JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 115, No. 2, pp. 345–368, November 2002 (© 2002)

Well-Posedness by Perturbations of Variational Problems[1]

B. LEMAIRE,[2] C. OULD AHMED SALEM,[3] and J. P. REVALSKI[4]

Communicated by J. P. Crouzeix

Abstract. In this paper, we consider the extension of the notion of well-posedness by perturbations, introduced by Zolezzi for optimization problems, to other related variational problems like inclusion problems and fixed-point problems. Then, we study the conditions under which there is equivalence of the well-posedness in the above sense between different problems. Relations with the so-called diagonal well-posedness are also given. Finally, an application to staircase iteration methods is presented.

Key Words. Well-posedness, convex functions, subdifferentials, optimization, maximal monotone operators, Yosida regularization, inclusion, fixed points.

1. Introduction

An initial, already classical, notion of well-posedness for an unconstrained minimization problem determined by an extended real-valued function $f\colon X \to \mathbb{R} \cup \{+\infty\}$, where X is a Banach space, is due to Tykhonov (Ref. 1). The problem of minimizing f on X is said to be Tykhonov well-posed if it has unique solution toward which every minimizing sequence of the problem converges. Clearly, this notion is motivated by the numerical methods producing optimizing sequences. The idea of the behavior of the minimizing

[1]Most of this work was done while the third author was visiting the University of Montpellier II during the second semester of the 1998/99 academic year. The same author was partially supported by the Bulgarian National Fund for Scientific Research under Contract MM-701/97.

[2]Professor, Département de Mathématiques, Université de Montpellier II, Montpellier, France.
[3]Assistant Professor, Université des Sciences et Techniques, Nouakchott, Mauritania.
[4]Professor, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria.

sequences was used by different authors to extend this concept to two directions: first for strengthened notions (also for constrained problems) and second for the case in which the optimal solutions are not unique (see e.g. Refs. 2–12). Unfortunately, these notions do not imply always appropriate continuous dependence of the solution on the data; therefore, in particular, they are not suitable for numerical methods, where the function f is approximated by a family or a sequence of functions (Refs. 5, 9, 12–16). For this reason, new notions of well-posedness have been introduced and studied. We will mention two: first, the well-posedness by perturbations, introduced by Zolezzi in Refs. 17–18, which incorporates both the idea of Tykhonov well-posedness and the continuous dependence of the solution on the data (the latter is termed often Hadamard well-posedness in optimization); and second the diagonal well-posedness (see Ref. 13), which relies on the approximation scheme of a given problem.

The aim of this paper is to study these latter notions, and mostly the first one, for related variational problems of fixed-points and inclusions and to investigate their links with minimization problems. For different reasons, both numerical and theoretical, we will consider the more general case when the requirement of the uniqueness of the solution is dropped.

In what follows, X will stand for a real Banach space and X^* will be its dual. For the norm in X and X^*, the symbol $\|\cdot\|$ will be used. The pairing between X and X^* will be designated by $\langle \cdot, \cdot \rangle$. Recall also that, given an extended real-valued function $f: X \to \mathbb{R} \cup \{+\infty\}$, the symbol $\operatorname{dom} f$ stands for the domain of f, i.e., the set

$$\{x \in X: f(x) < \infty\}.$$

The function f is called proper if its domain is nonempty.

For a subset $C \subset X$, we denote by $d(\cdot, C)$ the distance function generated by C, i.e., the function

$$d(x, C) := \inf\{\|x - y\|: y \in C\}, \qquad x \in X.$$

Given two sets C and D in X, as usual the excess of C over D is

$$c(C, D) := \sup\{d(x, D): x \in C\},$$

and the Hausdorff distance between C and D is

$$\operatorname{haus}(C, D) := \max\{e(C, D), e(D, C)\}.$$

In the rest of the paper we proceed as follows. In Section 2, we introduce well-posedness by perturbations for inclusion problems and fixed-point problems, having in mind the corresponding notion for minimization problems. Then, we study the links between these three types of problems for this concept. Section 3 is devoted to a brief discussion of the diagonal well-posedness and its connection with the well-posedness by perturbations. Finally, Section 4 contains applications to the convergence of staircase iteration methods.

2. Well-Posedness by Perturbations

The idea of well-posedness by perturbations proposed by Zolezzi in Refs. 17–18 for optimization problems is the following. One embeds the original problem into a parametrized family of similar problems and then requires the convergence of appropriately defined optimizing sequences to the solution of the original problem; for more details and results, see Refs. 17–18. We give below the precise definition of this notion for the case of nonunique solutions and extend it naturally for two (related) other types of problems: inclusion problems and fixed-point problems.

In what follows, P is a parameter space. Each $p \in P$ generates through the particular parametrization scheme a problem with solution set $S(p)$. The original problem corresponds to a certain fixed parameter. We will suppose that P is endowed with some convergence structure.

With each sequence of parameters $\{p_n\}$, we associate a sequence in the underlying space X, called asymptotically solving sequence. Let us see how this is done in the three cases which interest us.

2.1. Minimization Problems. See Zolezzi, Refs. 17–18. Let $f: X \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued function, and let

$$S := \operatorname{Argmin} f := \{x \in X: f(x) = \inf\{f(x'): x' \in X\} =: \inf f\}$$

be its set of minimizers in X. In this case, the parametrization scheme is given by a function $\tilde{f}: X \times P \to \mathbb{R} \cup \{+\infty\}$ such that

$$\tilde{f}(\cdot, p^*) = f, \qquad \text{for a certain } p^*.$$

We set

$$S(p) := \operatorname{Argmin} \tilde{f}(\cdot, p) := \{x \in X: \tilde{f}(x, p) = \inf \tilde{f}(\cdot, p)\}.$$

We will suppose that, for every $p \in P$, we have $\inf \tilde{f}(\cdot, p) > -\infty$. Given a sequence of parameters $\{p_n\}$, the sequence $\{x_n\} \subset X$ is called asymptotically solving corresponding to $\{p_n\}$ if it is $\tilde{f}$-asymptotically minimizing corresponding to $\{p_n\}$, which means that

$$\tilde{f}(x_n, p_n) - \inf \tilde{f}(\cdot, p_n) \to 0.$$

2.2. Inclusion Problems. Let Y be a real normed vector space with norm $\|\cdot\|_Y$, and let $T: X \rightrightarrows Y$ be a set-valued mapping between X and Y with domain

$$\mathrm{Dom}\,(T) := \{x \in X: Tx \neq \emptyset\}.$$

We will be interested in the problem of finding a solution to the following inclusion: find $x \in \mathrm{Dom}(T)$ such that $0 \in T(x)$. The solution set of this particular problem is

$$S := T^{-1}(0) = \{x \in X: 0 \in Tx\}.$$

Here, the parametrization scheme is given by a set-valued mapping $\tilde{T}: X \times P \rightarrow Y$ such that

$$\tilde{T}(\cdot, p^*) = T, \qquad \text{for some } p^*.$$

Let

$$S(p) := \{x \in X, 0 \in \tilde{T}(x, p)\}.$$

The sequence $\{x_n\} \subset X$ is called asymptotically solving, corresponding to the sequence of parameters $\{p_n\}$, if it is $\tilde{T}$-asymptotically stationary corresponding to $\{p_n\}$,

$$d(0, \tilde{T}(x_n, p_n)) \rightarrow 0,$$

or equivalently,

$$\forall n \in \mathbb{N}, \exists y_n \in \tilde{T}(x_n, p_n): \quad \|y_n\|_Y \rightarrow 0.$$

2.3. Fixed-Point Problems. Here, we are given a single-valued mapping $F: X \rightarrow X$, and we are looking for the fixed points of F. The particular solution set associated to F is

$$S := \mathrm{Fix}\, F = \{x \in X: x = Fx\},$$

and the parametrization is given by a single-valued mapping $\tilde{F}: X \times P \rightarrow X$ such that $\tilde{F}(\cdot, p^*) = F$ for a certain p^*.

In this case, a sequence $\{x_n\} \subset X$ is called asymptotically solving, corresponding to the sequence of parameters $\{p_n\}$, if it is $\tilde{F}$-asymptotically regular corresponding to $\{p_n\}$,

$$\|x_n - \tilde{F}(x_n, p_n)\| \rightarrow 0.$$

Now, we are ready to give a definition of well-posedness by perturbations in the three cases above. As we mentioned for the case of minimization problems, this notion was introduced by Zolezzi (Refs. 17–18). Here, we consider the more general case of nonunique solutions.

Definition 2.1. Let $p^* \in P$ be any particular problem from the above variational problems. Then, this problem is said to be well-posed by perturbations with respect to the particular perturbation scheme if the following conditions are satisfied:

(i) $S(p^*) \neq \emptyset$,
(ii) if $p_n \to p^*$, then any asymptotically solving sequence $\{x_n\}$ corresponding to $\{p_n\}$ converges to $S(p^*)$ in the sense that

$$d(x_n, S(p^*)) \to 0.$$

Let us remark that, for minimization problems, the above notion, in which we allow the nonuniqueness of the solution, is termed stability by perturbations in Ref. 19.

If the perturbed problem is independent of p, the well-posedness by perturbations for the above problems reduces to the concept of well-posedness (for the case of nonunique solution) as it is considered in Ref. 8. The latter or similar kind of well-posedness for minimization problems has also been studied in Refs. 2–3, 7–8. Of course, the particular case of minimization problems with a unique solution and without parameters is exactly the classical Tykhonov well-posedness.

Let us mention also that, with $Y = X$, fixed-point well-posedness by perturbations for $F: X \to X$ is nothing but inclusion well-posedness in the same sense for $I - F$ with the inherited parameterization scheme, where I is the identity in X.

It is well-known that the minimization of a proper lower-semicontinuous convex function $f: X \to \mathbb{R} \cup \{+\infty\}$ in X is equivalent to the inclusion problem generated by its subdifferential $\partial f: X \rightrightarrows X^*$,

$$\partial f(x) := \{x^* \in X^*: f(y) - f(x) \geq \langle y - x, x^* \rangle, \forall y \in X\}, \qquad x \in \operatorname{dom} f,$$

i.e.,

$$x \in \operatorname{Argmin} f, \qquad \text{iff } 0 \in \partial f(x).$$

We show that this equivalence extends to the case of well-posedness by perturbations under some conditions. For a bounded from below function $f: X \to \mathbb{R} \cup \{+\infty\}$ and $\epsilon > 0$, we use the symbol ϵ-Argmin f to denote the set of ϵ-minimizers of f, i.e., the set

$$\{x \in X: f(x) \leq \inf f + \epsilon\}.$$

Let us recall also the definition of ϵ-subdifferential $\partial_\epsilon f: X \rightrightarrows X^*$ of a given proper lower semicontinuous convex function f,

$$\partial_\epsilon f(x) := \{x^* \in X^*: f(y) - f(x) \geq \langle y - x, x^* \rangle - \epsilon, \forall y \in X\}, \qquad x \in \operatorname{dom} f.$$

One has always

$$\partial_\epsilon f(x) \neq \emptyset, \qquad \text{for every} \epsilon > 0 \text{ and } x \in \operatorname{dom} f.$$

Now, we are ready to prove the following theorem.

Theorem 2.1. Let X be a real Banach space and, for each p in P, let $\tilde{f}(\cdot, p)$ be a proper lower-semicontinuous convex function in X, where $\tilde{f}(\cdot, p^*) = f$ for some $p^* \in P$. Then, the minimization problem generated by f is $\tilde{f}$-well-posed by perturbations provided the inclusion problem determined by ∂f is $\partial\tilde{f}$-well-posed by perturbations. In addition, if the solution set $S = \operatorname{Argmin} f$ is bounded, or if, for every sequence $p_n \to p^*$, one has

$$e(S, (1/n) - \operatorname{Argmin} \tilde{f}(\cdot, p_n)) \to 0, \qquad \text{as } n \to \infty,$$

then the converse implication is also true.

Proof. Let the inclusion problem for ∂f be $\partial\tilde{f}$-well-posed by perturbations. The solution set S for the inclusion generated by ∂f is the same as the solution set of the minimization problem generated by f. Hence, we have to check only that, given a sequence of parameters $p_n \to p^*$, then every $\tilde{f}$-asymptotically minimizing sequence $\{x_n\}$ corresponding to $\{p_n\}$ converges to S. Let us fix such a sequence. Then, there exists a sequence of positive reals $\epsilon_n \to 0$ so that

$$\tilde{f}(x_n, p_n) < \inf \tilde{f}(\cdot, p_n) + \epsilon_n, \qquad \forall n \in \mathbb{N},$$

from where we see that

$$0 \in \partial_{\epsilon_n} \tilde{f}(\cdot, p_n)(x_n), \qquad \text{for every } n \in \mathbb{N}.$$

By the Brøndsted–Rockafellar theorem (Ref. 20), for each x_n there are $\bar{x}_n \in X$ and $x_n^* \in X^*$ such that

$$x_n^* \in \partial\tilde{f}(\bar{x}_n, p_n), \qquad \|x_n - \bar{x}_n\| \leq \sqrt{\epsilon_n}, \qquad \|x_n^*\| \leq \sqrt{\epsilon_n}.$$

In particular, $\{\bar{x}_n\}$ is a $\partial\tilde{f}$-asymptotically stationary sequence for the inclusion problem for ∂f and thanks to its $\partial\tilde{f}$-well-posedness by perturbations, we get

$$d(\bar{x}_n, S) \to 0.$$

Therefore, from the inequality

$$d(x_n, S) \leq \|x_n - \bar{x}_n\| + d(\bar{x}_n, S),$$

and $\epsilon_n \to 0$, we conclude that $d(x_n, S) \to 0$. The proof of this implication is completed.

For the other implication, suppose that the minimization problem generated by f is $\tilde{f}$-well-posed by perturbations, and in addition suppose that the common solution set S of this problem and of the inclusion problem for ∂f is either bounded or, for every sequence of parameters $p_n \to p^*$,

$$e(S, (1/n) - \text{Argmin}\, \tilde{f}(\cdot, p_n)) \to 0.$$

Take a particular sequence of parameters $p_n \to p^*$, and let $\{x_n\}$ be an asymptotically stationary sequence corresponding to $\{p_n\}$. This means that there exists $\{x_n^*\} \subset X^*$ so that $x_n^* \in \partial \tilde{f}(x_n, p_n)$ for every $n \in \mathbb{N}$ and in addition $\|x_n^*\| \to 0$. By definition, the former means that, for every $n \in \mathbb{N}$,

$$\tilde{f}(x_n, p_n) \leq \tilde{f}(x, p_n) + \langle x_n - x, x_n^* \rangle, \qquad \forall x \in \text{dom}\, \tilde{f}(\cdot, p_n). \tag{1}$$

Now, let $\{z_n\}$ be a sequence from X such that, for every $n \in \mathbb{N}$, we have

$$z_n \in \text{dom}\, \tilde{f}(\cdot, p_n) \quad \text{and} \quad z_n \in (1/n) - \text{Argmin}\, \tilde{f}(\cdot, p_n).$$

The latter means that

$$\tilde{f}(z_n, p_n) \leq \inf \tilde{f}(\cdot, p_n) + 1/n, \qquad \forall_n \in \mathbb{N}. \tag{2}$$

According to the supposed conditions, we choose such a sequence as follows;

Case 1. If S is bounded, the sequence $\{z_n\}$ is arbitrary with the property (2).

Case 2. If $e(S, (1/n) - \text{Argmin}\, \tilde{f}(\cdot, p_n)) \to 0$, we choose first for every $n \in \mathbb{N}$ a point $s_n \in S$ so that

$$\|x_n - s_n\| \leq d(x_n, S) + 1/n, \tag{3}$$

and then for every s_n, we take $z_n \in (1/n)\text{Argmin}\, \tilde{f}(\cdot, p_n)$ with the property that

$$\begin{aligned} \|s_n - z_n\| &\leq d(s_n, (1/n) - \text{Argmin}\, \tilde{f}(\cdot, p_n)) + 1/n \\ &\leq e(S, (1/n) - \text{Argmin}\, \tilde{f}(\cdot, p_n)) + 1/n. \end{aligned} \tag{4}$$

Observe that, in both cases, the sequence $\{z_n\}$ is asymptotically minimizing corresponding to $\{p_n\}$ and therefore $d(z_n, S) \to 0$.

Writing (1) for z_n and using (2), we have that, for every $n \in \mathbb{N}$,

$$\begin{aligned} \tilde{f}(x_n, p_n) &\leq \tilde{f}(z_n, p_n) + \langle x_n - z_n, x_n^* \rangle \\ &\leq \inf \tilde{f}(\cdot, p_n) + 1/n + \|x_n^*\| \|x_n - z_n\|. \end{aligned} \tag{5}$$

Now, if $\|x_n^*\| \cdot \|x_n - z_n\| \to 0$, we see that $\{x_n\}$ is asymptotically minimizing corresponding to $\{p_n\}$; hence, by the fact that the minimization problem for

f is $\tilde{f}$-well-posed by perturbations, we conclude that $d(x_n, S) \to 0$. So, the proof of the theorem will be complete if we show that

$$\|x_n^*\| \cdot \|x_n - z_n\| \to 0.$$

Suppose that

$$\|x_n^*\| \|x_n - z_n\| \not\to 0.$$

We will show that this is not possible by obtaining a contradiction. Indeed, passing to subsequences if needed and having in mind that $\|x_n^*\| \to 0$, without loss of generality we may think that there exists $a \in \mathbb{R}$ so that, for every $n \in \mathbb{N}$,

$$\|x_n^*\| \|x_n - z_n\| \geq a > 0 \quad \text{and} \quad \|x_n^*\| < 1.$$

For every $n \in \mathbb{N}$, let

$$\lambda_n = \sqrt{\|x_n^*\|}\, a / [\|x_n^*\| \cdot \|x_n - z_n\|].$$

Obviously, $\lambda_n \in (0, 1)$ for every $n \in \mathbb{N}$. For every n, consider now the segment $[x_n, z_n]$ and the point y_n on it defined by

$$y_n = \lambda_n x_n + (1 - \lambda_n) z_n, \qquad n \in \mathbb{N}.$$

By the convexity of $\tilde{f}(\cdot, p_n)$ and also using (2), (5), and the definition of λ_n, we get that, for every n,

$$\begin{aligned} \tilde{f}(y_n, p_n) &\leq \lambda_n \tilde{f}(x_n, p_n) + (1 - \lambda_n) \tilde{f}(z_n, p_n) \\ &\leq \lambda_n \inf \tilde{f}(\cdot, p_n) + \lambda_n / n + \lambda_n \|x_n^*\| \|x_n - z_n\| \\ &\quad + (1 - \lambda_n) \inf \tilde{f}(\cdot, p_n) + (1 - \lambda_n)/n \\ &= \inf \tilde{f}(\cdot, p_n) + 1/n + a \sqrt{\|x_n^*\|}. \end{aligned}$$

Since $\|x_n^*\| \to 0$, the last inequality implies that the sequence $\{y_n\}$ is asymptotically minimizing corresponding to $\{p_n\}$ and this entails that we would have $d(y_n, S) \to 0$. We will get a contradiction with this last conclusion in the two cases that we consider.

Case 1. S is bounded. Let $s \in S$ be arbitrary. Then, by the definition of λ_n, for every $n \in \mathbb{N}$ we have

$$\begin{aligned} \|y_n - s\| &\geq \|y_n - z_n\| - \|z_n - s\| \\ &= \lambda_n \|x_n - z_n\| - \|z_n - s\| \\ &= a / \sqrt{\|x_n^*\|} - \|z_n - s\|. \end{aligned}$$

Since S is bounded and $d(z_n, S) \to 0$, there exists a constant $M > 0$ so that

$$\|z_n - s\| \leq M, \quad \text{for every } n \in \mathbb{N} \text{ and every } s \in S.$$

This together with the last inequality and the fact that $a/\sqrt{\|x_n^*\|}) \to \infty$ show that $\|y_n - s\| \to \infty$ uniformly on $s \in S$. The last contradicts $d(y_n, S) \to 0$.

Case 2. $e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \to 0$. In this case, we will first show that

$$d(y_n, S) \geq \|z_n - y_n\| - 2/n - e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)), \quad \text{for every } n \in \mathbb{N}. \tag{6}$$

Indeed, otherwise for some $n \in \mathbb{N}$, there would exist $v_n \in S$ so that

$$\|y_n - v_n\| < \|z_n - y_n\| - 2/n - e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)).$$

Hence, using this last inequality, (3), and (4), we have

$$\begin{aligned} d(x_n, S) &\geq \|x_n - s_n\| - 1/n \\ &\geq \|x_n - z_n\| - \|z_n - s_n\| - 1/n \\ &\geq \|x_n - z_n\| - 2/n - e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \\ &= \|x_n - y_n\| + \|y_n - z_n\| - 2/n - e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \\ &> \|x_n - y_n\| + \|y_n - v_n\| \\ &\geq \|x_n - v_n\|; \end{aligned}$$

i.e.,

$$d(x_n, S) > \|x_n - v_n\|,$$

with $v_n \in S$, which is a contradiction. Hence, (6) is true.

As a consequence of (6) and the definitions of λ_n and y_n, we obtain that, for every $n \in \mathbb{N}$,

$$\begin{aligned} d(y_n, S) &\geq \|z_n - y_n\| - 2/n - e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \\ &= \lambda_n \|z_n - x_n\| - 2/n - e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \\ &= a/\sqrt{\|x_n^*\|} - 2/n - e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)), \end{aligned}$$

from where we see that $d(y_n, S) \to \infty$, because $\|x_n^*\| \to 0$ and $e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \to 0$, once again a contradiction.

Therefore, in both cases, we have obtained a contradiction with the supposition that $\|x_n^*\| \|x_n - z_n\| \not\to 0$. The proof of Theorem 2.1 is completed. □

Let us mention that, for the second implication in the above theorem, the completeness of the space X was not needed.

Remark 2.1. A close look at the proof of the second implication shows that the condition $e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \to 0$ can be replaced by $e(S, \epsilon_n - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \to 0$ for any sequence $\{\epsilon_n\}$ such that ϵ_n goes to zero by strictly positive values.

Remark 2.2. If the family $\{\tilde{f}(\cdot, p)\}$ reduces to a single function f, then the above result is obtained in Ref. 7, where the proof for the second implication passes through an intermediate result. Observe that, in this particular case, we have always the condition $e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \to 0$ fulfilled. Therefore, the result from Ref. 7 is a consequence from the theorem above.

Remark 2.3. It is easy to see that, if the minimization problem for f is $\tilde{f}$-well-posed by perturbations, then we have always that, for every sequence of parameters $p_n \to p^*$, $e((1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n), S) \to 0$. Hence, our supposition that

> "f is $\tilde{f}$-well-posed by perturbations and for every sequence of parameters $p_n \to p^*$, $e(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \to 0$"

is in fact equivalent to

> "f is $\tilde{f}$-well-posed by perturbations and, for every sequence of parameters $p_n \to p^*$, haus $(S, (1/n) - \text{Argmin}\,\tilde{f}(\cdot, p_n)) \to 0$."

On the other hand, we see from the next example that, in the case of unbounded set of solutions, the supposition made in the theorem is crucial.

Example 2.1. Let $X = \mathbb{R}$, $P = [0, 1]$, and let the function $\tilde{f}$ be defined as follows:

$$\tilde{f}(x, p) := \begin{cases} -x + 1, & \text{if } p = 0,\ x \in (-\infty,\ 1), \\ 0, & \text{if } p = 0,\ x \in [1, +\infty), \\ -px + 1, & \text{if } p > 0,\ x \in (-\infty,\ 1/p), \\ 0, & \text{if } p > 0,\ x \in [1/p, +\infty). \end{cases}$$

It is easily seen that

$$S := S(0) = [1, +\infty),$$

$$\inf \tilde{f}(\cdot, p) = 0, \qquad \text{for every } p \in P,$$

and that the problem generated by $f := \tilde{f}(\cdot, 0)$ is $\tilde{f}$-well-posed by perturbations. On the other hand, if $p_n = 1/n, n \in \mathbb{N}$, one sees that, for every $n \in \mathbb{N}$,

$$\partial\tilde{f}(x, p_n) = \{-1/n\}, \qquad \text{for every } x \in (-\infty, n).$$

Hence, for example, the constant sequence $x_n = 1/2, n \in \mathbb{N}$, is a $\partial\tilde{f}$-asymptotically stationary sequence corresponding to $\{p_n\}$, which stays away from S. Therefore, the inclusion problem for ∂f is not $\partial\tilde{f}$-well-posed by perturbations. Observe that here, for every $n \in \mathbb{N}$,

$$(1/n) - \operatorname{Argmin} \tilde{f}(\cdot, p_n) = [n-1, +\infty);$$

consequently, for $n \geq 2$, we have

$$e(S, (1/n) - \operatorname{Argmin} \tilde{f}(\cdot, p_n)) = n - 2 \to +\infty;$$

i.e., the condition from Theorem 2.1 is not fulfilled.

As always, the case of a unique solution is of special importance. The following is an immediate corollary from Theorem 2.1.

Corollary 2.1. Let X be a real Banach space and, for all p in P, let $\tilde{f}(\cdot, p)$ be a proper lower-semicontinuous convex function on X, where $\tilde{f}(\cdot, p^*) = f$ for some p^*. Suppose that f has a unique minimizer. Then, the minimization problem for f is $\tilde{f}$-well-posed by perturbations if and only if the inclusion problem for ∂f is $\partial\tilde{f}$-well-posed by perturbations.

Now, we pass to the study of the relation between the well-posedness of certain inclusion problems and the same notion of appropriately linked fixed-point problems. More precisely, we will be interested in inclusion problems generated by monotone operators. Recall that a set-valued operator $A: X \rightrightarrows X^*$ is said to be monotone if it satisfies the following condition:

$$\langle y - x, y^* - x^* \rangle \geq 0, \quad \text{for every } (x, x^*), (y, y^*) \in \operatorname{Gr}(A),$$

where the symbol $\operatorname{Gr}(A)$ means the graph of A,

$$\operatorname{Gr}(A) := \{(x, x^*) \in X \times X^* : x^* \in Ax\}.$$

A monotone operator $A: X \rightrightarrows X^*$ is said to be maximal if its graph is not contained properly in the graph of any other monotone operator from X to X^*. The subdifferentials of proper convex lower-semicontinuous functions as well as many differential operators turn out to be maximal monotone.

Further in this section, we will restrict our considerations to the (still enough general) case of a reflexive real Banach space X. In this case [see e.g. Diestel (Ref.21)], we may suppose that both the norm in X and the dual norm in X^* are Fréchet differentiable (except at the origin), locally uniformly rotund norms. In particular, these norms are strictly convex and moreover satisfy the following Kadec–Klee property:

"if $x_n \to x$ weakly in X and $\|x_n\| \to \|x\|$, then $x_n \to x$ in the norm of X,"

and the same for the weak turology and the norm topology in X^*. In such a situation, the duality mapping J between X and X^*, given by

$$Jx := \{x^* \in X^* : \langle x, x^* \rangle = \|x\|^2 = \|x^*\|^2\}, \qquad x \in X,$$

which is always a maximal monotone operator with $\mathrm{Dom}(J) = X$, is also one-to-one, surjective, and norm-to-norm continuous.

Using this mapping, one can define, for every maximal monotone operator $A: X \rightrightarrows X^*$, its resolvent J_λ^A of order $\lambda > 0$ as the operator from X into X in the following way: for any $x \in X$, $J_\lambda^A x$ is the unique [by the classical results of Browder (Ref. 22) and Rockafellar (Ref. 23)] solution x_λ of the inclusion

$$0 \in J(x_\lambda - x) + \lambda A x_\lambda. \tag{7}$$

The resolvent J_λ^A maps X into $\mathrm{Dom}(A)$ and is norm-to-weak continuous. It is used to define another important mapping related to A, the Yosida approximation A_λ of A of order λ, as follows:

$$A_\lambda x := (1/\lambda)\, J(x - x_\lambda), \qquad x \in X. \tag{8}$$

The mapping A_λ is an everywhere defined single-valued maximal monotone operator between X and X^*. One has always

$$A_\lambda x \in A(J_\lambda^A x), \qquad \text{for every } x \in X.$$

Moreover, for every $\lambda > 0$,

$$J_\lambda^A x = x - \lambda J^{-1} A_\lambda x, \qquad \text{for every } x \in X. \tag{9}$$

In the special case of a Hilbert space X, J is the identity and the above formulas take the well-known forms

$$J_\lambda^A = (J + \lambda A)^{-1} \quad \text{and} \quad A_\lambda = (J - J_\lambda^A)/\lambda.$$

Now, we turn back to the inclusion problems. First, since the inclusion (7) has a unique solution, it is verifiable easily that, if $A: X \rightrightarrows X^*$ is a maximal monotone operator, then for any $\lambda > 0$, the inclusion problem

$$\text{find } x \in \mathrm{Dom}(A) \text{ such that } 0 \in Ax$$

is equivalent to the fixed-point problem

$$\text{find } x \in X \text{ such that } x = J_\lambda^A x.$$

On the other hand, because of (8), the latter problem is equivalent to the inclusion (which is in fact an equality)

$$\text{find } x \in X \text{ such that } A_\lambda x = 0.$$

In the next theorem, we see that the well-posedness by perturbations for the different variational problems generated by a maximal monotone operator $A: X \rightrightarrows X^*$ and its approximations, defined above, are equivalent.

Theorem 2.2. Let $A: X \rightrightarrows X^*$ be a maximal monotone operator. Consider a parametrization $\tilde{A}(\cdot, p), p \in P$, so that, for every $p \in P, \tilde{A}(\cdot, p)$ is a maximal monotone operator between X and X^* such that $\tilde{A}(\cdot, p^*) = A$, for some $p^* \in P$. Let $\lambda > 0$. Then, the following assertions are equivalent:

(a) the inclusion problem generated by A is $\tilde{A}$-well-posed by perturbations;

(b) the fixed-point problem for J_λ^A is $J_\lambda^{\tilde{A}}$-well-posed by perturbations;

(c) the inclusion problem for A_λ is $\tilde{A}_\lambda$-well-posed by perturbations.

Proof. First we prove that (a) $\Rightarrow$ (b). Let the inclusion problem for A be $\tilde{A}$-well-posed by perturbations with (nonempty) solution set S. Then, as pointed above, the fixed-point problem generated by the resolvent J_λ^A has the same set of solutions S. Take a sequence of parameters $p_n \to p^*$ and suppose that $\{u_n\}$ is a $J_\lambda^{\tilde{A}}$-asymptotically regular sequence corresponding to $\{p_n\}$. For every $n \in \mathbb{N}$, let

$$w_n = J_\lambda^{\tilde{A}(\cdot, p_n)} u_n.$$

By definition, this means that

$$0 \in J(w_n - u_n) + \lambda \tilde{A}(w_n, p_n).$$

The last entails that, for every $n \in \mathbb{N}$, the functional

$$y_n^* := -(1/\lambda) J(w_n - u_n)$$

satisfies the inclusion

$$y_n^* \in \tilde{A}(w_n, p_n).$$

On the other hand, by the definition of J, we have that, for every $n \in \mathbb{N}$,

$$\|y_n^*\| = (1/\lambda) \|w_n - u_n\|.$$

But since the sequence $\{u_n\}$ is asymptotically regular, we obtain that $\|w_n - u_n\| \to 0$, from where $\|y_n^*\| \to 0$. Therefore, the sequence $\{w_n\}$ is $\tilde{A}$-asymptotically stationary corresponding to $\{p_n\}$. Since the inclusion problem for A is $\tilde{A}$-well-posed by perturbations, we get $d(w_n, S) \to 0$, from where, using once again that $\|w_n - u_n\| \to 0$, we obtain that $d(u_n, S) \to 0$ as well. The proof of this implication is completed.

Conversely, to prove that (b)$\Rightarrow$(a), suppose that the fixed point for the resolvent J_λ^A is $J_\lambda^{\tilde{A}}$-well-posed by perturbations with (nonempty) solution set S. Let again $\{p_n\}$ be a sequence of parameters converging to p^*. Let further $\{x_n\}$ be an $\tilde{A}$-asymptotically stationary sequence corresponding to $\{p_n\}$. This means that, for every $n = 1, 2, \ldots$, there exists $x_n^* \in \tilde{A}(x_n, p_n)$ so that $\|x_n^*\| \to 0$. For every $n = 1,2, \ldots$, put

$$y_n := x_n - J^{-1}(-\lambda x_n^*).$$

Then, it is seen easily that, for every $n = 1,2, \ldots$, we have

$$0 \in J(x_n - y_n) + \lambda \tilde{A}(x_n, p_n),$$

which shows that, for every $n = 1, 2 \ldots$,

$$x_n = J_\lambda^{\tilde{A}(\cdot, p_n)} y_n.$$

From here, using the definition of y_n and J, we get that, for each $n = 1, 2, \ldots$,

$$\begin{aligned} \|y_n - J_\lambda^{\tilde{A}(\cdot, p_n)} y_n\| &= \|y_n - x_n\| \\ &= \|J^{-1}(-\lambda x_n^*)\| \\ &= \lambda \|x_n^*\|. \end{aligned}$$

As $\|x_n^*\| \to 0$, we conclude that the sequence $\{y_n\}$ is asymptotically regular corresponding to $\{p_n\}$ with respect to the parametrization $J_\lambda^{\tilde{A}}$. Thanks to the $J_\lambda^{\tilde{A}}$-well-posedness by perturbations of the fixed-point problem generated by J_λ^A, we get that $d(y_n, S) \to 0$. Since $\|x_n - y_n\| \to 0$, we deduce that $d(x_n, S) \to 0$. The proof of the equivalence between the assertions (a) and (b) is completed.

As to the equivalence between (b) and (c), we see from the definition of the Yosida approximations (8) that the inclusion problem for A_λ and the fixed-point problem for J_λ^A have the same set of solutions. On the other hand, given a sequence of parameters $\{p_n\} \subset P$, it is seen easily from the relation (9) that the $\tilde{A}_\lambda$-asymptotically stationary sequences for the inclusion problem A_λ corresponding to $\{p_n\}$ are $J_\lambda^{\tilde{A}}$-asymptotically regular sequences for the fixed-point problem J_λ^A corresponding to $\{p_n\}$ and vice-versa, from where the equivalence between (b) and (c) follows. □

3. Links with Diagonal Well-Posedness

In this section, we outline briefly the natural relation of the notion of well-posedness by perturbations with the so-called diagonal well-posedness considered in Refs.13, 15. We will do this only for minimization problems. The natural counterparts for inclusion problems and fixed-point problems as well as the corresponding relations are left to the reader.

The idea of diagonal well-posedness is motivated by problems where the initial (bounded from below) function $f: X \to \mathbb{R} \cup \{+\infty\}$ to be minimized is approximated in some way by a sequence of (bounded from below) functions $f_n: X \to \mathbb{R} \cup \{+\infty\}$, and one considers the corresponding minimization problems generated by f_n in X. In such a situation, one possible approach is: instead of searching for a notion of well-posedness for the particular problem determined by f, it is better sometimes, from a practical point of view, to look for a notion of well-posedness for the whole sequence $\{f_n\}$.

More precisely (for more details, see Refs. 13, 15), let X be a real Banach space and we are given a sequence of usually lower-semicontinuous bounded from below functions $f_n: X \to \mathbb{R} \cup \{+\infty\}$. The sequence $\{x_n\} \subset X$ is said to be diagonally minimizing for $\{f_n\}$ if

$$f_n(x_n) - \inf f_n \to 0.$$

Definition 3.1. The sequence $\{f_n\}$ is called diagonally well-posed iff

(i) $\forall n \in \mathbb{N}$, $\operatorname{Argmin} f_n \neq \emptyset$;
(ii) any diagonally minimizing sequence $\{x_n\}$ for $\{f_n\}$ is diagonally convergent to S_n, i.e.,

$$d(x_n, S_n) \to 0.$$

Of course, if $f_n \equiv f$, for every n, then the above is exactly the well-posedness of f.

As we mentioned, there are natural extensions of this idea for inclusion and fixed point-problems: one has a sequence of set-valued operators

$$T_n: X \rightrightarrows (Y, \|\cdot\|_Y), \; n \in \mathbb{N},$$

for inclusion problems, and a sequence of single-valued operators

$$F_n: X \to X, \; n \in \mathbb{N}$$

for fixed-point problems. Then, one defines the corresponding diagonally solving sequences: diagonally stationary sequences $\{x_n\} \subset X$, if $d(0, T_n(x_n)) \to 0$ for inclusion problems, and diagonally regular sequences $\{x_n\}$, if $\|x_n - F_n(x_n)\| \to 0$ for fixed-point problems. Finally, the notion of diagonal well-posedness in the two cases requires the existence of solutions for the perturbed problems and the diagonal convergence of the corresponding diagonally solving sequences.

For details and a result similar to Theorem 2.1, as well as for other properties of this notion, the reader is referred to Refs. 13, 15. Here, we want to mention only that, using the same techniques as in Theorem 2.2, one can prove the following analogous result.

Theorem 3.1. Let X be a reflexive real Banach space and, for every $n \in \mathbb{N}$, let A_n be a maximal monotone operator between X and X^*. Let $\lambda > 0$. Then, the following assertions are equivalent:

(a) the sequence of inclusion problems $\{A_n\}$ is diagonally well-posed;
(b) the sequence of fixed-point problems $\{J_\lambda^{A_n}\}$ is diagonally well-posed;
(c) the sequence of inclusion problems $\{(A_n)_\lambda\}$ is diagonally well-posed.

Further, we give an obvious natural relation of the notion of diagonal well-posedness with the notion of well-posedness by perturbations.

Theorem 3.2. Let X be a real Banach space, and let P be a convergence space of parameters. Suppose that, for all $p \in P, \tilde{f}(\cdot, p)$ is a bounded from below proper extended real-valued function on X so that $S(p) := \operatorname{Argmin} \tilde{f}(\cdot, p) \neq \emptyset$, $\tilde{f}(\cdot, p^*) = f$, $S := S(p^*)$, and $\operatorname{haus}(S(p), S) \overset{p \to p^*}{\to} 0$. Then, the minimization problem for f is $\tilde{f}$-well-posed by perturbations iff, for all $\{p_n\}$ converging to p^*, $\{\tilde{f}(\cdot, p_n)\}$ is diagonally well-posed.

Proof. The proof is elementary using (in both directions) the triangle inequality for the distance function. □

Let us mention that the assumption

$$\operatorname{haus}(S(p), S) \xrightarrow{p \to p^*} 0$$

cannot be avoided in general. This is seen for instance by Example 2.1 if one takes $p_n = 1/n$ and $x_n = n - 1$, $n \in \mathbb{N}$.

We finish the section by pointing out that exact analogues of the above result are true if one considers inclusion problems or fixed-point problems.

4. Staircase Iteration and Well-Posedness

4.1. General Discussion. Let us describe the general scheme of a staircase iteration process and its relation to well-posedness by perturbations. A standard situation, which occurs often in numerical analysis is the following;

it is given a problem $\mathscr{P}$ with solution set S and there exists a sequence of approximate problems $\mathscr{P}_n$, $n = 1, 2, \ldots$, of the same type as $\mathscr{P}$ and an iterative method adapted to the type of problems considered. The idea is to apply this iterative method successively to $\mathscr{P}_n$ for every $n = 1, 2, \ldots$, but for each particular n stopping after a finite number of iterates k_n and then restarting with a new approximate problem from the last iterate. This gives rise to a staircase iterative process.

More precisely, let x_1^0 be a given initial point in the underlying space and, for every $n = 1, 2, \ldots$, denote by Q_n the iteration mapping for the problem $\mathscr{P}_n$ given by the method. Then, for fixed $n \geq 1$, one makes k_n iterations of the usual type,

$$x_n^k := Q_n x_n^{k-1}, \qquad k = 1, 2, \ldots, k_n.$$

To pass to the next problem $\mathscr{P}_{n+1}$ one just sets as an initial point

$$x_{n+1}^0 := x_n^{k_n}.$$

Of course, the aim of the stopping rule is to obtain a sequence $\{x_n^{k_n}\}$ which converges to S in a reasonable sense.

In some cases, after a possible reformulation of the problem $\mathscr{P}$ as a fixed-point problem generated by some mapping F in a Banach space X, the sequence $\{Q_n\}$ is defined via a certain embedding $\tilde{F}$ of F, as in Section 2, in the following way: there is a sequence $\{p_n\}$ of parameters so that $Q_n := \tilde{F}(\cdot, p_n)$. Hence, if the problem $\mathscr{P}$ is well-posed by perturbations with respect to this embedding, the question amounts to finding a stopping rule (the choice of k_n) so that the sequence $\{x_n^{k_n}\}$ is asymptotically regular corresponding to $\{p_n\}$.

In general, for the three types of problems (minimization, inclusion, and fixed-point) considered in Section 2, this is done as follows. In the notation of Section 2, given a sequence of parameters $\{p_n\}$, let

$$s_n(x) := \tilde{f}(p_n, x) - \inf \tilde{f}(\cdot, p_n) \text{ [resp. } s_n(x) := d(0, \tilde{T}(x, p_n)) \text{ or}$$
$$s_n(x) := \|x - \tilde{F}(x, p_n)\|\text{]}.$$

Then, obviously, $\{x_n\}$ is asymptotically solving corresponding to $\{p_n\}$ iff

$$\lim_{n \to +\infty} s_n(x_n) = 0.$$

Now, let us assume that, for each fixed n, we can build a sequence $\{x_n^k\}$ such that

$$\lim_{k \to +\infty} s_n(x_n^k) = 0;$$

i.e., in a certain sense, the sequence $\{x_n^k\}_{k=1}^{\infty}$ is asymptotically solving for the fixed n (or p_n). Let $\{\epsilon_n\}$ be a sequence of positive reals converging to 0, and set k_n so that

$$s_n(x_n^{k_n}) \leq \epsilon_n.$$

It is clear that

$$s_n(x_n^{k_n}) \to 0, \qquad \text{as } n \to \infty.$$

In particular, coming back to the above staircase iterative process, we see that, if the iteration Q_n for each fixed n generates a sequence $\{x_n^k\}_{k=1}^{\infty}$ so that

$$\lim_{k \to +\infty} \|x_n^k - x_n^{k-1}\| = 0,$$

then k_n must be chosen so that

$$\|x_n^{k_n} - x_n^{k_n+1}\| \leq \epsilon_n,$$

where ϵ_n converges to 0 with positive values.

4.2. Application: Staircase Proximal Method. Let $\{T_n\}$ be a sequence of maximal monotone operators on the real Hilbert space X. Let us consider the staircase iterative process with

$$Q_n := J_\lambda^{T_n},$$

where

$$J_\lambda^{T_n} := (J + \lambda T_n)^{-1}$$

denotes the resolvent of parameter $\lambda > 0$ associated with T_n. It is known (Ref. 24) that the iteration process generated by Q_n at fixed n generates an asymptotically regular sequence for this fixed n, i.e.,

$$\lim_{k \to +\infty} \|x_n^k - Q_n x_n^k\| = 0,$$

in the following cases: T_n is the subdifferential of a bounded-from-below proper lower-semicontinuous convex function f_n or $T_n^{-1}(0)$ is nonempty. So, in these cases, we can choose k_n as described above. Therefore, if T_n is defined from some maximal monotone operator T via an embedding $\tilde{T}(\cdot, \cdot)$ such that $T = \tilde{T}(\cdot, p)$ and $T_n(\cdot) = \tilde{T}(\cdot, p_n)$ for some sequence of parameters $p_n \to p$, and if the inclusion problem for T is $\tilde{T}$-well-posed by perturbations, thanks to Theorem 2.2, the sequence $\{x_n^{k_n}\}$ generated satisfies $d(x_n^{k_n}, T^{-1}(0)) \to 0$.

The above described method is called staircase proximal method. We will give two examples which lead to convergence in this method. In what follows, given a nonempty set $C \subset X$, the symbol $\delta_C(\cdot)$ denotes the usual indicator function of C; i.e.,

$$\delta_C(x) = 0, \text{ if } x \in C,$$

$$\delta_C(x) = +\infty, \text{ otherwise.}$$

Example 4.1. Exterior Penalization. Let $X := \mathbb{R}^m, m \in \mathbb{N}$, and let $f_0: X \to \mathbb{R}$ be a bounded-from-below lower-semicontinuous function which is coercive, i.e.,

$$f_0(x) \to \infty \qquad \text{if } \|x\| \to \infty.$$

Let C be a nonempty closed subset of X, and set

$$f(\cdot) := f_0(\cdot) + \delta_C(\cdot) \quad \text{and} \quad S := \operatorname{Argmin} f.$$

The coercivity of f_0 implies that S is nonempty and bounded.

Put $P = [0, +\infty)$, and define the embedding

$$\tilde{f}(\cdot, p) := f_0(\cdot) + (1/p)\Phi(\cdot), \qquad \text{if } p > 0,$$

$$\tilde{f}(\cdot, 0) := f(\cdot),$$

where $\Phi: X \to \mathbb{R}_+$ is a lower-semicontinuous function such that $C = \Phi^{-1}(0)$. It can be seen that the minimization problem for f is $\tilde{f}$-well-posed by perturbations. Indeed, let $\{p_n\} \subset P$ be a sequence of parameters converging to 0 and take a sequence $\{x_n\} \subset X$ which is asymptotically minimizing for $\{p_n\}$, i.e.,

$$\tilde{f}(x_n, p_n) - \inf \tilde{f}(\cdot, p_n) \to 0. \tag{10}$$

First of all, since Φ is positive and $C = \Phi^{-1}(0)$, we see easily that

$$\begin{aligned} \inf f_0 &\le \inf \tilde{f}(\cdot, p_n) \\ &\le \inf \{f_0(x): x \in C\} \\ &=: \inf_C f_0, \qquad \text{for every } n. \end{aligned} \tag{11}$$

Now, suppose that $d(x_n, S)$ does not converge to zero. Without loss of generality, we may assume that

$$d(x_n, S) \ge \epsilon, \qquad \text{for some } \epsilon > 0.$$

Since f_0 is coercive, then because of (10)–(11) the sequence $\{x_n\}$ is bounded. Therefore, again without loss of generality, we can assume that $x_n \to x_0$ for some $x_0 \in X$. Obviously, $x_0 \notin S$. We claim that $x_0 \in C$. Otherwise, $\Phi(x_0) > 0$ and, because of the lower semicontinuity of Φ, we would have $\Phi(x_n) > t > 0$, for some t and large n. But this would imply $\tilde{f}(x_n, p_n) \to \infty$, hence $\inf \tilde{f}(\cdot, p_n) \to \infty$, contradicting (11). Thus, $x_0 \in C$. Since $x_0 \notin S$, for some $t \in \mathbb{R}$, we have

$$f_0(x_0) > t > \inf_C f_0,$$

which imply that

$$f_0(x_n) > t > \inf_C f_0, \qquad \text{for large } n,$$

this time because of the lower semicontinuity of f_0. This is a contradiction, having in mind (10), (11), and the obvious inequality $\tilde{f}(x_n, p_n) \geq f_0(x_n)$.

The contradiction achieved shows that the minimization problem for f is $\tilde{f}$-well-posed by perturbations. Now, suppose in addition that f_0, C, Φ are convex. Therefore, since S is bounded, by Theorem 2.1 the inclusion problem for ∂f is $\partial\tilde{f}$-well-posed by perturbations. Hence, if we set $f_n := \tilde{f}(\cdot, p_n)$ for a sequence of parameters $\{p_n\}$ converging to zero, the staircase proximal method described above for $T_n := \partial f_n$ is convergent, i.e.,

$$d(x_n^{k_n}, S) \to 0.$$

Example 4.2. Variational Approximation. Let us assume here that X is a real Hilbert space, $C \subset X$ is nonempty closed and convex, and $f_0: X \to \mathbb{R}$ is bounded from below, continuous, and strongly convex; the latter as usual means that, for some $a > 0$, we have

$$f_0(\lambda x + (1-\lambda)y) \leq \lambda f_0(x) + (1-\lambda) f_0(y) - a\lambda(1-\lambda)\|x - y\|^2,$$
$$\text{for every } x, y \in X \text{ and } \lambda \in (0, 1).$$

Setting

$$f(\cdot) := f_0(\cdot) + \delta_C(\cdot),$$

these suppositions imply that

$$S := \operatorname{Argmin} f = \{\hat{x}\}, \qquad \text{for some } \hat{x} \in C.$$

Let $P := [0, +\infty)$, and define

$$\tilde{f}(\cdot, p) := f_0 + \delta_{C_p}, \qquad \text{if } p > 0,$$
$$\tilde{f}(\cdot, 0) := f,$$

where C_p is a nonempty closed convex subset of X, such that weak $-\limsup_{p \to 0} C_p \subset C$ and, for each sequence $\{p_n\}$ converging to 0, there exists $\hat{x}_n \in C_{p_n}$ so that $\hat{x}_n \to \hat{x}$. As usual, weak $-\limsup_{p\to 0} C_p \subset C$ means that, whenever we have $p_n \to 0$ and x_n converges weakly to x, with $x_n \in C_{p_n}$ for every n, then $x \in C$.

Here again the minimization problem for f is $\tilde{f}$-well-posed by perturbations. To see this, let $p_n \to 0$ and take a sequence as in (10). Since

$$\inf \tilde{f}(\cdot, p_n) = \inf_{C_n} f_0 \leq f_0(\hat{x}_n) \quad \text{and} \quad \hat{x}_n \to \hat{x},$$

it follows that $\{x_n\}$ is bounded and

$$\overline{\lim} \inf f(\cdot, p_n) \leq f_0(\hat{x}).$$

Arguing by contradiction as above, we may suppose that

$$\|x_n - \hat{x}\| \geq \epsilon > 0, \qquad \text{for some } \epsilon,$$

and that x_n converges weakly to some x_0. Because obviously $x_n \in C_{p_n}$ for large n, the latter implies that $x_0 \in C$. Since f_0 is weakly lower semicontinuous, and using the above, we get

$$\begin{aligned} f_0(x_0) &\leq \underline{\lim}\, f_0(x_n) \\ &= \underline{\lim}\, \tilde{f}(x_n, p_n) \\ &\leq \overline{\lim}\, \tilde{f}(x_n, p_n) \\ &\leq f_0(\hat{x}). \end{aligned}$$

This implies $x_0 = \hat{x}$ and

$$\begin{aligned} \lim f_0(x_n) &= \lim \tilde{f}(x_n, p_n) \\ &= \lim \inf \tilde{f}(\cdot, p_n) \\ &= f_0(\hat{x}). \end{aligned}$$

But this is a contradiction already: if we take the point

$$y_n := (1/2)(x_n + \hat{x}_n), n = 1, 2 \ldots,$$

we have $y_n \in C_{p_n}$ and it is seen easily that

$$\begin{aligned} \lim \tilde{f}(y_n, p_n) &= \lim f_0(y_n) \\ &= f_0(\hat{x}). \end{aligned}$$

Thus, applying the strong convexity assumption of f_0 for the points $y_n, x_n, \hat{x}_n$, and having in mind the above chain of equalities, we get

$$\|x_n - \hat{x}_n\| \to 0.$$

Therefore, by Corollary 2.1, the inclusion problem for ∂f is $\partial \tilde{f}$-well-posed by perturbations. Hence, if $f_n := \tilde{f}(\cdot, p_n)$, for some sequence $\{p_n\}$ converging to 0, then the staircase proximal method described above is convergent; i.e.,

$$x_n^{k_n} \to \hat{x}.$$

We refer to Ref. 25 for another discussion on such a coupling between the proximal method and approximation; we refer to Ref. 26 on the general staircase iteration method in connection with regularization.

5. Conclusions

In this paper, we have considered the extension of the concept of well-posedness by perturbations, introduced by Zolezzi (Refs. 17–18) for optimization problems, to other related variational problems, like inclusion problems and fixed-point problems. This notion unifies the two underlying ideas of stability of the solution: continuous dependence on the data and convergence of the solving sequences.

We have shown that, under some assumptions, the well-posedness in the above sense of a minimization problem for a proper lower-semicontinuous convex function f, defined in a Banach space, is equivalent to the same notion for the inclusion problem determined by the corresponding subdifferential ∂f (Theorem 2.1). On the other hand, the well-posedness by perturbations of the inclusion problem for a maximal monotone operator A in a reflexive Banach space has turned out to be equivalent to the same well-posedness of the fixed-point problem for the resolvent of A (Theorem 2.2).

In Section 3, we outlined briefly the relation of the well-posedness by perturbations for the above variational problems with another similar concept, the so-called diagonal well-posedness (Theorems 3.1 and 3.2). Finally, in Section 4, we showed how Theorems 2.1 and 2.2 can be used to assure the convergence of the proximal staircase method (Examples 4.1 and 4.2).

References

1. Tykhonov, A. N., *On the Stability of the Functional Optimization Problem*, USSR Journal of Computational Mathematics and Mathematical Physics, Vol. 6, pp. 631–634, 1966.
2. Auslender, A., Crouzeix, J.P., and Cominetti, R., *Convex Functions with Unbounded Level Sets and Applications to Duality Theory*, SIAM Journal on Optimization, Vol. 3, pp. 669–687, 1993.
3. Bednarczuk, E., and Penot, J. P., *Metrically Well-Set Minimization Problems*, Applied Mathematics and Optimization, Vol. 26, pp. 273–285, 1992.

4. BEER, G., and LUCCHETTI, R., *Well-Posed Optimization Problems and a New Topology for the Closed Subsets of a Metric Space*, Rocky Mountain Journal of Mathematics, Vol. 23, pp. 1197–1220, 1993.
5. DONTCHEV, A. L., and ZOLEZZI, T., *Well-Posed Optimization Problems*, Lectures Notes in Mathematics, Springer Verlag, Berlin, Germany, Vol. 1543, 1993.
6. FURI, M., and VIGNOLI, A., *About Well-Posed Optimization Problems for Functionals in Metric Spaces*, Journal of Optimization Theory and Applications, Vol. 5, pp. 225–229, 1970.
7. LEMAIRE, B., *Bonne Position, Conditionnement, et Bon Comportement Asymptotique*, Report 5, Séminaire d'Analyse Convexe de Montpellier, 1992.
8. LEMAIRE, B., *Well-Posedness, Conditioning, and Regularization of Minimization, Inclusion, and Fixed-Point Problems*, Pliska Studia Mathematica Bulgarica, Vol. 12, pp. 71–84, 1998.
9. OULD AHMED SALEM, C., *Approximation de Points Fixes d'une Contraction*, Thèse de Doctorat, Université Montpellier 11, 1998.
10. REVALSKI, J. P., *Generie Well-Posedness in Some Classes of Optimization Problems*, Acta Universitatis Carolinae-Mathematica et Physica, Vol. 28, pp. 117–125, 1987.
11. REVALSKI, J. P., and ZHIVKOV, N. V., *Well-Posed Constrained Optimization Problems in Metric Spaces*, Journal of Optimization Theory and Applications, Vol. 76, pp. 143–163, 1993.
12. ZOLEZZI, T., *On Equiwellset Minimum Problems*, Applied Mathematics and Optimization, Vol. 4, pp. 209–223, 1978.
13. BAHRAOUI, M.A., and LEMAIRE, B., *Convergence of Diagonally Stationary Sequences in Convex Optimization*, Set-Valued Analysis, Vol. 2, pp. 1–13, 1994.
14. BOYER, R., *Quelques Algorithmes Diagonaux en Optimisation Convexe*, Thèse de Doctorat, Université de Provence, 1974.
15. LEMAIRE, B., *Bounded Diagonally Stationary Sequences in Convex Optimization*, Journal of Convex Analysis (to appear).
16. LORIDAN, P., *Sur un Procédé d'Optimisation Utilisant Simultanément les Méthodes de Pénalisation et des Variations Locales*, SIAM Journal on Control, Vol. 11, pp. 159–172, 1973 and Vol. 11, pp. 173–184, 1973.
17. ZOLEZZI, T., *Well-posed Criteria in Optimization with Application to the Calculus of Variations*, Nonlinear Analysis: Theory, Methods, and Applications, Vol. 25, pp. 437–453, 1995.
18. ZOLEZZI, T., *Extended Well-Posedness of Optimization Problems*, Journal of Optimization Theory and Applications, Vol. 91, pp. 257–268, 1996.
19. LUCCHETTI, R., and ZOLEZZI, T., *On Well-Posedness and Stability Analysis in Optimization*, Mathematical Programming with Data Perturbations, Edited by A. V. Fiacco, Lecture Notes in Pure and Applied Mathematics, Marcel Dekker, New York, NY, Vol. 195, pp. 223–251, 1998.
20. BRØNDSTED, A., and ROCKAFELLAR, R. T., *On the Subdifferentiability of Convex Functions*, Proceedings of the American Mathematical Society, Vol. 16, pp. 605–611, 1965.
21. DIESTEL, J., *Geometry of Banach Spaces: Selected Topics*, Lecture Notes in Mathematics, Springer Verlag, Berlin, Germany, 1975.

22. BROWDER, F. E., *Nonlinear Maximal Monotone Operators in Banach Space*, Mathematische Annalen, Vol. 175, pp. 89–113, 1968.
23. ROCKAFELLAR, R. T., *On the Maximality of Sums of Nonlinear Monotone Operators*, Transactions of the American Mathematical Society, Vol. 149, pp. 75–88, 1970.
24. ROCKAFELLAR, R. T., *Monotone Operators and the Proximal Point Algorithm*, SIAM Journal on Control and Optimization, Vol. 14, pp. 877–898, 1976.
25. KAPLAN, A., and TICHATSCHKE, R., *Stable Methods for Ill-Posed Problems*, Akademie Verlag, Berlin, Germany, 1994.
26. LEMAIRE, B., *Regularization of Fixed-Point Problems and Staircase Iteration*, Ill-Posed Variational Problems and Regularization Techniques, Edited by M. Théra and R. Tichatschke, Lecture Notes in Economics and Mathematical Systems, Springer Verlag, Berlin, Germany, Vol. 477, pp. 151–166, 1999.