1$, i.e., for an arbitrary point $x \in X$ at which $0 < D(U(x)) \le D(m)$, there exists a point $x' \neq x$ such that $D(U(x')) + \rho(x, x') \le D(U(x))$. Indeed, for an arbitrary point $x \in X$ at which $0 < D(U(x)) \le D(m)$, it is obvious that the function U doesn't attain its minimum. Thus, there exists a point $x' \neq x$ such that the inequality in (2.1) holds. Hence,

$$\rho(x,x') \le \frac{U(x) - U(x')}{k(U(x))} = \int_{U(x')}^{U(x)} \frac{d\tau}{k(U(x))} \le \int_{U(x')}^{U(x)} \frac{d\tau}{k(\tau)} = D(U(x)) - D(U(x')).$$

Here the second inequality follows from the fact that the function k is increasing. Therefore, the function $D(U(\cdot))$ satisfies the CLC with constants $D(m)$ and $k = 1$.

By virtue of Theorem 1.1, for every $x_0 \in \mathrm{dom}D(U(\cdot))$ satisfying $D(U(x_0)) \le D(m)$, there exists a point $\bar{x} \in X$ such that $D(U(\bar{x})) = 0$ and $\rho(x_0, \bar{x}) \le D(U(x_0))$. It is obvious that the function $D(U(\cdot))$ attains its minimum at those and only those points at which the function U attains its minimum. So, $\bar{x}$ is the desired point. ◻

In Theorem 2.2, the minimum point $\bar{x}$ may not be strict. A simple example is provided by the function $U(x) = \max\{0, x\}$, $x \in \mathbb{R}$, which satisfies the assumptions of Theorem 2.2 with $\gamma = 0$, $m = +\infty$, $k(\tau) \equiv 1$, and has no strict minima. The same remark is valid for Theorem 1.1.

Note also that Theorem 1.1 follows from Theorem 2.2 for $k(\tau) = \text{const.} = k$.

Theorem 2.2 implies the following assertion.

COROLLARY 2.3. *Let the function U satisfy the generalized CLC* (2.1) *with a constant $m \in \mathbb{R} \cup \{+\infty\}$ and a function $k \in \mathcal{K}$. Then, there exists a function $D \in \mathcal{D}$ such that for every $x_0 \in \mathrm{dom}U$ satisfying $U(x_0) \le m$, there exists a point $\bar{x} \in X$ of minimum of U satisfying the inequality*

$$\rho(\bar{x}, x_0) \le D(U(x_0) - U(\bar{x})). \tag{2.2}$$

The assumptions of Corollary 2.3 are not necessary conditions for the existence of the corresponding function $D \in \mathcal{D}$. Consider the corresponding example.

Example 2.4. Define a set $X \subset \ell_2$ and a function $U : X \to \mathbb{R}$ as follows. Set

$$x_n := \Big(\underbrace{0, \dots, 0, \frac{1}{n}}_{n}, 0, 0, \dots\Big) \in \ell_2, \quad n = 3, 4, \dots,$$

$$X := \{x_n : n = 3, 4, \dots\} \cup \{0\}, \quad U(x_n) := e^{-n}, \quad n = 3, 4, \dots, \quad U(0) := 0.$$

It is obvious that the metric space X is compact, the function U is continuous and bounded below by $\gamma := 0$, and $0 \in X$ is the only minimum of U. Let $k : (0, +\infty) \to (0, +\infty)$ be an arbitrary increasing function satisfying (2.1) with $m = +\infty$. Then

$$k(e^{-n}) \le \sup_{j>n} \frac{U(x_n) - U(x_j)}{\rho(x_n, x_j)} = \sup_{j>n} \frac{e^{-n}(1 - e^{-(j-n)})}{\sqrt{n^{-2} + j^{-2}}} = ne^{-n} \quad \forall\, n = 3, 4, \dots .$$

Thus, since k is increasing, we have $k(\tau) \le ne^{-n}$ for each $\tau \in (e^{-n-1}, e^{-n}]$, for each $n = 3, 4, \dots$. Therefore,

$$\int_0^{e^{-3}} \frac{d\tau}{k(\tau)} \ge \sum_{n=1}^{\infty} \frac{e^{-n} - e^{-n-1}}{ne^{-n}} = \sum_{n=1}^{\infty} \frac{e - 1}{en} = +\infty.$$

So, any function $k : (0, +\infty) \to (0, +\infty)$ for which (2.1) holds does not belong to $\mathcal{K}$. Therefore, the generalized CLC does not hold.

Now set

$$D(\tau) := -\frac{1}{\ln(\tau)}, \quad \tau \in (0, e^{-3}], \quad D(0) := 0.$$

Since

$$D'(\tau) = \frac{1}{\tau \ln^2(\tau)} > 0, \quad D''(\tau) = -\frac{2 + \ln(\tau)}{\tau^2 \ln^3(\tau)} \leq 0 \quad \forall\, \tau \in (0, e^{-3}],$$

we obtain $D \in \mathcal{D}$. Moreover,

$$\rho(0, x_n) = \frac{1}{n} = D(e^{-n}) = D(U(x_n)) \quad \forall\, n = 3, 4, \dots .$$

So, the function U does not satisfy the generalized CLC but estimate (2.2) holds. ∎

It turns out that under additional assumptions Corollary 2.3 is not only sufficient, but also a necessary condition. Let us show this. For every $D \in \mathcal{D}$ set

$$k_D(\tau) := \frac{\tau}{D(\tau)} \quad \forall\, \tau > 0, \quad k_D(0) = k_D(0+).$$

The function k_D has the following properties.

PROPOSITION 2.5. *For every $D \in \mathcal{D}$ the function k_D is absolutely continuous, increasing, and $k_D(\tau) > 0$ for every $\tau > 0$.*

Proof. Absolute continuity and positivity of k_D are obvious. The function k_D is increasing since

$$k'_D(\tau) = \frac{D(\tau) - \tau D'(\tau)}{D^2(\tau)} \geq 0 \quad \text{for almost all} \quad \tau > 0.$$

Here the last inequality holds since D is concave and $D(0) = 0$. ∎

So, the assumptions of Corollary 2.3 are necessary and sufficient conditions in the following case.

PROPOSITION 2.6. *Let $m \in \mathbb{R} \cup \{+\infty\}$ and let a function $D \in \mathcal{D}$ be such that the function $1/k_D$ be summable in a neighborhood of zero. Assume that for every $x_0 \in \mathrm{dom} U$ such that $U(x_0) \leq m$ there exists a point of minimum $\bar{x} \in X$ of the function U such that* (2.2) *holds. Then the function U satisfies the assumptions of Corollary* 2.3 *with the function k_D, i.e., the function U satisfies the generalized CLC* (2.1) *with the constant m and the function $k = k_D$ for $\gamma = \min_{x \in X} U(x)$.*

Proof. Without loss of generality assume that $\gamma = 0$. Proposition 2.5 implies that $k_D \in \mathcal{K}$. Moreover, for every $x \in \mathrm{dom} U$ satisfying inequality $U(x) \leq m$, there exists a point x' such that $U(x') = 0$ and $\rho(\bar{x}, x') \leq D(U(x))$. Hence,

$$U(x') + k_D(U(x))\rho(x, x') = \frac{U(x)}{D(U(x))}\rho(x, x') \leq U(x).$$

Thus, U satisfies the assumptions of Corollary 2.3 with the function k_D. ∎

Consider some examples of classes of functions U for which Theorem 2.2 provides necessary and sufficient conditions for existence of minima.

PROPOSITION 2.7. *Let $\Omega \subset \mathbb{R}^n$ be an open set, $U : \Omega \to \mathbb{R}$ be an analytic function, $X \subset \Omega$ be compact, and $\gamma = \min_{x\in X} U(x)$. Then there exist $c > 0$ and a positive integer p such that the restriction of U to X satisfies the generalized CLC with the function $k(\tau) \equiv c\tau^{1-\frac{1}{p}}$.*

Proof. The Lojasiewicz inequality (see, for example, [22]) implies that there exist $c > 0$ and positive integer p such that for every $x \in X$ there exists $\bar{x} \in X$ satisfying the relations $U(\bar{x}) = \gamma$ and $|x - \bar{x}| \le |U(x)|^{\frac{1}{p}}/c$. Set $D(\tau) := \tau^{\frac{1}{p}}/c$. Proposition 2.6 implies that the restriction of U to X satisfies the generalized CLC with the function k_D. Obviously

$$k_D(\tau) = \frac{\tau}{D(\tau)} = c\tau^{1-\frac{1}{p}} \quad \forall \tau > 0. \qquad \square$$

PROPOSITION 2.8. *Given a convex function $U : \mathbb{R} \to \mathbb{R}$, assume that its set of points of minimum H is nonempty, and $\gamma := \min_{x\in X} U(x)$. Then there exists a function $k \in \mathcal{K}$ such that the function U satisfies the generalized CLC* (2.1) *with the function k.*

Proof. First, let us prove an auxiliary proposition. Let $V : \mathbb{R}_+ \to \mathbb{R}_+$ be a convex function such that $V(0) = 0$ and $V(x) > 0$ for every $x > 0$. Let us show that V satisfies the generalized CLC with a function $k \in \mathcal{K}$.

It is obvious that V is one-to-one and absolutely continuous. Hence, the inverse function $W := V^{-1}$ exists, it is absolutely continuous, concave, and increasing. The identity $W(V(x)) \equiv x$ implies that $W'(V(x))V'(x) = 1$ for almost all $x \ge 0$. Take $k(\tau) := 1/W'(\tau)$, $\tau > 0$, $k(0) = 0$. It is a straightforward task to ensure that $k \in \mathcal{K}$. Indeed, it is obvious that $k(\tau) > 0$ for every $\tau > 0$. Moreover, k is increasing, since W is increasing. Finally, $1/k$ is integrable in a neighborhood of zero, since $1/k(\tau) \equiv W'(\tau)$ and W is absolutely continuous.

Let us show that V satisfies the generalized CLC with the function k. For every $x > 0$, for each $x' \in (0, x)$ we have $V(x') < V(x)$, and since V is convex we have

$$V(x') + k(V(x))|x' - x| = V(x') + \frac{x - x'}{W'(V(x))} = V(x') + V'(x)(x - x') \le V(x)$$

if V is differentiable at x. Since V is differentiable almost everywhere, the obtained inequality holds for all $x > 0$, $x' \in (0, x)$. So, V satisfies the generalized CLC with the function k.

Let us now pass to the proof of the initial proposition. Since U is convex, H is either a segment, a half-line, or the entire line. If $H = \mathbb{R}$ then the statement of the proposition obviously holds. If $H = (-\infty, a]$ (or $H = [a, +\infty)$) then the above proved auxiliary proposition implies that the restriction of U to $[a, +\infty)$ (or $(-\infty, a]$) satisfies the generalized CLC with some $k \in \mathcal{K}$ and, hence, U satisfies the generalized CLC with the same k. Consider the case in which $H = [a, b]$. Denote by V_1 and V_2 the restrictions of U to $(-\infty, a]$ and to $[b, +\infty)$, respectively. The above proved auxiliary proposition implies that V_1 and V_2 satisfy the generalized CLC with some functions $k_1 \in \mathcal{K}$ and $k_2 \in \mathcal{K}$, respectively. Hence, U satisfies the generalized CLC with $k(\tau) \equiv \min\{k_1(\tau), k_2(\tau)\}$. It is a straightforward task to ensure that $k \in \mathcal{K}$. $\square$

3. Generalized Caristi-like condition for differentiable functions. Let us discuss the generalized CLC (2.1) for differentiable functions defined on a Banach space. Let X be a Banach space with a norm $\|\cdot\|$. The norm in the space X^* that is the topological dual of X we also denote by $\|\cdot\|$. Denote by $S \subset X$ the unit sphere in X, i.e., $S = \{x \in X : \|x\| = 1\}$. Recall that the function $U : X \to \mathbb{R}$ is said to be differentiable at a point $x \in X$ in a direction $h \in X$ if the limit

$$U'(x;h) := \lim_{\varepsilon \to 0+} \frac{U(x+\varepsilon h) - U(x)}{\varepsilon}$$

exists.

Proposition 3.1. *Let there exist a function $k : (0, +\infty) \to (0, +\infty)$ such that for all $x \in X$ for which $U(x) > \gamma$ there exists a direction $h = h(x) \in S$ for which U is differentiable at the point x in the direction h and the following inequality holds*

$$k(U(x) - \gamma) < -U'(x;h). \tag{3.1}$$

Then the function U satisfies condition (2.1) *with the function k and $m = +\infty$.*

Proof. Assume again that $\gamma = 0$. Fix an arbitrary $x \in X$ for which $U(x) > 0$. Take an arbitrary $h \in S$ such that inequality (3.1) holds. Then for sufficiently small $\varepsilon > 0$, the inequality

$$k(U(x)) < -\frac{U(x+\varepsilon h) - U(x)}{\varepsilon}$$

holds, and therefore (2.1) holds for $x' := x + \varepsilon h$. ∎

Corollary 3.2. *Let the function U be Gâteaux differentiable. Assume that there exists a function $k : (0, +\infty) \to (0, +\infty)$ such that for all $x \in X$ for which $U(x) > \gamma$ the inequality*

$$k(U(x) - \gamma) < \|\nabla U(x)\|$$

holds (here ∇U is the Gâteaux derivative of the function U). Then U satisfies condition (2.1) *with the function k and $m = +\infty$.*

Let us apply Theorem 2.2 to functions that are differentiable in some directions.

Proposition 3.3. *Let us assume that the function U does not attain its infimum. Then for every function $k \in \mathcal{K}$ there exists a minimizing sequence $\{x_n\} \subset X$ such that $\{U(x_n)\}$ monotonically decreases and for every n, for every direction $h \in S$ in which the function U is differentiable at the point x_n, the following inequality holds:*

$$-U'(x_n;h) \le k(U(x_n) - \gamma) \quad \forall\, n. \tag{3.2}$$

Proof. Assume again that $\gamma = 0$. Take an arbitrary function $k \in \mathcal{K}$. According to Theorem 2.2, there exists a minimizing sequence $\{x_n\} \subset X$ such that $\{U(x_n)\}$ monotonically decreases and

$$U(x') + k(U(x_n))\|x_n - x'\| - U(x_n) > 0 \quad \forall\, x' \ne x, \quad \forall\, n.$$

Then for each direction $h \in S$ in which the function U is differentiable at the point x_n, for all $\varepsilon > 0$ the inequality $U(x_n + \varepsilon h) + \varepsilon k(U(x_n)) - U(x_n) > 0$ holds. Dividing the last inequality by $\varepsilon > 0$ and passing to the limit as $\varepsilon \to 0+$, we obtain the desired inequality (3.2). ∎

Consider the following important corollary of Proposition 3.3.

Proposition 3.4. *Let the function U be Gâteaux differentiable. Then for every function $k \in \mathcal{K}$ there exists a minimizing sequence $\{x_n\} \subset X$ such that $\{U(x_n)\}$ monotonically decreases, and for every n the following inequality holds:*

$$\|\nabla U(x_n)\| \le k(U(x_n) - \gamma) \quad \forall\, n. \tag{3.3}$$

Proof. If the function U attains its minimum at a point $\bar{x} \in X$, then we can take the stationary sequence $x_n = \bar{x}$ as the minimizing sequence. If the function U does not attain its minimum, then this proposition follows from Proposition 3.3, since for the Gâteaux differentiable function U, if inequality (3.2) holds for every $h \in S$, then (3.3) holds. ◻

Proposition 3.4 strengthens the well-known theorem on minimizing sequences of Gâteaux differentiable functions (see [9, Chapter 5, section 3, Corollary 7]). This theorem guarantees the existence of the minimizing sequence $\{x_n\} \subset X$ for which $\{\nabla U(x_n)\}$ tends to zero. At the same time, Proposition 3.4 guarantees that for every function $k \in \mathcal{K}$ such that $k(\tau) \to 0$ as $\tau \to 0$, there exists a minimizing sequence $\{x_n\} \subset X$ such that not only $\{\nabla U(x_n)\}$ tends to zero, but the convergence estimate (3.3) holds.

In Theorem 2.2 and Corollary 2.3, as the function $k \in \mathcal{K}$ we can take the function $k(\tau) = a\tau^p$ with $p < 1$, $a > 0$. At the same time, when $p = 1$ we cannot take this function, since if condition (2.1) holds for the function $k(\tau) \equiv a\tau$ (which does not belong to $\mathcal{K}$), then the function U may have no points of minimum. An appropriate example is provided by the function $U(x) = e^{\alpha x}$, $x \in \mathbb{R}$, where $\alpha \neq 0$. For this function condition (2.1) holds for $k(\tau) \equiv a\tau$ with any positive $a < |\alpha|$ and $m = +\infty$ (see Corollary 3.2); however, the function U does not have points of minimum.

In Propositions 3.3 and 3.4 we also can take the function $k(\tau) = a\tau^p$ with $p < 1$, $a > 0$. At the same time we cannot take $k(\tau) \equiv a\tau$. An appropriate example is provided by the function $U(x) = e^{-x^2}$, $x \in \mathbb{R}$. For this function, relation (3.3) takes the form $|2x_n|e^{-x_n^2} \leq ae^{-x_n^2}$, and, therefore, any sequence $\{x_n\}$ satisfying (3.3) is bounded. Therefore, such a sequence $\{x_n\}$ is not minimizing.

4. Criteria for the generalized Caristi-like condition. In this section, we study the existence of the function $k \in \mathcal{K}$ corresponding to the function U by virtue of the generalized CLC (2.1). These conditions will be obtained in terms of the given function U. Everywhere in this section we assume that $\gamma = \inf_{x \in X} U(x) = 0$.

Given a function $U : X \to \mathbb{R} \cup \{+\infty\}$ and a number m that is either positive or $m = +\infty$, define the function $K_U : (0, m) \to \mathbb{R}_+ \cup \{+\infty\}$ as follows:

$$(4.1) \qquad \begin{cases} K_U(\tau) = \inf\limits_{x:\, U(x) = \tau} \left(\sup\limits_{x':\, x' \neq x} \dfrac{U(x) - U(x')}{\rho(x, x')} \right) & \forall\, \tau \in U(X), \\ K_U(\tau) = +\infty \quad \forall\, \tau \notin U(X). \end{cases}$$

Denote the envelope of the function K_U by $\underline{K_U}\!\uparrow$, i.e.,

$$\underline{K_U}\!\uparrow\!(\tau) := \inf_{s \geq \tau} K_U(s).$$

It is obvious that $\underline{K_U}\!\uparrow$ increases on $(0, m)$, $\underline{K_U}\!\uparrow\!(\tau) \leq K_U(\tau)$ for every τ, and $\underline{K_U}\!\uparrow$ is the largest increasing function that does not exceed K_U.

Assume that U satisfies the generalized CLC (2.1) with the constant m and a function $k \in \mathcal{K}$.

Let us show that the function $\underline{K_U}\!\uparrow$ is a pointwise upper bound of all increasing functions k satisfying (2.1), i.e., $\underline{K_U}\!\uparrow\!(\tau) \geq k(\tau)$ for every increasing function k satisfying (2.1) for every τ. Indeed, take an increasing function k satisfying (2.1).

At first, note that $k(\tau) \leq K_U(\tau)$ for every τ. Indeed, for $\tau \notin U(X)$ this is obvious, whereas for every $\tau \in U(X)$, for every $x \in X$ such that $U(x) = \tau$, by virtue of (2.1) we have

$$k(\tau) = k(U(x)) \leq \frac{U(x) - U(x')}{\rho(x, x')} \leq \sup_{x' \colon x' \neq x} \frac{U(x) - U(x')}{\rho(x, x')},$$

and hence $k(\tau) \leq K_U(\tau)$ by the definition of K_U.

Now we show that $k(\tau) \leq \underline{K_U}\uparrow(\tau)$ for every τ. Indeed, let $k(\tau) > \underline{K_U}\uparrow(\tau)$. The definition of the function $\underline{K_U}\uparrow$ implies that $k(\tau) > \inf_{s \geq \tau} K_U(s)$. Moreover, $k(\tau) \leq K_U(\tau)$. Thus, there exists $s > \tau$ such that $k(\tau) > K_U(s)$. Hence, since k increases we have $k(s) > K_U(s)$, which is impossible. So, $\underline{K_U}\uparrow$ is a pointwise upper bound of all increasing functions k satisfying (2.1).

This reasoning implies the following proposition.

PROPOSITION 4.1. *There exists a function $k \in \mathcal{K}$ that satisfies condition* (2.1) *if and only if the function $1/\underline{K_U}\uparrow$ is summable in a neighborhood of zero.*

Proof. Assume that there exists a function $k \in \mathcal{K}$ that satisfies condition (2.1). Then, since $k(\tau) \leq \underline{K_U}\uparrow(\tau)$ for every τ, the function $1/\underline{K_U}\uparrow$ is summable in a neighborhood of zero.

Assume now that the function $1/\underline{K_U}\uparrow$ is summable in a neighborhood of zero. Consider two cases. At first, assume that $\underline{K_U}\uparrow$ is bounded in a neighborhood of zero. Without loss of generality, reducing if necessary the value of finite $m > 0$ we shall assume that $\underline{K_U}\uparrow$ is bounded on $(0, m]$. Take arbitrary $\lambda \in (0, 1)$. Set $k(\tau) := \lambda \underline{K_U}\uparrow(\tau)$, $\tau \in (0, m]$. Since $1/\underline{K_U}\uparrow$ is summable, we have $\underline{K_U}\uparrow(\tau) \neq 0$ for every $\tau > 0$. So, $k(\tau) < \underline{K_U}\uparrow(\tau)$ for every $\tau > 0$. Therefore, for every $x \in X$ such that $0 < U(x) \leq m$, by virtue of the definitions of K_U and $\underline{K_U}\uparrow$ we obtain

$$k(U(x)) < \underline{K_U}\uparrow(U(x)) \leq K_U(U(x)) \leq \sup_{x' \colon x' \neq x} \frac{U(x) - U(x')}{\rho(x, x')}.$$

Hence, there exists $x' \in X$ satisfying the inequality in (2.1). Thus, k satisfies condition (2.1). Moreover, $1/k$ is summable in a neighborhood of zero and is increasing since $1/\underline{K_U}\uparrow$ is summable in a neighborhood of zero and is increasing. Thus, $k \in \mathcal{K}$.

Consider now the case in which $\underline{K_U}\uparrow(\tau) = +\infty$ for every $\tau > 0$. Set $k(\tau) := 1$. Obviously, $k \in \mathcal{K}$. Moreover, k satisfies condition (2.1), since for every $x \in X$ such that $0 < U(x) \leq m$, by virtue of the definitions of K_U and $\underline{K_U}\uparrow$ we have

$$k(U(x)) = 1 < \underline{K_U}\uparrow(U(x)) \leq K_U(U(x)) \leq \sup_{x' \colon x' \neq x} \frac{U(x) - U(x')}{\rho(x, x')}.$$

Hence, there exists $x' \in X$ satisfying the inequality in (2.1). ◻

Remark 4.2. Let us note that if $\underline{K_U}\uparrow$ is bounded in a neighborhood of zero, then $\underline{K_U}\uparrow$ is a pointwise supremum of all increasing functions k satisfying (2.1), i.e., for every τ in a neighborhood of zero, the value $\underline{K_U}\uparrow(\tau)$ is the supremum of all $k(\tau)$ such that k is an increasing function satisfying (2.1). Without loss of generality, reducing if necessary the value of finite $m > 0$ we shall assume that $\underline{K_U}\uparrow$ is bounded on $(0, m]$. It is a straightforward task to ensure that for arbitrary $\lambda \in (0, 1)$ the function $k := \lambda \underline{K_U}\uparrow$ satisfies relation (2.1). Thus, $\underline{K_U}\uparrow$ is the pointwise supremum of all increasing functions k satisfying (2.1).

If $\underline{K_U}\uparrow(\tau) = +\infty$ for every $\tau > 0$ in a neighborhood of zero, then this proposition is obvious.

Define a function $\widetilde{K}_U : (0,m) \to \mathbb{R}_+ \cup \{+\infty\}$ as follows:

$$\widetilde{K}_U(\tau) = \inf_{x:\,U(x)=\tau} \left(\limsup_{x'\to x} \frac{U(x)-U(x')}{\rho(x,x')} \right)$$
$$\forall\, \tau \in U(X), \quad \widetilde{K}_U(\tau) = +\infty \quad \forall\, \tau \notin U(X).$$

It is obvious that $\widetilde{K}_U(\tau) \le K_U(\tau)$ for all τ. Therefore, Proposition 4.1 implies the following assertion.

PROPOSITION 4.3. *If the function* $1/\underline{\widetilde{K}_U}\!\uparrow$ *is summable in a neighborhood of zero then there exists a function* $k \in \mathcal{K}$ *such that* (2.1) *holds.*

For the rest of this section we put $m = +\infty$.

Before discussing the obtained propositions, compare the boundaries of applicability of Theorems 1.1 and 2.2 to the class of positively homogenous functions.

Let the function $U : \mathbb{R}^n \to \mathbb{R}$ be nonegative, continuous, Gâteaux differentiable on $\mathbb{R}^n \setminus \{0\}$, and positively homogenous of degree $d \ge 1$, i.e.,

$$U(\tau x) \equiv \tau^d U(x), \quad x \in \mathbb{R}^n, \quad \tau > 0.$$

It is obvious that U attains its minimum at zero. However, Theorem 1.1 is applicable to this function only when $d = 1$. At the same time, Theorem 2.2 is applicable even in the $d > 1$ case. Namely, the following proposition holds.

Let

$$S := \{x \in \mathbb{R}^n : |x| = 1\}, \;\; \eta := \min_{x\in S} |U'(x)|, \;\; \mu := \max_{x\in S} |U(x)|, \;\; \widetilde{\eta} := \inf_{x:\,U(x)=1} |U'(x)|.$$

PROPOSITION 4.4. *For a function* U, *the relations*

$$\widetilde{K}_U(\tau) \equiv \widetilde{\eta}\tau^{\frac{d-1}{d}}, \quad \widetilde{K}_U(\tau) \ge k(\tau) := \eta\mu^{\frac{1-d}{d}}\tau^{\frac{d-1}{d}} \quad \forall\, \tau > 0 \tag{4.2}$$

hold. These relations imply that condition (2.1) *holds for* U.

Proof. Since the function U is positively homogenous, $U'(\lambda x) \equiv \lambda^{d-1}U'(x)$ for every $\lambda > 0$, for every $x \in \mathbb{R}^n$. Therefore,

$$\begin{aligned}\widetilde{K}_U(\tau) = \inf_{x:\,U(x)=\tau} |U'(x)| &= \inf\{|U'(x)| : U(\tau^{-\frac{1}{d}}x) = 1, \ x \in X\}\\ &= \inf\{|U'(\tau^{\frac{1}{d}}\xi)| : U(\xi) = 1, \ \xi \in X\}\\ &= \tau^{\frac{d-1}{d}} \inf\{|U'(\xi)| : U(\xi) = 1, \ \xi \in X\} = \widetilde{\eta}\tau^{\frac{d-1}{d}}\end{aligned}$$

for $\tau > 0$. The inequality $\widetilde{K}_U(\tau) \ge k(\tau)$ follows from the relations

$$\begin{aligned}\widetilde{\eta} &= \inf\{|U'(\lambda e)| : U(\lambda e) = 1, \ \lambda > 0, \ e \in S\}\\ &= \inf\{\lambda^{d-1}|U'(e)| : \lambda^d U(e) = 1, \ \lambda > 0, \ e \in S\}\\ &= \inf_{e\in S,\,U(e)\ne 0} |U'(e)|\,U(e)^{\frac{1-d}{d}} \ge \eta\mu^{\frac{1-d}{d}}.\end{aligned}$$

Since the functions $\widetilde{K}_U$ and k are increasing, Proposition 4.3 implies that U satisfies the condition (2.1) with $m = +\infty$ and $\gamma = 0$. ∎

Let us show that in (4.2) the inequality $\widetilde{K}_U(\tau) \geq k(\tau)$ cannot be replaced by the equality. Indeed, consider the function $U : \mathbb{R}^2 \to \mathbb{R}$, $U(x) := x_1^2 + ax_2^2$, $x \in \mathbb{R}^2$, where $a > 1$ is given. Direct computations show that $\mu = a$, $\eta = 2$, $\widetilde{\eta} = 2$, and thus $\widetilde{K}_U(1) > k(1)$.

Let us compute K_U and $\widetilde{K}_U$ for the function $U : \mathbb{R} \to \mathbb{R}$, $U(x) \equiv e^{-x}$. It is easy to see that $K_U(\tau) \equiv \widetilde{K}_U \equiv \tau$. It is obvious that $K_U \notin \mathcal{K}$.

Let us discuss Propositions 4.1 and 4.3. Proposition 4.1 implies that if the function $1/\underline{K_U}\uparrow$ is summable then there exists a function k that satisfies the assumptions of Theorem 2.2. At the same time, if $1/K_U$ is summable but the function K_U is not increasing, then the function U may not attain its minimum, so the function $k \in \mathcal{K}$ mentioned in Proposition 4.1 may not exist. Therefore, the assumption that the function k is increasing in Theorem 2.2 is essential. Let us illustrate the aforementioned with the following example.

Example 4.5. Take an arbitrary decreasing sequence of positive numbers a_j, $j = 0, 1, 2, \dots$, for which $\sum_{j=0}^{\infty} a_j = 1$. Define the function $U : \mathbb{R}_+ \to \mathbb{R}$ in the following way. For every half-open interval $[n, n+1)$, $n = 0, 1, 2, \dots$, consider two half-open intervals $[n, n+a_n)$ and $[n+a_n, n+1)$, define U as a linear function with the angular coefficient -1 on the first half-open interval, and define U on the second half-open interval as a constant so that $U(0) = 1$ and the function U is continuous, i.e.,

$$U(x) = \begin{cases} -x + n + \displaystyle\sum_{j=n}^{\infty} a_j & \text{if} \quad x \in [n, n+a_n), \\ \displaystyle\sum_{j=n+1}^{\infty} a_j & \text{if} \quad x \in [n+a_n, n+1), \end{cases} \qquad \forall\, x \in \mathbb{R}_+, \quad \forall\, n = 0, 1, 2, \dots .$$

This function is decreasing, bounded below, and does not attain its infimum, which is equal to zero. For all $n = 1, 2, \dots$ the preimage of the point $\tau_n := \sum_{j=n}^{\infty} a_j$ is the half-open interval $[n-1+a_{n-1}, n)$, the preimage of the point $\tau \in (\tau_n, \tau_{n-1})$ is a certain point $x \in (n, n+a_n)$, and the preimage of the point τ_0 is zero.

Let us compute the function K_U, defined by formula (4.1). Take an arbitrary positive integer n. It is a straightforward task to ensure that if $x \in [n, n+a_n)$ and $x' > x$, then

$$U(x') \geq -x' + n + \sum_{j=n}^{\infty} a_j,$$

and if $x \in [n+a_n, n+1)$ and $x' > x$, then

$$U(x') \geq \frac{a_{n+1}}{x - n - 1 - a_{n+1}}(x' - x) + \sum_{j=n+1}^{+\infty} a_j.$$

The first relation implies that for all $x \in [n, n+a_n)$, for all $x' > x$ we have

$$\frac{U(x) - U(x')}{\rho(x, x')} \leq \frac{(-x + n + \sum_{j=n}^{\infty} a_j) - (-x' + n + \sum_{j=n}^{\infty} a_j)}{x' - x} = 1,$$

whereas the definition of U implies

$$\frac{U(x) - U(x')}{\rho(x, x')} = \frac{(-x + n + \sum_{j=n}^{\infty} a_j) - (-x' + n + \sum_{j=n}^{\infty} a_j)}{x' - x} = 1$$

if, in addition, $x' \in (x, n + a_n)$. Hence,

$$\sup_{x' : x' \neq x} \frac{U(x) - U(x')}{\rho(x, x')} = 1.$$

The second relation implies that for all $x \in [n + a_n, n + 1)$, for all $x' > x$ we have

$$\frac{U(x) - U(x')}{\rho(x, x')} \leq \frac{a_{n+1}}{n + 1 + a_{n+1} - x},$$

whereas the definition of U implies

$$\frac{U(x) - U(x')}{\rho(x, x')} = \frac{a_{n+1}}{n + 1 + a_{n+1} - x} \quad \text{for} \quad x' = n + 1 + a_{n+1}.$$

Hence,

$$\sup_{x' : x' \neq x} \frac{U(x) - U(x')}{\rho(x, x')} = \frac{a_{n+1}}{n + 1 + a_{n+1} - x}.$$

Take an arbitrary positive integer n and arbitrary $\tau \in (\tau_n, \tau_{n-1})$. The point τ has the only preimage $x \in (n, n + a_n)$. Therefore,

$$K_U(\tau) = \sup_{x' : x' \neq x} \frac{U(x) - U(x')}{\rho(x, x')} = 1.$$

The preimage of the point τ_n is the half-open interval $[n - 1 + a_{n-1}, n)$. Therefore,

$$\begin{aligned} K_U(\tau) &= \inf_{x \in [n-1+a_{n-1}, n)} \left(\sup_{x' : x' \neq x} \frac{U(x) - U(x')}{\rho(x, x')} \right) \\ &= \inf_{x \in [n-1+a_{n-1}, n)} \left(\frac{a_n}{n + a_n - x} \right) = \frac{a_n}{1 + a_n - a_{n-1}}. \end{aligned}$$

Therefore, the value of $K_U(\tau)$ is equal to 1 for every $\tau \in (0, 1]$ except for the points τ_n, $n = 1, 2, \ldots$, at which it is equal to $\frac{a_n}{1+a_n-a_{n-1}}$. It is obvious that the function K_U is summable; however, it is not increasing. At the same time, the function U does not attain its minimum. Therefore, for the constructed function U, by virtue of Theorem 2.2 there exists no function $k \in \mathcal{K}$ for which the generalized CLC (2.1) holds. Moreover, $\underline{K_U} \uparrow (\tau) = 0$ for all $\tau \in (0, 1)$. ∎

In this example, the function K_U is not increasing. At the same time, there exists a function U for which there exists a function $k \in \mathcal{K}$ such that the generalized CLC (2.1) holds; however, the function K_U defined by formula (4.1) is not increasing. Consider the corresponding example.

Example 4.6. Take an arbitrary decreasing sequence of positive numbers τ_n, $n = 0, 1, 2, \ldots$, $\tau_0 = 1$. Define the function $U : \mathbb{R}_+ \to \mathbb{R}$ as follows. For every half-open interval $(\tau_{n+1}, \tau_n]$, $n = 0, 1, 2, \ldots$, consider two half-open intervals $(\tau_{n+1}, (\tau_n + \tau_{n+1})/2]$ and $((\tau_n + \tau_{n+1})/2, \tau_n]$, put $U(x) \equiv \tau_{n+1}$ on the first interval, and define U as a linear function with angular coefficient 2 on the second interval so that $U(0) = 0$ and the function U is continuous, i.e.,

$$U(x) = \begin{cases} \tau_{n+1} & \text{if} \quad x \in (\tau_{n+1}, (\tau_n + \tau_{n+1})/2], \\ 2x - \tau_n & \text{if} \quad x \in ((\tau_n + \tau_{n+1})/2, \tau_n], \end{cases} \qquad \forall\, x \in \mathbb{R}_+, \quad \forall\, n = 0, 1, 2, \ldots,$$

$$U(0) = 0.$$

This function is increasing and zero is the only minimum point of U.

Let us show that $K_U(\tau) = 2$ for each $\tau \notin \{\tau_1, \tau_2, \tau_3, \ldots\}$. Since U is absolutely continuous and $0 \leq U'(x) \leq 2$ for almost all x, the Newton–Leibniz formula implies that $U(x) - U(x') \leq 2(x - x')$ for any x, for any $x' < x$. Hence, $K_U(\tau) \leq 2$ for every $\tau > 0$. Moreover, it is obvious that $U'(x) = 2$ for each x such that $U(x) \notin \{\tau_1, \tau_2, \tau_3, \ldots\}$, and hence $K_U(\tau) \geq 2$ for each $\tau \notin \{\tau_1, \tau_2, \tau_3, \ldots\}$.

Let us show that the function K_U is not increasing in every neighborhood of zero. Indeed, if K_U is increasing in a neighborhood of zero, then $K_U(\tau) \geq 2$ in this neighborhood by virtue of the monotonicity of K_U. Then, Theorem 2.2 implies that $\rho(\tau_n, 0) \leq U(\tau_n)/2 = \tau_n/2$ for all sufficiently large n, which is impossible. Hence, K_U is not increasing in every neighborhood of zero.

In particular, when $\tau_n = 2^{-n}$ for every n the assumption of Theorem 2.2 holds with $k(\tau) \equiv 2/3$. ∎

By virtue of Theorem 2.2, the generalized CLC is a sufficient condition for the function U to attain its minimum. However, Example 2.4 shows that the generalized CLC is not a necessary minimality condition. Another simple example is provided by an arbitrary function defined on a complete metric space for which there exists an unbounded minimizing sequence and the set of the minimum points is bounded. An example of such a function is $U : \mathbb{R}_+ \to \mathbb{R}$, $U(x) = xe^{-x}$. Example 2.4 shows that the generalized CLC in Theorem 2.2 is not a necessary minimality condition even when X is compact.

5. Generalizations of the Bishop–Phelps and Ekeland variational principles. In this section we deduce the generalizations of the EVP and the BPVP from Theorem 2.2. First, let us recall the BPVP (see [25] and section 2.2 of [18]).

Throughout this section, we assume that (X, ρ) is a complete metric space, $U : X \to \mathbb{R} \cup \{+\infty\}$ is a bounded below lower semicontinuous function, and

$$\gamma := \inf_{x \in X} U(x).$$

BPVP. *For all $c > 0$, for all $x_0 \in \mathrm{dom} U$ there exists a point $\bar{x} \in X$ such that*

$$U(\bar{x}) + c\rho(x_0, \bar{x}) \leq U(x_0), \tag{5.1}$$

$$U(x) + c\rho(x, \bar{x}) > U(\bar{x}) \quad \forall\, x \neq \bar{x}. \tag{5.2}$$

As was noted, the BPVP is equivalent to Theorem 1.1, whereas Theorem 2.2 generalizes Theorem 1.1. Therefore, a generalization of the BPVP can be deduced from Theorem 2.2. Let us do this.

Theorem 5.1 (generalized BPVP). *For any function $k \in \mathcal{K}$, for all $x_0 \in X$ there exists a point $\bar{x} \in X$ such that*

$$\rho(x_0, \bar{x}) \leq D_k(U(x_0) - \gamma) - D_k(U(\bar{x}) - \gamma) = \int_{U(\bar{x})-\gamma}^{U(x_0)-\gamma} \frac{d\tau}{k(\tau)}, \tag{5.3}$$

$$\text{and if} \quad U(\bar{x}) > \gamma, \quad \text{then} \quad U(x) + k(U(\bar{x}) - \gamma)\rho(x, \bar{x}) > U(\bar{x}) \quad \forall\, x \neq \bar{x}. \tag{5.4}$$

Proof. Assume again that $\gamma = 0$. Set $D := D_k$,

$$\overline{X} := \big\{x \in X : \ \rho(x_0, x) \leq D(U(x_0)) - D(U(x))\big\}.$$

The set $\overline{X}$ is nonempty, since it contains x_0, and it is closed, since the function U is lower semicontinuous. The function D is continuous since the function $1/k$ is summable.

In order to prove the theorem, consider the contrary. Assume that for all $\bar{x} \in \overline{X}$ the inequality $U(\bar{x}) > 0$ holds and the relation (5.4) does not hold. Then for all $x \in \overline{X}$, we have $U(x) > 0$ and there exists a point $x' \in X$, $x' \neq x$, such that the inequality in (2.1) holds with $m = +\infty$. Then $x' \in \overline{X}$, since

$$\begin{aligned}\rho(x_0, x') \leq \rho(x_0, x) + \rho(x, x') &\leq D(U(x_0)) - D(U(x)) + \frac{U(x) - U(x')}{k(U(x))} \\ &\leq D(U(x_0)) - D(U(x)) + \int_{U(x')}^{U(x)} \frac{d\tau}{k(\tau)} \\ &= \int_{U(x)}^{U(x_0)} \frac{d\tau}{k(\tau)} + \int_{U(x')}^{U(x)} \frac{d\tau}{k(\tau)} \\ &= \int_{U(x')}^{U(x_0)} \frac{d\tau}{k(\tau)}.\end{aligned}$$

Therefore, by virtue of Theorem 2.2, there exists $\bar{x} \in \overline{X}$ for which the equality $U(\bar{x}) = 0$ holds. This contradicts the assumption that $U(x) > 0$ for all $x \in \overline{X}$. ∎

The BPVP follows from the generalized BPVP for the constant function $k(\tau) \equiv c$. So, it is natural to call Theorem 5.1 the generalized BPVP.

In the next section, we show that the generalized BPVP is a significant strengthening of the BPVP. Precisely, we will prove that for all sufficiently small $c > 0$ there exist functions $k \in \mathcal{K}$ such that the estimates (5.3) and (5.4) are significantly better than the estimates (5.1) and (5.2).

Let us turn to the EVP (which was formulated in the introduction) and deduce a generalization of the EVP from Theorem 2.2.

Theorem 5.2 (generalized EVP). *Let $k \in \mathcal{K}$. For all $x_0 \in X$, $\varepsilon > 0$ such that $U(x_0) \leq \gamma + \varepsilon$, for all $\lambda > 0$ there exists a point $\bar{x} \in X$ such that* (1.1) *holds and*

$$\text{if} \quad U(\bar{x}) > \gamma, \quad \text{then}$$
$$U(x) + \lambda^{-1}\left(k(U(\bar{x}) - \gamma) \int_0^{\varepsilon} \frac{d\tau}{k(\tau)}\right)\rho(x, \bar{x}) > U(\bar{x}) \quad \forall\, x \neq \bar{x}. \tag{5.5}$$

Proof. Assume again that $\gamma = 0$. Take an arbitrary function $k \in \mathcal{K}$ and $\lambda > 0$. Set $c := \frac{1}{\lambda}\int_0^{\varepsilon} \frac{d\tau}{k(\tau)}$. Applying the generalized BPVP to the functions U and ck, we obtain that there exists a point $\bar{x} \in X$ such that

$$\rho(x_0, \bar{x}) \leq \int_{U(\bar{x})}^{U(x_0)} \frac{d\tau}{ck(\tau)},$$
$$\text{and if} \quad U(\bar{x}) > 0, \quad \text{then} \quad U(x) + ck(U(\bar{x}))\rho(x, \bar{x}) > U(\bar{x}) \quad \forall\, x \neq \bar{x}.$$

The first inequality implies $U(\bar{x}) \leq U(x_0)$ and $\rho(x_0, \bar{x}) \leq \int_0^{\varepsilon} \frac{d\tau}{ck(\tau)} = \lambda$. The second inequality implies

$$U(x) + \left(\frac{k(U(\bar{x}))}{\lambda} \int_0^{\varepsilon} \frac{d\tau}{k(\tau)}\right)\rho(x, \bar{x}) = U(x) + ck(U(\bar{x}))\rho(x, \bar{x}) > U(\bar{x}) \quad \forall\, x \neq \bar{x}. \quad ∎$$

The EVP follows from the generalized EVP for the constant function $k(t) \equiv 1$. For this reason, Theorem 5.2 is called the generalized EVP.

Let us compare the generalized BPVP and the generalized EVP. Let $k \in \mathcal{K}$, $x_0 \in X$, and let positive numbers ε, λ, c be given. Denote by $\mathcal{GBP}(x_0; k)$ the set of points $\bar{x} \in X$ corresponding to the generalized BPVP, that is, those for which relations (5.3) and (5.4) hold. By $\mathcal{GE}(x_0; k; \varepsilon, \lambda)$ we will denote the set of points $\bar{x} \in X$ corresponding to the generalized EVP, that is, those for which relations (1.1) and (5.5) hold.

Let us show that when $\lambda = \int_0^\varepsilon \frac{d\tau}{k(\tau)}$, the following inclusion holds:

$$\mathcal{GBP}(x_0; k) \subset \mathcal{GE}(x_0; k; \varepsilon, \lambda). \tag{5.6}$$

Take an arbitrary point $\bar{x} \in \mathcal{GBP}(x_0; k)$. Inequality (5.3) implies that (1.1) holds for the point $\bar{x}$. Moreover, since $\frac{k(U(\bar{x}))}{\lambda} \int_0^\varepsilon \frac{d\tau}{k(\tau)} = k(U(\bar{x}))$, relation (5.4) implies (5.5). Therefore, $\bar{x} \in \mathcal{GE}(x_0; k; \varepsilon, \lambda)$, which proves inclusion (5.6).

Define

$$\overline{X}(x_0; k) := \big\{x \in X : \rho(x_0, x) \le D_k(U(x_0)) - D_k(U(x))\big\}.$$

The above-mentioned arguments prove the validity of the equality

$$\mathcal{GBP}(x_0; k) = \mathcal{GE}(x_0; k; \varepsilon, \lambda) \cap \overline{X}(x_0).$$

Note that inclusion (5.6) can be strict. A simple example of the aforementioned is the function $U(x) = \min\{1, e^x\}$, $x \in \mathbb{R}$. It is a straightforward task to ensure that if $\lambda > 1$, $k(\tau) \equiv \lambda$, then the relations $x \in \mathcal{GE}(0; k; 1, \lambda)$, $x \notin \mathcal{GBP}(0; k)$ hold for all $x \in (0, \lambda]$.

Let us compare the BPVP and the EVP. Denote by $\mathcal{BP}(x_0; c)$, the set of points $\bar{x} \in X$ corresponding to the BPVP, i.e., those for which relations (5.1) and (5.2) hold. Denote by $\mathcal{E}(x_0; \varepsilon, \lambda)$ the set of points $\bar{x} \in X$ corresponding to the EVP, i.e., those for which relations (1.1) and (1.2) hold.

Since the BPVP coincides with the generalized BPVP for $k(\tau) \equiv c$ and the EVP coincides with the generalized EVP for $k(\tau) \equiv 1$, the relations

$$\mathcal{BP}(x_0; c) = \mathcal{GBP}(x_0; c), \quad \overline{X}(x_0; c) = \left\{x \in X : \rho(x_0, x) \le \frac{U(x) - U(x_0)}{c}\right\},$$
$$\mathcal{E}(x_0; \varepsilon, \lambda) = \mathcal{GE}(x_0; 1; \varepsilon, \lambda)$$

hold. Therefore,

$$\mathcal{BP}\left(x_0; \frac{\varepsilon}{\lambda}\right) = \mathcal{E}(x_0; \varepsilon, \lambda) \cap \overline{X}\left(x_0; \frac{\varepsilon}{\lambda}\right),$$

and, in particular,

$$\mathcal{BP}\left(x_0; \frac{\varepsilon}{\lambda}\right) \subset \mathcal{E}(x_0; \varepsilon, \lambda), \tag{5.7}$$

wherein the inclusion (5.7), similar to (5.6), can be strict. The corresponding example is presented above in the discussion of the inclusion (5.6).

Relation (5.6) shows that on the corresponding choice of λ, the set of points $\mathcal{GBP}(x_0; k)$, corresponding to the generalized BPVP, is contained in the set of points $\mathcal{GE}(x_0; k; \varepsilon, \lambda)$, corresponding to the generalized EVP. So, the generalized EVP is weaker than the generalized BPVP. Similarly, relation (5.7) shows that same holds true for the BPVP and the EVP.

6. Comparison of the variational principles of Bishop–Phelps and Ekeland with their generalized analogues. Let (X, ρ) be a complete metric space and let $U : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function bounded from below. For simplicity we assume in this section that $\gamma = \inf_{x \in X} U(x) = 0$.

6.1. Comparison of the BPVP and the generalized BPVP. Let us compare the BPVP and the generalized BPVP. At first, we obtain two auxiliary propositions.

LEMMA 6.1. *Let $k \in \mathcal{K}$, $x_0 \in X$, and $U(x_0) > 0$. Then for all $\varepsilon > 0$ there exists $\bar{\mu} > 0$ such that if $0 < \mu \le \bar{\mu}$, then for every point $\bar{x} = \bar{x}(\mu) \in X$ for which*

$$\rho(x_0, \bar{x}) \le \int_{U(\bar{x})}^{U(x_0)} \frac{d\tau}{\mu k(\tau)}, \tag{6.1}$$

$$\text{if} \quad U(\bar{x}) > 0, \quad \text{then} \quad U(x) + \mu k(U(\bar{x}))\rho(x, \bar{x}) > U(\bar{x}) \quad \forall\, x \neq \bar{x}, \tag{6.2}$$

and the inequality

$$U(\bar{x}) \le \left(\frac{1}{2} + \varepsilon\right) U(x_0) \tag{6.3}$$

holds.

Proof. Take arbitrary $k \in \mathcal{K}$, $x_0 \in X$, $U(x_0) > 0$, $\varepsilon > 0$, $\widehat{x} \in X$ such that $U(\widehat{x}) \le \varepsilon U(x_0)$. Take $\bar{\mu} > 0$ such that

$$\bar{\mu} k(U(x_0))\rho(x_0, \widehat{x}) \le \varepsilon U(x_0).$$

Take an arbitrary $\mu \le \bar{\mu}$. Let $\bar{x} \in X$ satisfy relations (6.1) and (6.2). If $U(\bar{x}) = 0$, then (6.3) obviously holds. Thus, assume that $U(\bar{x}) > 0$. Since $U(\bar{x}) > 0$, we have $k(U(\bar{x})) > 0$. Therefore,

$$\begin{aligned} U(\bar{x}) &\le U(\widehat{x}) + \mu k(U(\bar{x}))\rho(\bar{x}, \widehat{x}) \le \varepsilon U(x_0) + \mu k(U(\bar{x}))\rho(\bar{x}, \widehat{x}) \\ &\le \varepsilon U(x_0) + \mu k(U(\bar{x}))(\rho(\bar{x}, x_0) + \rho(x_0, \widehat{x})) \\ &\le \varepsilon U(x_0) + \mu k(U(\bar{x})) \int_{U(\bar{x})}^{U(x_0)} \frac{d\tau}{\mu k(\tau)} + \mu k(U(\bar{x}))\rho(x_0, \widehat{x}) \\ &\le \varepsilon U(x_0) + k(U(\bar{x})) \frac{U(x_0) - U(\bar{x})}{k(U(\bar{x}))} + \bar{\mu} k(U(x_0))\rho(x_0, \widehat{x}) \\ &\le \varepsilon U(x_0) + U(x_0) - U(\bar{x}) + \varepsilon U(x_0) = (1 + 2\varepsilon)U(x_0) - U(\bar{x}). \end{aligned}$$

Here the first inequality follows from (6.2); the second, from definition of the point $\widehat{x}$; the third, from the triangle inequality; the fourth, from (6.1); the fifth, from the fact that the function k is increasing as well as the fact that $\mu \le \bar{\mu}$; the sixth, from the definition of $\bar{\mu}$. Obviously, (6.3) follows from the obtained inequality. ∎

COROLLARY 6.2. *If the assumptions of Lemma* 6.1 *hold, then there exists $\bar{\mu} > 0$ such that for all positive $\mu \le \bar{\mu}$, for all points $\bar{x} \in X$ satisfying relations* (6.1) *and* (6.2), *the inequality $U(\bar{x}) \le \frac{2}{3} U(x_0)$ holds.*

Taking $k(t) \equiv 1$ in Lemma 6.1 we obtain the following proposition.

COROLLARY 6.3. *Let $x_0 \in X$. There exists $\bar{c} > 0$ such that if $c \le \bar{c}$, then for any point $\bar{x} \in X$ satisfying relations* (5.1) *and* (5.2), *the inequality $U(\bar{x}) \le \frac{2}{3} U(x_0)$ holds.*

LEMMA 6.4. *Let $k \in \mathcal{K}$, $x_0 \in X$, and $U(x_0) > 0$. Then for all $\varepsilon > 0$ there exists $\bar{\mu} > 0$ such that if $\mu \le \bar{\mu}$, then there exists $\bar{x} = \bar{x}(\mu) \in X$ for which relations* (6.1), (6.2), *and inequality $U(\bar{x}) \le \varepsilon U(x_0)$ hold.*

Proof. Take a positive integer n such that $(\frac{2}{3})^{n+1}U(x_0) \le \varepsilon$. Applying Corollary 6.2 and the generalized BPVP to the functions U and k at the point x_0 we obtain that there exist a positive number μ_1 and a point $x_1 \in X$ such that

$$U(x_1) \le \frac{2}{3}U(x_0), \quad \rho(x_0, x_1) \le \int_{U(x_1)}^{U(x_0)} \frac{d\tau}{\mu_1 k(\tau)},$$
$$\text{and if} \quad U(x_1) > 0, \quad \text{then} \quad U(x) + \mu_1 k(U(x_1))\rho(x, x_1) > U(x_1) \ \ \forall\, x \ne x_1.$$

If $U(x_1) > 0$, then applying Corollary 6.2 and the generalized BPVP to the functions U and k at the point x_1 we obtain that there exist a positive number $\mu_2 \le \mu_1$ and a point $x_2 \in X$ such that

$$U(x_2) \le \frac{2}{3}U(x_1), \quad \rho(x_1, x_2) \le \int_{U(x_2)}^{U(x_1)} \frac{d\tau}{\mu_2 k(\tau)},$$
$$\text{and if} \quad U(x_2) > 0, \quad \text{then} \quad U(x) + \mu_2 k(U(x_2))\rho(x, x_2) > U(x_2) \ \ \forall\, x \ne x_2.$$

If $U(x_1) = 0$, then $\bar{\mu} := \mu_1$ and $\bar{x} := x_1$ are the desired values.

Repeating this procedure n times we obtain points $x_1, x_2, \dots, x_n \in X$ and numbers $\mu_1, \mu_2, \dots, \mu_n > 0$ such that the relations

$$\rho(x_{j-1}, x_j) \le \int_{U(x_j)}^{U(x_{j-1})} \frac{d\tau}{\mu_j k(\tau)}, \tag{6.4}$$
$$\text{if} \quad U(x_j) > 0, \quad \text{then} \quad U(x) + \mu_j k(U(x_j))\rho(x, x_j) > U(x_j) \quad \forall\, x \ne x_j,$$

and

$$U(x_j) \le \frac{2}{3}U(x_{j-1}) \tag{6.5}$$

hold for all $j = \overline{1, n}$, and

$$\mu_1 \ge \mu_2 \ge \cdots \ge \mu_n. \tag{6.6}$$

Applying Corollary 6.2 and the generalized BPVP to the functions U and k at the point x_n we obtain the following. There exists a positive number $\bar{\mu} \le \mu_n$ such that for all positive $\mu \le \bar{\mu}$ there exists $\bar{x} \in X$ such that

$$\rho(x_n, \bar{x}) \le \int_{U(\bar{x})}^{U(x_n)} \frac{d\tau}{\mu k(\tau)}, \tag{6.7}$$
$$\text{if} \quad U(\bar{x}) > 0, \quad \text{then} \quad U(x) + \mu k(U(\bar{x}))\rho(x, \bar{x}) > U(\bar{x}) \quad \forall\, x \ne \bar{x}, \tag{6.8}$$
$$U(\bar{x}) \le \frac{2}{3}U(x_n). \tag{6.9}$$

Thus,

$$\rho(x_0, \bar{x}) \le \rho(\bar{x}, x_n) + \sum_{j=1}^{n} \rho(x_{j-1}, x_j) \le \int_{U(\bar{x})}^{U(x_n)} \frac{d\tau}{\mu k(\tau)} + \sum_{j=1}^{n} \int_{U(x_j)}^{U(x_{j-1})} \frac{dt}{\mu_j k(\tau)}$$
$$\le \int_{U(\bar{x})}^{U(x_n)} \frac{d\tau}{\mu k(\tau)} + \sum_{j=1}^{n} \int_{U(x_j)}^{U(x_{j-1})} \frac{d\tau}{\mu k(\tau)} = \int_{U(\bar{x})}^{U(x_0)} \frac{d\tau}{\mu k(\tau)}.$$

Therefore, (6.1) holds. Here the first inequality follows from the triangle inequality; the second, from (6.4); and the third, from (6.6) and the relations $\bar{\mu} \le \mu_n$ and $\mu \le \bar{\mu}$. Further, relation (6.2) coincides with (6.8). Finally, (6.5) and (6.9) imply

$$U(\bar{x}) \le \frac{2}{3}U(x_n) \le \left(\frac{2}{3}\right)^2 U(x_{n-1}) \le \cdots \le \left(\frac{2}{3}\right)^{n+1} U(x_0) \le \varepsilon U(x_0).$$

Here the last inequality holds owing to the choice of n. ∎

PROPOSITION 6.5. *Let $x_0 \in X$, $k \in \mathcal{K}$, $U(x_0) > 0$, and $k(\tau) \to 0$ as $\tau \to 0$. Then for all $\theta \in (0,1)$ there exists $\bar{c} = \bar{c}(\theta) > 0$ such that for all positive $c < \bar{c}$ there exist a number $\mu = \mu(c) > 0$ and a point $\bar{x}_\mu \in X$ for which*

$$\rho(x_0, \bar{x}_\mu) \le \int_{U(\bar{x}_\mu)}^{U(x_0)} \frac{d\tau}{k_\mu(\tau)}, \tag{6.10}$$
$$\text{if} \quad U(\bar{x}_\mu) > 0, \quad \text{then} \quad U(x) + k_\mu(U(\bar{x}_\mu))\rho(x, \bar{x}_\mu) > U(\bar{x}_\mu) \quad \forall\, x \ne \bar{x}_\mu,$$

and for every point $\bar{x}_c$ satisfying relations (5.1) *and* (5.2), *the estimates*

$$\int_0^{U(x_0)} \frac{d\tau}{k_\mu(\tau)} \le \theta \frac{U(x_0) - U(\bar{x}_c)}{c}, \quad k_\mu(U(\bar{x}_\mu)) \le \theta c \tag{6.11}$$

hold. Here $k_\mu(\tau) \equiv \mu k(\tau)$.

Proof. Set

$$I := \int_0^{U(x_0)} \frac{dt}{k(\tau)}, \quad \varepsilon := \frac{\theta^2}{3I}.$$

Since $k(\tau) \to 0$ as $\tau \to 0$, by virtue of Lemma 6.4 there exists $\bar{\mu} > 0$ such that for all positive $\mu \le \bar{\mu}$ there exists a point $\bar{x} = \bar{x}_\mu$ for which relations (6.1), (6.2), and the inequality $k(U(\bar{x}_\mu)) \le \varepsilon U(x_0)$ hold. Corollary 6.3 implies that there exists a positive number $\bar{c} \le \theta\bar{\mu}U(x_0)/3I$ such that for any positive $c \le \bar{c}$, for all $\bar{x}_c$ satisfying relations (5.1) and (5.2), the inequality $U(\bar{x}_c) \le \frac{2}{3}U(x_0)$ holds.

Take an arbitrary positive $c \le \bar{c}$. Set

$$\mu := \frac{3cI}{\theta U(x_0)}.$$

We have

$$\mu = \frac{3cI}{\theta U(x_0)} \le \frac{3\bar{c}I}{\theta U(x_0)} = \bar{\mu}.$$

Therefore, there exists a point $\bar{x}_\mu$ for which relations (6.1), (6.2) hold and $k(U(\bar{x}_\mu)) \le \varepsilon U(x_0)$. In particular, for the point $\bar{x}_\mu$, relations (6.10) hold. The BPVP implies that there exist points $\bar{x}_c$ that satisfy relations (5.1) and (5.2). Take such a point. We have

$$\int_0^{U(x_0)} \frac{d\tau}{\mu k(\tau)} = \frac{I}{\mu} = \theta\frac{U(x_0)}{3c} \le \theta\frac{U(x_0) - U(\bar{x}_c)}{c}.$$

Here the second equality follows from the definition of the number μ, and the inequality follows from the relation $U(\bar{x}_c) \le \frac{2}{3}U(x_0)$. Further, we have

$$\mu k(U(\bar{x}_\mu)) \le \mu\varepsilon U(x_0) = \frac{\mu\theta^2 U(x_0)}{3I} = \theta c.$$

Here the inequality follows from the relation $k(U(\bar{x}_\mu)) \le \varepsilon U(x_0)$; the first equality, from the definition of the number ε; the second, from the definition of μ. Thus, $\bar{x}_\mu$ is the desired point. ∎

Proposition 6.5 shows the following. Let $x_0 \in X$ and a function $k \in \mathcal{K}$ be such that $U(x_0) > 0$ and $k(\tau) \to 0$ as $\tau \to 0$. Then for every sufficiently small $\theta > 0$ there exists $c_0 > 0$ such that for every $c \in (0, c_0]$ and every point $\bar{x}_c$ satisfying the BPVP for the chosen c, there exist $\mu > 0$ and a point $\bar{x}$ corresponding to the function $k_\mu := \mu k$ by virtue of the generalized BPVP (i.e., a point for which the generalized BPVP holds with the function k_μ) such that

$$D_{k_\mu}(U(x_0)) - D_{k_\mu}(U(\bar{x})) \leq \theta \frac{U(x_0) - U(\bar{x})}{c}, \qquad k_\mu(U(\bar{x})) \leq \theta c.$$

In other words, for every sufficiently small c the generalized BPVP provides the estimates that are θ times better than the estimates (5.1), (5.2) that are guaranteed by the BPVP.

6.2. Comparison of the EVP and the generalized EVP. Let us compare the EVP and the generalized EVP. Here the situation is analogous to the one described above for the comparison of the BPVP and the generalized BPVP. Namely, the following proposition holds.

PROPOSITION 6.6. *Let $x_0 \in X$, $k \in \mathcal{K}$, $U(x_0) > 0$, and $k(\tau) \to 0$ as $\tau \to 0$. Then for all $\theta \in (0,1)$ there exists $\bar{c} = \bar{c}(\theta) > 0$ such that, for all $\lambda > 0$ and $\varepsilon > 0$ for which $U(x_0) \leq \varepsilon$ and $\frac{\varepsilon}{\lambda} < \bar{c}$, there exist $\bar{\lambda} > 0$ and a point $\bar{x} \in X$ such that*

$$\begin{gathered} U(\bar{x}) \leq U(x_0), \quad \rho(x_0, \bar{x}) \leq \bar{\lambda}, \\ \textit{if} \quad U(\bar{x}) > 0, \quad \textit{then} \quad U(x) + \frac{k(U(\bar{x}))}{\bar{\lambda}} \int_0^{\bar{\varepsilon}} \frac{d\tau}{k(\tau)} \rho(x, \bar{x}) > U(\bar{x}) \quad \forall\, x \neq \bar{x}, \end{gathered} \tag{6.12}$$

and the following estimates hold:

$$\bar{\lambda} \leq \theta \lambda, \quad \frac{k(U(\bar{x}))}{\bar{\lambda}} \int_0^{\bar{\varepsilon}} \frac{d\tau}{k(\tau)} \leq \theta \frac{\varepsilon}{\lambda}. \tag{6.13}$$

Here $\bar{\varepsilon} = U(x_0)$.

Proof. Take arbitrary $\theta \in (0,1)$, take $\bar{c} = \bar{c}(\theta)$ from Proposition 6.5 and arbitrary $\lambda > 0$ and $\varepsilon > 0$ for which $\frac{\varepsilon}{\lambda} < \bar{c}$. Set $c := \varepsilon/\lambda$. By virtue of the BPVP there exists $\bar{x}_c \in \mathcal{BP}(x_0; c)$. Since $c < \bar{c}$, Proposition 6.5 implies that there exist $\mu > 0$ and a point $\bar{x}_\mu \in \mathcal{GBP}(x_0; k_\mu)$ such that for any point $\bar{x}_c \in \mathcal{BP}(x_0; c)$ inequalities (6.11) hold. Set

$$\bar{x} := \bar{x}_\mu, \quad \bar{\varepsilon} := U(x_0), \quad \bar{\lambda} := \int_0^{\bar{\varepsilon}} \frac{d\tau}{\mu k(\tau)}.$$

Relations (6.11) imply

$$\begin{gathered} \bar{\lambda} = \int_0^{U(x_0)} \frac{d\tau}{\mu k(\tau)} \leq \theta \frac{U(x_0) - U(\bar{x}_c)}{c} \leq \theta \frac{U(x_0)}{c} \leq \theta \frac{\varepsilon}{c} = \theta \lambda, \\ \frac{k(U(\bar{x}))}{\bar{\lambda}} \int_0^{\bar{\varepsilon}} \frac{d\tau}{k(\tau)} = \mu k(U(\bar{x})) \leq \theta c = \theta \frac{\varepsilon}{\lambda}. \end{gathered}$$

It follows from (5.6) that $\bar{x} \in \mathcal{GE}(x_0; k; \bar{\varepsilon}, \bar{\lambda})$. So, (6.12) holds for $\bar{x}$. Thus, the point $\bar{x}$ is the desired one. ∎

Proposition 6.5 shows the following. Let a point $x_0 \in X$ and a function $k \in \mathcal{K}$ such that $k(\tau) \to 0$ as $\tau \to 0$ be given. Then, for every sufficiently small $\theta > 0$, if $\varepsilon \geq U(x_0)$ and $\frac{\varepsilon}{\lambda}$ is sufficiently small, then there exist $\bar{\lambda} > 0$ and a point $\bar{x} \in X$, which corresponds to k, $\bar{\lambda}$, and $\bar{\varepsilon} = U(x_0)$ by virtue of the generalized EVP, such that

$$\bar{\lambda} \leq \theta\lambda, \qquad \frac{k(U(\bar{x}))}{\bar{\lambda}} \int_0^{\bar{\varepsilon}} \frac{d\tau}{k(\tau)} \leq \theta \frac{\varepsilon}{\lambda}.$$

In other words, for every sufficiently small $\frac{\varepsilon}{\lambda}$ the generalized EVP provides estimates that are θ times better than the estimates (1.1), (1.2) that are guaranteed by the EVP.

7. Some applications and generalizations.

7.1. Existence of solutions in extremal problems with constraints. Let $(E, \|\cdot\|)$ be a Banach space and let a function $U : E \to \mathbb{R}$ and a nonempty closed set $X \subset E$ be given. Denote by $S \subset E$ the unit sphere E, i.e., $S = \{x \in E : \|x\| = 1\}$. For all $x \in X$, denote by $T_X(x)$ the contingent cone (Bouligand cone) to the set X at the point x. In other words, $T_X(x)$ is the set of vectors $h \in E$ for each of which there exists a sequence of positive numbers $\{\varepsilon_i\}$ and a sequence $\{o_i\} \subset E$ such that $\varepsilon_i \to 0$, $o_i\varepsilon_i^{-1} \to 0$ as $i \to \infty$, and $x + \varepsilon_i h + o_i \in X\ \forall\, i$.

Assume that the function U is bounded below, lower semicontinuous, and Hadamard differentiable at every point $x \in X$, i.e., U is differentiable at every point $x \in X$ for every direction $h \in S$, and the following equality holds:

$$U'(x; h) = \lim_{\substack{\varepsilon \to 0+ \\ g \to h}} \frac{U(x + \varepsilon g) - U(x)}{\varepsilon} \quad \forall\, h \in E, \quad \forall\, x \in X.$$

Set $\gamma := \inf_{x \in X} U(x)$.

Proposition 7.1. *Assume that there exists a function $k \in \mathcal{K}$ for which the condition*

$$\forall\, x \in X : \quad U(x) > \gamma \quad \exists\, h \in T_X(x) \cap S : \quad U'(x; h) + k(U(x) - \gamma) < 0 \tag{7.1}$$

holds. Then for every $x_0 \in X$ there exists a point $\bar{x} \in X$ at which the minimum of the function U on X is attained and

$$\rho(x_0, \bar{x}) \leq D_k(U(x_0) - \gamma) = \int_0^{U(x_0)-\gamma} \frac{d\tau}{k(\tau)}.$$

Proof. Assume again that $\gamma = 0$. Consider the metric space (X, ρ) with the metric ρ induced from E. Its obvious that (X, ρ) is complete. Take the restriction of the function U to X and prove that it satisfies the generalized CLC (2.1) with the function k.

Take an arbitrary $x \in X$ for which $U(x) > 0$. Take an arbitrary $h \in T_X(x)$ that meets the assumption (7.1). Then there exists a sequence of positive numbers $\{\varepsilon_i\}$ and a sequence $\{o_i\} \subset E$ such that $\varepsilon_i \to 0$, $o_i\varepsilon_i^{-1} \to 0$ as $i \to \infty$, and

$$x_i := x + \varepsilon_i h + o_i \in X \quad \forall\, i.$$

Set $h_i := h + \frac{o_i}{\varepsilon_i}$. Then $h_i \to h$ as $i \to \infty$. Since U is Hadamard differentiable, the inequality in (7.1) implies

$$\frac{U(x_i) - U(x)}{t} + k(U(x)) = \frac{U(x + \varepsilon_i h) - U(x)}{\varepsilon_i} + k(U(x)) < 0$$

for sufficiently large i. Multiplying this inequality by $\varepsilon_i > 0$ and using the fact that $\|h_i\| \to \|h\| = 1$ as $i \to \infty$, we obtain

$$U(x_i) - U(x) + \|x - x_i\| k(U(x)) \leq 0.$$

So, for sufficiently large i, for the function U the generalized CLC (2.1) holds with $x' = x_i$ and $m = +\infty$. Hence, Theorem 2.2 implies that for every $x_0 \in X$ there exists a desired point $\bar{x} \in X$. □

7.2. Extremal points of the perturbed function. Let us consider the problem of stability of the minima of lower semicontinuous functions under perturbations (see [2]).

Given a function $\Delta : X \to \mathbb{R} \cup \{+\infty\}$, denote by Ξ the set of minimum points of U, and denote by Ξ_Δ the set of minimum points of the restriction of the function Δ on Ξ.

Proposition 7.2. *Assume that*

- *the function U satisfies the generalized CLC* (2.1) *with some $k \in \mathcal{K}$;*
- *$\Xi_\Delta \neq \emptyset$ and $\theta := \min\{\Delta(x) : x \in \Xi\}$;*
- *for all $x \in X \setminus \Xi$ and $\bar{x} \in \Xi$ such that $\rho(x, \bar{x}) \leq D_k(U(x))$, the following inequality holds:*

$$\Delta(x) > \theta - D_k^{-1}(\rho(x, \bar{x})). \tag{7.2}$$

Then Ξ_Δ coincides with the set of minimal points of the function $U + \Delta$.

Proof. Assume again that $\gamma = 0$.

If $x \in \Xi_\Delta$ then $U(x) + \Delta(x) = \theta$. If $x \in \Xi \setminus \Xi_\Delta$ then $U(x) + \Delta(x) = \Delta(x) > \theta$. Assume now that $x \in X \setminus \Xi$. Theorem 2.2 implies that there exists a point $\bar{x} \in \Xi$ such that

$$U(\bar{x}) = 0 \quad \text{and} \quad \rho(x, \bar{x}) \leq D_k(U(x)).$$

Hence, it follows from (7.2) that

$$(U + \Delta)(x) \geq D_k^{-1}(\rho(x, \bar{x})) + \Delta(x) > \theta.$$

So, $(U + \Delta)(x) = \theta$ for all $x \in \Xi_\Delta$ and $(U + \Delta)(x) > \theta$ for all $x \notin \Xi_\Delta$. Therefore, the set Ξ_Δ coincides with the set of minimum points of the function $U + \Delta$. □

Remark 7.3. The assumption $\Xi_\Delta \neq \emptyset$ automatically holds if Ξ is compact or the restriction of Δ to Ξ is lower semicontinuous and satisfies the generalized CLC.

Remark 7.4. In the case when $k(t) = \text{const.} = k$, the assumption for the function Δ in Proposition 7.2 takes the following form: $\Delta(x) > \theta - k\rho(x, \bar{x})$ for all $x \in X \setminus \Xi$, for all $\bar{x} \in \Xi$ such that $\rho(x, \bar{x}) \leq \frac{U(x)}{k}$. In particular, this assumption holds if the function Δ satisfies the Lipschitz condition with a constant $c < k$. In this case, in [6] it was proved that the function $U + \Delta$ satisfies the CLC with the constant $k - c$ on the set $X \setminus \Xi$, i.e., for all $x \in X \setminus \Xi$ there exists $x' \in X$, $x' \neq x$ such that $(U + \Delta)(x') + (k - c)\rho(x, x') \leq (U + \Delta)(x)$.

Let us present the sufficient conditions for the generalized CLC to hold.

Proposition 7.5. *Assume that the function U satisfies generalized CLC* (2.1) *with a function $k \in \mathcal{K}$, and a function $V : X \to \mathbb{R} \cup \{+\infty\}$ satisfies the following relation:*

$$U(x) \leq V(x) \quad \forall\, x \in X, \quad U(x) = V(x) \quad \forall\, x \in \Xi.$$

Then the function V satisfies the generalized CLC with the function k_D with $D = D_k$.

Proof. Assume again that $\gamma = 0$. Theorem 2.2 implies that for every $x \in X$, if $U(x) > 0$ then there exists x' such that $U(x') = 0$ and $\rho(x, x') \leq D(U(x))$. Thus, $V(x) \geq U(x) > 0$ and

$$V(x') + k_D(U(x))\rho(x, x') = k_D(U(x))\rho(x, x') = \frac{U(x)}{D(U(x))}\rho(x, x') \leq U(x) \leq V(x).$$

Therefore, the function V satisfies the generalized CLC with the function k_D. ∎

7.3. Local variant of Theorem 2.2. Let $x_0 \in X$ and $\delta \in (0, +\infty]$. Set $B(x_0, \delta) := \{x \in X : \rho(x, x_0) \leq \delta\}$.

THEOREM 7.6. *Let $x_0 \in \operatorname{dom} U$, $k \in \mathcal{K}$, and*

$$\delta \geq D_k(U(x_0) - \gamma) = \int_0^{U(x_0)-\gamma} \frac{dt}{k(t)}.$$

Assume that for every $x \in B(x_0, \delta)$ satisfying the inequalities

$$U(x) > \gamma \quad \textit{and} \quad \rho(x_0, x) \leq D_k(U(x_0) - \gamma) - D_k(U(x) - \gamma) = \int_{U(x)-\gamma}^{U(x_0)-\gamma} \frac{d\tau}{k(\tau)}$$

there exists $x' \in X$ such that $x' \neq x$ and

$$U(x') + k(U(x) - \gamma)\rho(x, x') \leq U(x). \tag{7.3}$$

Then there exists a point $\bar{x} \in B(x_0, \delta)$ at which the minimum of the function U on the set $B(x_0, \delta)$ is attained, and moreover, $U(\bar{x}) = \gamma$.

Proof. Assume again that $\gamma = 0$. Set

$$\overline{X} = \big\{x \in B(x_0, \delta) : \rho(x, x_0) \leq D_k(U(x_0)) - D_k(U(x))\big\}.$$

This set is nonempty and closed, since $x_0 \in \overline{X}$; the function U is lower semicontinuous and the function $1/k$ is summable. Therefore, the metric space $(\overline{X}, \rho)$ is complete. Let us show that for the restriction of U to $\overline{X}$ the assumptions of Theorem 2.2 hold.

Take arbitrary points $x \in \overline{X}$ and $x' \in B(x_0, \delta)$ such that $U(x) > 0$, $x' \neq x$, and the inequality (7.3) holds. Then $x' \in \overline{X}$, since

$$\begin{aligned}\rho(x_0, x') &\leq \rho(x_0, x) + \rho(x, x') \leq D_k(U(x_0)) - D_k(U(x)) + \frac{U(x) - U(x')}{k(U(x))} \\ &\leq D_k(U(x_0)) - D_k(U(x)) + \int_{U(x')}^{U(x)} \frac{d\tau}{k(\tau)} = D_k(U(x_0)) - D_k(U(x')).\end{aligned}$$

Here the second inequality follows from the fact that $x \in \overline{X}$ and from (7.3); the third one follows from the function k being increasing. So, the assumptions of Theorem 2.2 hold for the restriction of the function U to $\overline{X}$. The existence of the desired point $\bar{x}$ follows from Theorem 2.2. ∎

Acknowledgment. The authors are grateful to the anonymous referees for attentive reading of the paper and useful remarks.

REFERENCES

[1] A. Arutyunov, *Relaxation and perturbation of optimal control problem*, Proc. Steklov Inst. Math., 220 (1998), pp. 23–30.

[2] A. Arutyunov, *Second-order conditions in extremal problems. The abnormal points*, Trans. Amer. Math. Soc., 350 (1998), pp. 4341–4365.

[3] A. Arutyunov, *Caristi's condition and existence of a minimum of a lower bounded function in a metric space. Applications to the theory of coincidence points*, Proc. Steklov Inst. Math., 291 (2015), pp. 24–37.

[4] A. Arutyunov and R. Vinter, *A simple "finite approximations" proof of the Pontryagin maximum principle under reduced differentiability hypotheses*, Set-Valued Var. Anal., 12 (2004), pp. 5–24.

[5] A. Arutyunov, N. Bobylev, and S. Korovin, *One remark to Ekeland's variational principle*, Comput. Math. Appl., 34 (1997), pp. 267–271.

[6] A. Arutyunov, B. Gel'man, E. Zhukovskiy, and S. Zhukovskiy, *Caristi-like condition. Existence of solutions to equations and minima of functions in metric spaces*, Fixed Point Theory, 20 (2019), pp. 31–58.

[7] A. Arutyunov, E. Zhukovskiy, and S. Zhukovskiy, *Coincidence points of set-valued mappings in partially ordered spaces*, Dokl. Math., 88 (2013), pp. 727–729.

[8] A. Arutyunov, E. Zhukovskiy, and S. Zhukovskiy, *Coincidence points principle for set-valued mappings in partially ordered spaces*, Topology Appl., 201 (2016), pp. 330–343.

[9] J.-P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley-Interscience, New York, 1984.

[10] E. Avakov, A. Arutyunov, and E. Zhukovskii, *Covering mappings and their applications to differential equations unsolved for the derivative*, Differ. Equ., 45 (2009), pp. 627–649.

[11] T. Bao and B. Mordukhovich, *Variational principles for set-valued mappings with applications to multiobjective optimization*, Control and Cybernet., 36 (2007), pp. 531–562.

[12] T. Bao, B. Mordukhovich, and A. Soubeyran, *Minimal points, variational principles, and variable preferences in set optimization*, J. Nonlinear Convex Anal., 16 (2015), pp. 1511–1537.

[13] J. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions*, Trans. Amer. Math. Soc., 303 (1987), pp. 517–527.

[14] J. Borwein and Q. Zhu, *Techniques of Variational Analysis*, Springer, New York, 2005.

[15] J. Caristi, *Fixed point theorems for mappings satisfying the inwardness condition*, Trans. Amer. Math. Soc., 215 (1976), pp. 241–251.

[16] F. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983.

[17] M. Frigon, *On some generalizations of Ekeland's principle and inward contractions in gauge spaces*, J. Fixed Point Theory Appl., 10 (2011), pp. 279–298.

[18] A. Granas and J. Dugundji, *Fixed Point Theory*, Springer, New York, 2003.

[19] A. Ioffe and V. Tikhomirov, *Some remarks on variational principles*, Math. Notes, 61 (1997), pp. 248–253.

[20] A. Kruger, S. Plubtieng, and T. Seangwattana, *Borwein–Preiss variational principle revisited*, J. Math. Anal. Appl., 435 (2016), pp. 1183–1193.

[21] Y. Li and S. Shi, *A generalization of Ekeland's ε-variational principle and its Borwein–Preiss smooth variant*, J. Math. Anal. Appl., 246 (2000), pp. 308–319.

[22] B. Malgrange, *Ideals of Differentiable Functions*, Oxford University Press, Oxford, 1966.

[23] B. Mordukhovich, *Variational Analysis and Generalized Differentiation* II. *Applications*, Springer, Berlin, 2006.

[24] B. Mordukhovich and B. Wang, *Necessary suboptimality and optimality conditions via variational principles*, SIAM J. Control Optim., 41 (2002), pp. 623–640.

[25] R. Phelps, *Support cones in Banach spaces and their applications*, Adv. Math., 13 (1974), pp. 1–19.

[26] F. Sullivan, *A characterization of complete metric spaces*, Proc. Amer. Math. Soc., 83 (1981), pp. 345–346.

[27] W. Takahashi, *Minimization theorems and fixed point theorems*, in Nonlin. Anal. Math. Econ. 829, T. Maruyama, ed., 1993, Research Institute for Mathematical Sciences, Kyoto University, Kyoto, pp. 175–191.

[28] J. Ume, *A minimization theorem in quasi-metric spaces and its applications*, Int. J. Math. Math. Sci., 31 (2002), pp. 443–447.