On two questions for Hybrid Proximal Point methods allowing for constant relative errors.

Rolando Gárciga Otero* Alfredo Iusem†
B. F. Svaiter‡

Instituto de Matemática Pura e Aplicada,
Estrada Dona Castorina 110,
Rio de Janeiro, RJ, 22460-320, Brazil.

April 25, 2001

Abstract

We show that the hybrid steps on the Proximal Point methods, allowing for constant relative errors, are necessary in order to ensure boundedness of the iterates, even in the optimization case. Moreover, we show that such conditions do not imply that the sequence of errors results summable "a posteriori".

Keywords: proximal point method, relative error, inexact solutions, hybrid steps.

1 Introduction

The Proximal Point method, whose origins can be traced back e.g. to [7], [9], [8], was presented in [10] as a method for finding zeroes of monotone operators in Hilbert spaces. The *exact* Proximal Point Method is defined as follows: Given a Hilbert space H and a maximal monotone point-to-set operator $T : H \to \mathcal{P}(H)$, take a bounded sequence of regularization

*(**rgarciga@impa.br**).
†(**iusp@impa.br**). The work of this author was partially supported by CNPq grant no. 301280/86
‡(**benar@impa.br**). The work of this author was partially supported by CNPq grant no. 301200/93-9

parameters $\{\lambda_k\} \subset \mathbb{R}_{++}$ and any initial $z^0 \in H$, and then, given z^k, define z^{k+1} as the only $z \in H$ such that

$$0 \in T(z) + \lambda_k(z - z^k). \tag{1}$$

It was proved in [10] that when T has zeroes the sequence $\{x^k\}$ defined by (1) is weakly convergent to a zero of T. This convergence is not strong in general, a fact that was proved by Güler in [4], through an example in the space ℓ^2 for the optimization case.

The resolvent operator P_k of T, for λ_k fixed, is given by

$$P_k = (T + \lambda_k I)^{-1}.$$

Hence, the exact Proximal Point Method define the sequence $\{z^k\}$ by the iteration $z^{k+1} = P_k(z^k)$.

Inexact variants in Hilbert spaces were considered as early as in [10], where the iteration

$$z^{k+1} \approx P_k(z^k) \tag{2}$$

is considered. Here $\|z^{k+1} - P_k(z^k)\| = \epsilon_k$ is the error in the k-th iteration, and convergence of the sequence $\{z^k\}$ to a zero of T is ensured under the assumption that $\sum_{k=0}^{\infty} \epsilon_k < \infty$. Such summability condition demands that the precision in the computations increase with the iteration index k.

New variations of the Proximal Point Algorithm, accepting constant relative errors, have been recently presented in [11], [12], [13]. The basic feature of these methods is to combine, in each iteration, an approximated solution of the proximal problem with a projection [11] or an extragradient step [12, 13]. For this reason, these are hybrid methods (proximal-projection or proximal-extragradient). In both of these hybrid methods, the proximal problem (1) is decomposed in the following equivalent system:

$$\begin{aligned} &v \in T(z), \\ &v + \lambda_k(z - z^k) = 0. \end{aligned} \tag{3}$$

An approximated solution of these system is used to compute the next iterate by means of a projection or an extragradient step. The error tolerance is given by a *fixed* $\sigma \in [0, 1)$. In the Hybrid Projection-Proximal method, an approximated solution of the proximal system is a pair $(\tilde{z}^k, \tilde{v}^k)$ with

$$\begin{aligned} &\tilde{v}^k \in T(\tilde{z}^k), \\ &\tilde{v}^k + \lambda_k(\tilde{z}^k - z^k) = e^k \approx 0. \end{aligned} \tag{4}$$

The pair $(\tilde{z}^k, \tilde{v}^k)$ is acceptable if

$$\|e^k\| \leq \sigma \max\{\|\tilde{v}^k\|, \lambda_k \|\tilde{z}^k - z^k\|\}. \tag{5}$$

Then the next iterate is obtained as the orthogonal projection of z^k onto the hyperplane $H_k = \{z \in H : \langle \tilde{v}^k, z - \tilde{z}^k \rangle = 0\}$, i.e.

$$z^{k+1} = z^k - \frac{\langle \tilde{v}^k, z^k - \tilde{z}^k \rangle}{\|\tilde{v}^k\|^2} \tilde{v}^k. \tag{6}$$

As for the Hybrid Extragradient-Proximal method, in the system (3), the inclusion of $\tilde{v}^k \in T(\tilde{z}^k)$ may also be relaxed, by means of an enlargement of T [2]. To simplify our exposition, we consider only approximated solutions as in (4). In this case, the pair $(\tilde{z}^k, \tilde{v}^k)$ (as in (4)) is acceptable if

$$\|e^k\| \leq \sigma \lambda_k \|\tilde{z}^k - z^k\|, \tag{7}$$

where again $\sigma \in [0, 1)$ is a fixed (relative) error tolerance. Then the next iterate is obtained by an extragradient step,

$$z^{k+1} = z^k - \lambda_k^{-1} \tilde{v}^k. \tag{8}$$

If T has zeroes, then, sequences $\{z^k\}$ defined by (4),(5),(6) and (4),(7),(8) respectively, converge weakly to a zero of T; otherwise (i.e., if $T^{-1} = \emptyset$) these sequences are unbounded [11, 13].

The main objective of this work is to answer the following questions, concerning these hybrid methods:

1.) Is boundedness of the iterates guaranteed without the hybrid steps?

   Or, equivalently:

   Does the classical PPA converges under the "relative error criteria" of the hybrid methods?

2.) Do the error measures, mainly the distances of the inexact solutions to the resolvent, i.e. $\|\tilde{z}^k - P_k(z^k)\|$, result summable "a posteriori"?

It should be observed that Solodov and Svaiter gave in [11] a negative answer to question (1.), taking $T(x, y) = (x, -y)$ in $\mathbb{R}^2$. This operator T is not a subdifferential, so that question (1.) remains open for the optimization case. We prove here that, even for the optimization case, the answer for both questions is negative in general. We give appropriate examples establishing this fact in the next section.

Note that error criterium (7) is stronger than (5). Hence particular instances of both methods for the optimization case (i.e. $T = \partial g$ for some proper, closed and convex $g : H \to \mathbb{R} \cup \{\infty\}$), are given below:

1. Choose $z^0 \in H$.
2. Given z^k, find $\tilde{v}^k, \tilde{z}^k \in H$ such that

$$\begin{aligned} \tilde{v}^k &\in \partial g(\tilde{z}^k), && (9)\\ \tilde{v}^k &+ \lambda_k(\tilde{z}^k - z^k) = e^k, && (10)\\ \|e^k\| &\leq \sigma\lambda_k\|\tilde{z}^k - z^k\|. && (11) \end{aligned}$$

3. Hybrid step:

$$\begin{aligned} z^{k+1} &= z^k - \frac{\langle \tilde{v}^k, z^k - \tilde{z}^k\rangle}{\|\tilde{v}^k\|^2}\tilde{v}^k. \quad (Projection) && (12)\\ z^{k+1} &= z^k - \lambda_k^{-1}\tilde{v}^k. \quad (Extragradient) && (13) \end{aligned}$$

2 Answering the two open questions

We start with Güler's example on the existence of a function for which the exact Proximal Point algorithm converges weakly but not strongly.

Proposition 1. *(Güler) There exists a proper, closed and convex function in ℓ^2 which has minimizers and such that, given any bounded positive sequence $\{\lambda_k\}$, there exists a point $x \in dom\ f$ for which the exact Proximal Point algorithm* (1)*, starting at x, converges weakly but not strongly to a minimizer of f.*

Proof. See [4] Corollary 5.1.

We state next our main result.

Theorem 1. *There exist a Hilbert space H and a proper, closed and convex function g in H which has minimizers and such that, given any bounded positive sequence $\{\lambda_k\}$, there exists a point $z \in dom\ g$ such that*

i) The inexact Proximal Point algorithm allowing for constant relative errors as in (9)-(11)*, without the hybrid steps, starting at z, may generate unbounded sequences of iterates.*

ii) The Hybrid Proximal Point methods allowing for constant relative error in (9)-(11) *and* (12) *or* (9)-(11) *and* (13)*, starting at z, accept inexact iterates with nonsummable distances to the resolvent.*

Proof. Let f be the function provided by Proposition 1 and $\{x^k\}$ the sequence generated by the exact Proximal Point Algorithm (1) starting at point x, also given by Proposition 1. Then the sequence $\{x^k\}$ converges weakly to some x^*, but does not converge strongly. Therefore,

$$\sum_{k=1}^{\infty} \|x^{k+1} - x^k\| = \infty, \tag{14}$$

because otherwise the sequence $\{x^k\}$ would converge strongly to x^*. Since $\{x^k\}$ is generated by Algorithm (1), there exists a sequence $\{u^k\}_{k\geq 1} \in \ell^2$ such that for all k:

$$\begin{aligned} u^{k+1} &\in \partial f(x^{k+1}), & (15)\\ u^{k+1} &+ \lambda_k(x^{k+1} - x^k) = 0. & (16) \end{aligned}$$

Let us consider the product space $H = \ell^2 \times \mathbb{R}$ provided with the 2-norm $\|\cdot\|_2$ (i.e. $\|(x,t)\|_2 = (\|x\|^2_{\ell^2} + |t|^2)^{1/2}$). Trivially, H is a Hilbert space. Define the function $g : H \to \mathbb{R} \cup \{\infty\}$ as

$$g(x,t) = f(x),$$

for all $(x,t) \in H$. Consequently g is a proper, closed and convex function with subdifferential at any (x,t) given by $\partial g(x,t) = (\partial f(x), 0)$. In particular, if S is the set of minimizers of f, then $S \times \mathbb{R}$ is the set of minimizers of g, hence nonempty.

Now we are ready to prove item (i). Choose any $t^0 \in \mathbb{R}$ and define the sequence $\{t^k\} \in \mathbb{R}$ as

$$t^{k+1} = t^k + \frac{\sigma}{\sqrt{1-\sigma^2}}\|x^{k+1} - x^k\| = t^0 + \frac{\sigma}{\sqrt{1-\sigma^2}} \sum_{n=0}^{k} \|x^{n+1} - x^n\|. \tag{17}$$

Define also, for $k = 0, 1, \cdots$

$$z^k = (x^k, t^k), \quad e^k = \lambda_k(0, t^{k+1} - t^k), \quad v^{k+1} = (u^{k+1}, 0). \tag{18}$$

Then

$$\begin{aligned} v^{k+1} &\in \partial g(z^{k+1}), & (19)\\ v^{k+1} &+ \lambda_k(z^{k+1} - z^k) = e^k, & (20) \end{aligned}$$

where (20) follows from (16), (19) from (15) and the residual e^k satisfies

$$\begin{aligned}\|e^k\|_2 &= \lambda_k|t^{k+1}-t^k| = \lambda_k\frac{\sigma}{\sqrt{1-\sigma^2}}\|x^{k+1}-x^k\| \\ &= \lambda_k\sigma\left(\|x^{k+1}-x^k\|^2+|t^{k+1}-t^k|^2\right)^{1/2} = \lambda_k\sigma\|z^{k+1}-z^k\|,\end{aligned}$$

using (17). It follows that the pair (z^{k+1}, v^{k+1}) satisfies (9)-(11), and can be accepted as an inexact solution of the k-th proximal subproblem. Hence, the sequence $\{z^k\}$ is generated by the inexact Proximal Point Algorithm, applied to an optimization problem and using the relative error tolerance of the hybrid methods. Trivially, from (14), (17) and (18),

$$\lim_{k\to\infty}\|z^k\|_2 = \infty,$$

which proves item (i).

For (ii) let us consider first the extragradient case. i.e. the method in (9)-(11) and (13). Through an induction argument, it can be proved that starting at $z^0 = (x, t^0)$ the inexact solutions can be taken as

$$\tilde{z}^k = (x^{k+1}, t^0 + \frac{\sigma}{\sqrt{1-\sigma^2}}\|x^{k+1}-x^k\|),\ \tilde{v}^k = (u^{k+1}, 0),$$

and the iterates as

$$z^k = (x^k, t^0).$$

Then $P_k(z^k) = z^{k+1}$ for all k and

$$\|\tilde{z}^k - P_k(z^k)\| = \|\tilde{z}^k - z^{k+1}\| = |\tilde{t}^k - t^0| = \frac{\sigma}{\sqrt{1-\sigma^2}}\|x^{k+1}-x^k\|,$$

which implies $\sum_{k=0}^{\infty}\|\tilde{z}^k - P_k(z^k)\| = \infty$. Note that z^{k+1} is also the projection of z^k onto $H_k = \{z \in H : \langle\tilde{v}^k, z - \tilde{z}^k\rangle = 0\}$, hence the same example works for the Hybrid Projection Proximal Algorithm. □

Remark: Observe that in the proof of item (ii), $\|e^k\| = \lambda_k|\tilde{t}^k - t^0|$ for all k. Consequently

$$\sum_{k=0}^{\infty}\|e^k\| \geq \underline{\lambda}\sum_{k=0}^{\infty}\|\tilde{z}^k - P_k(z^k)\| = \infty,$$

provided $\lambda_k \geq \underline{\lambda} > 0$ for all k, so that the residuals are not summable either.

Remark: Though the example given by Theorem (1) does not prove nonsummability of the distances of the hybrid iterates to the resolvent, because $z^{k+1} = P_k(z^k)$ for all k, a slight modification of it (e.g. with $g(x,t) = f(x) + t^2/2$) provides an example in which the condition $\sum_{k=0}^{\infty} \|z^{k+1} - P_k(z^k)\| < \infty$ does not hold either.

References

[1] Eckstein, J. Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming. *Mathematics of Operations Research* **18** (1993) 202-226.

[2] Burachik, R.S., Iusem, A.N., Svaiter, B.F. Enlargement of monotone operators with applications to variational inequalities *Set-Valued Analysis* **5** (1997) 159-180.

[3] Burachik, R.S., Svaiter, B.F. A relative error tolerance for a family of generalized proximal point methods (to be published).

[4] Güler, O. On the convergence of the proximal point algorithm for convex minimization. *SIAM Journal on Control and Optimization* **29** (1991) 403-419.

[5] Kassay, G., The proximal point algorithm for reflexive Banach spaces. *Studia Mathematica* **30** (1985) 9-17.

[6] Kiwiel, K. Proximal minimization methods with generalized Bregman functions. *SIAM Journal on Control and Optimization* **35** (1997) 1142-1168.

[7] Krasnoselskii, M.A. Two observations about the method of succesive approximations. *Uspekhi Matematicheskikh Nauk* **10** (1955) 123-127.

[8] Martinet, B. Régularisation d'inéquations variationnelles par approximations successives. *Revue Française de Informatique et Recherche Opérationnelle* **4** (1970) Ser. R-3, 154-158.

[9] Moreau, J. Proximité et dualité dans un espace hilbertien. *Bulletin de la Societé Mathématique de France* **93** (1965).

[10] Rockafellar, R.T. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization* **14** (1976) 877-898.

[11] Solodov, M.V., Svaiter, B.F. A hybrid projection–proximal point algorithm. *Journal of Convex Analysis* **6** (1999) 59-70.

[12] Solodov, M.V., Svaiter, B.F. An inexact hybrid generalized proximal point algorithms and some new results on the theory of Bregman functions. *Mathematics of Operations Research* **51** (2000) 479-494.

[13] Solodov, M.V., Svaiter, B.F. An inexact hybrid extragradient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Analysis* **7** (1999) 323-345.