

A generalization of the regularization proximal point method

Oganeditse A. Boikanyo ^{1*}; Gheorghe Moroşanu ²

(1) *Department of Mathematics, University of Botswana, Private Bag 00704, Gaborone, Botswana*

(2) *Department of Mathematics and its Applications, Central European University,
Nador u. 9, H-1051 Budapest, Hungary*

Dedicated to Professor Stepan A. Tersian on the occasion of his 60th birthday.

Copyright 2012 © Oganeditse A. Boikanyo and Gheorghe Moroşanu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

This paper deals with the generalized regularization proximal point method which was introduced by the authors in [Four parameter proximal point algorithms, *Nonlinear Anal.* 74 (2011), 544-555]. It is shown that sequences generated by it converge strongly under minimal assumptions on the control parameters involved. Thus the main result of this paper unify many results related to the prox-Tikhonov method, the contraction proximal point algorithm and/or the regularization method as well as some results of the above quoted paper.

Keywords : Maximal monotone operator; nonexpansive map; proximal point algorithm; prox-Tikhonov method, regularization method; resolvent operator

1 Introduction

Throughout this paper, H will be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Recall that a map $T : H \rightarrow H$ is called nonexpansive if for every $x, y \in H$ we have $\|Tx - Ty\| \leq \|x - y\|$. An operator $A : D(A) \subset H \rightarrow 2^H$ is said to be monotone if

$$\langle x - x', y - y' \rangle \geq 0, \quad \forall (x, y), (x', y') \in G(A).$$

In other words, its graph $G(A) = \{(x, y) \in H \times H : x \in D(A), y \in Ax\}$ is a monotone subset of the product space $H \times H$. An operator A is called maximal monotone if in

*Corresponding author. Email address: boikanyoa@gmail.com Tel: (+267) 3555045

addition to being monotone, its graph is not properly contained in the graph of any other monotone operator. For a maximal monotone operator A , the resolvent of A , defined by $J_\beta^A := (I + \beta A)^{-1}$, is well defined on the whole space H , single-valued and nonexpansive for every $\beta > 0$.

One of the oldest and most effective iterative method for solving the set valued equation

$$\text{find an } x \in D(A) \text{ such that } 0 \in A(x), \quad (1.1)$$

where A is a maximal monotone operator, is the so called proximal point algorithm (PPA) which was first introduced by Martinet [8] in 1970. Rockafellar [10] generalized the PPA of Martinet by defining a sequence (x_n) such that

$$x_{n+1} = J_{\beta_n} x_n + e_n, \quad n = 0, 1, \dots, \quad (1.2)$$

for any starting point $x_0 \in H$, where (e_n) is considered to be the sequence of computational errors and $(\beta_n) \subset (0, \infty)$. The sequence (x_n) is known to converge weakly to a solution of problem (1.1), if $\liminf_{n \rightarrow \infty} \beta_n > 0$ and $\sum_{n=0}^{\infty} \|e_n\| < \infty$, see [10], but fails in general to converge strongly [4]. As a result different proximal point algorithms which converge strongly have been constructed by several authors, see for example [11, 13]. One such algorithm which generates a sequence of proximal iterates according to the rule

$$x_{n+1} = \alpha_n u + (1 - \alpha_n) J_{\beta_n} x_n + e_n, \quad \text{for any } u, x_0 \in H \text{ and all } n \geq 0, \quad (1.3)$$

where $(\alpha_n) \subset (0, 1)$ and $(\beta_n) \subset (0, \infty)$, was introduced independently by Xu [13] and Kamimura and Takahashi [5]. Different assumptions on the sequences of control parameters (α_n) and (β_n) have been used to derive strong convergence results of the contraction proximal point algorithm (1.3) above to the solution of problem (1.1) which is nearest to u , see for example [1, 13] for details. The generalized contraction proximal point algorithm

$$x_{n+1} = \alpha_n u + \lambda_n x_n + \gamma_n J_{\beta_n} x_n + e_n, \quad n \geq 0, \quad (1.4)$$

where again $u, x_0 \in H$ are given, $\alpha_n \in (0, 1)$, $\lambda_n, \gamma_n \in [0, 1]$ with $\alpha_n + \lambda_n + \gamma_n = 1$, and $\beta_n \in (0, \infty)$, which was introduced by Yao and Noor [15] also converges strongly (under appropriate assumptions) to the solution of problem (1.1) which is nearest to u . Just as in the case of the scheme (1.3), different sets of conditions on the control parameters $\alpha_n, \lambda_n, \gamma_n$ and β_n have been used to prove strong convergence of the iterative process (1.4), see [2, 15, 3]. Another proximal method which generates strongly convergent sequences is the prox-Tikhonov method of Lehdili and Moudafi [6] which was extended by Xu [14] in the following way

$$x_{n+1} = J_{\beta_n}(\alpha_n u + (1 - \alpha_n)x_n + e_n), \quad \text{for all } n \geq 0, \quad (1.5)$$

where $u, x_0 \in H$ are given, $\alpha_n \in (0, 1)$ and $\beta_n \in (0, \infty)$. The authors [1] have shown that for $\alpha_n \rightarrow 0$ and $e_n \rightarrow 0$ as $n \rightarrow \infty$, the regularization method is equivalent to the scheme (1.3) above. Therefore, the results already proved for the contraction proximal point algorithm also hold for the regularization method and vice versa. The authors [2] generalized the regularization method as

$$x_{n+1} = J_{\beta_n}(\alpha_n u + \lambda_n x_n + \gamma_n T x_n + e_n) \quad \text{for } n = 0, 1, \dots, \quad (1.6)$$

where $T : H \rightarrow H$ is a nonexpansive map, $\beta_n \in (0, \infty)$ and $\alpha_n, \lambda_n, \gamma_n \in [0, 1]$ with $\alpha_n + \lambda_n + \gamma_n = 1$. They showed that for $\emptyset \neq A^{-1}(0) \subset \text{Fix}(T)$, where $\text{Fix}(T) := \{x \in H : x = Tx\}$, the sequence generated by this method is also strongly convergent (under some conditions on $\alpha_n, \lambda_n, \gamma_n, \beta_n$ and e_n) to a solution of (1.1) which is nearest to u . The purpose of this paper is to investigate if the method used in [12] can be applied to the scheme (1.6) (which is different from (1.4) except when $\lambda_n = 0$ for all n and $T = I$, the identity operator [2]) in order to get a strong convergence result of a sequence generated by it under minimal assumptions on the control parameters α_n and β_n , thereby refining previously obtained results associated with the iterative process (1.6).

2 Preliminaries

Our analysis will be based on the following two lemmas

Lemma 2.1 (Xu [13]). *Let (s_n) be a sequence of non-negative real numbers satisfying*

$$s_{n+1} \leq (1 - a_n)s_n + a_nb_n + c_n, \quad n \geq 0,$$

where $(a_n), (b_n)$ and (c_n) satisfy the conditions: (i) $(a_n) \subset (0, 1)$, with $\prod_{n=0}^{\infty} (1 - a_n) = 0$, (ii) $c_n \geq 0$ for all $n \geq 0$ with $\sum_{n=0}^{\infty} c_n < \infty$, and (iii) $\limsup_{n \rightarrow \infty} b_n \leq 0$. Then $\lim_{n \rightarrow \infty} s_n = 0$.

Remark 2.1. *If $\lim_{n \rightarrow \infty} a_n = 0$, then $\prod_{n=0}^{\infty} (1 - a_n) = 0$ if and only if $\sum_{n=0}^{\infty} a_n = \infty$.*

Lemma 2.2 (Maingé [7]). *Let (s_n) be a sequence of real numbers that does not decrease at infinity, in the sense that there exists a subsequence (s_{n_j}) of (s_n) such that $s_{n_j} < s_{n_j+1}$ for all $j \geq 0$. Define an integer sequence $(\tau(n))_{n \geq n_0}$ as*

$$\tau(n) = \max\{n_0 \leq k \leq n : s_k < s_{k+1}\}.$$

Then $\tau(n) \rightarrow \infty$ as $n \rightarrow \infty$ and for all $n \geq n_0$

$$\max\{s_{\tau(n)}, s_n\} \leq s_{\tau(n)+1}. \tag{2.7}$$

We will also need the following lemma whose proof can be easily reproduced.

Lemma 2.3 (Xu [14]). *Let $A : D(A) \subset H \rightarrow 2^H$ be a maximal monotone operator. For any $x \in H$ and $\mu \geq \beta > 0$, the following inequality holds:*

$$\|x - J_{\beta}x\| \leq 2\|x - J_{\mu}x\|.$$

The next lemma is well known, it can be found for example in [9, p. 20].

Lemma 2.4. *Any maximal monotone operator $A : D(A) \subset H \rightarrow 2^H$ satisfies the demiclosedness principle. In other words, given any two sequences (x_n) and (y_n) satisfying $x_n \rightarrow x$ and $y_n \rightharpoonup y$ with $(x_n, y_n) \in G(A)$, then $(x, y) \in G(A)$.*

3 Main Result

We will use some ideas of the papers [7, 12] to prove our main result below.

Theorem 3.1. *Let $A : D(A) \subset H \rightarrow 2^H$ be a maximal monotone operator and $T : H \rightarrow H$ a nonexpansive map with $\emptyset \neq F := A^{-1}(0) \subset \text{Fix}(T)$, where $\text{Fix}(T)$ is the fixed point set of T . For arbitrary but fixed vectors $x_0, u \in H$, let (x_n) be the sequence generated by (1.6), where $\beta_n \in (0, \infty)$ and $\alpha_n, \lambda_n, \gamma_n \in [0, 1]$ with $\alpha_n + \lambda_n + \gamma_n = 1$. Assume that $\lim_{n \rightarrow \infty} \alpha_n = 0$ with $\sum_{n=0}^{\infty} \alpha_n = \infty$ and $\beta_n \geq \beta$ for some $\beta > 0$. If either $\sum_{n=0}^{\infty} \|e_n\| < \infty$ or $\|e_n\|/\alpha_n \rightarrow 0$, then (x_n) converges strongly to the point of F nearest to u .*

Proof. We have shown in the proof of Theorem 5 [2] that the exact iterative process of (1.6), namely, the sequence (v_n) defined by

$$v_{n+1} = J_{\beta_n}(\alpha_n u + \lambda_n v_n + \gamma_n T v_n) \quad \text{for } n = 0, 1, \dots, \quad (3.8)$$

for any $v_0 \in H$ is bounded. (In fact, we showed that (x_n) is itself bounded). Now observe that from the nonexpansivity of T and of the resolvent operator, we have

$$\begin{aligned} \|x_{n+1} - v_{n+1}\| &\leq \lambda_n \|x_n - v_n\| + \gamma_n \|T x_n - T v_n\| + \|e_n\| \\ &\leq (1 - \alpha_n) \|x_n - v_n\| + \|e_n\|. \end{aligned}$$

It then follows from Lemma 2.1 that $\|x_n - v_n\| \rightarrow 0$ as $n \rightarrow \infty$. Therefore, it is enough to show that $v_n \rightarrow P_F u$, where $P_F u$ denotes the metric projection of u on F . For this purpose, we first show that for any $p \in F$, we have

$$\begin{aligned} (1 + \alpha_n) \|v_{n+1} - p\|^2 &\leq (1 - \alpha_n) \|v_n - p\|^2 + 2\alpha_n \langle u - p, v_{n+1} - p \rangle \\ &\quad - \lambda_n \|v_{n+1} - v_n\|^2 - \gamma_n \|T v_n - v_{n+1}\|^2. \end{aligned} \quad (3.9)$$

Indeed, multiplying

$$v_{n+1} - p + \beta_n A v_{n+1} \ni \alpha_n (u - p) + \lambda_n (v_n - p) + \gamma_n (T v_n - p)$$

scalarly by $v_{n+1} - p$ and using the monotonicity of A , we have

$$\begin{aligned} 2 \|v_{n+1} - p\|^2 &\leq 2\alpha_n \langle u - p, v_{n+1} - p \rangle + 2\lambda_n \langle v_n - p, v_{n+1} - p \rangle + 2\gamma_n \langle T v_n - p, v_{n+1} - p \rangle \\ &= 2\alpha_n \langle u - p, v_{n+1} - p \rangle + \lambda_n (\|v_n - p\|^2 + \|v_{n+1} - p\|^2 - \|v_{n+1} - v_n\|^2) \\ &\quad + \gamma_n (\|T v_n - p\|^2 + \|v_{n+1} - p\|^2 - \|T v_n - v_{n+1}\|^2) \\ &\leq (1 - \alpha_n) (\|v_{n+1} - p\|^2 + \|v_n - p\|^2) + 2\alpha_n \langle u - p, v_{n+1} - p \rangle \\ &\quad - \lambda_n \|v_{n+1} - v_n\|^2 - \gamma_n \|T v_n - v_{n+1}\|^2. \end{aligned}$$

Rearranging terms, we readily get (3.9). Denote $s_n := \|v_n - P_F u\|^2$. Then it follows from (3.9) and the boundedness of (v_n) that

$$s_{n+1} - s_n + \lambda_n \|v_{n+1} - v_n\|^2 + \gamma_n \|T v_n - v_{n+1}\|^2 \leq \alpha_n M, \quad (3.10)$$

for some positive constant M . On the other hand, we have from (3.8)

$$\begin{aligned} \|v_{n+1} - J_{\beta} v_{n+1}\| &\leq 2 \|v_{n+1} - J_{\beta_n} v_{n+1}\| \\ &\leq 2(\alpha_n \|u - v_{n+1}\| + \lambda_n \|v_n - v_{n+1}\| + \gamma_n \|T v_n - v_{n+1}\|), \end{aligned} \quad (3.11)$$

where the first inequality follows from Lemma 2.3. In order to prove the result, we consider two possible cases on the sequence (s_n) .

CASE 1: (s_n) is eventually decreasing (i.e., there exists $N \geq 0$ such that (s_n) is decreasing for all $n \geq N$). In this case, (s_n) is convergent. Therefore, we derive from (3.10) and (3.11)

$$\lim_{n \rightarrow \infty} \|v_{n+1} - J_\beta v_{n+1}\| = 0 = \lim_{n \rightarrow \infty} \|v_n - J_\beta v_n\|.$$

Note that A_β is a maximal monotone operator, thus so is A_β^{-1} , where A_β denotes the Yosida approximation of A . By Lemma 2.4, it follows that $\omega_w((v_n)) \subset A^{-1}(0) =: F$, where $\omega_w((v_n))$ denotes the set of weak cluster points of (v_n) . Now, extract a subsequence (v_{n_k}) of (v_n) converging weakly to some $y \in F$ such that

$$\limsup_{n \rightarrow \infty} \langle u - P_F u, v_n - P_F u \rangle = \lim_{k \rightarrow \infty} \langle u - P_F u, v_{n_k} - P_F u \rangle = \langle u - P_F u, y - P_F u \rangle \leq 0,$$

where $P_F u$ denotes the projection of u on F . Then from (3.9), we have

$$\|v_{n+1} - P_F u\|^2 \leq (1 - \alpha_n) \|v_n - P_F u\|^2 + 2\alpha_n \langle u - P_F u, v_{n+1} - P_F u \rangle,$$

and hence from Lemma 2.1, we get $v_n \rightarrow P_F u$ as desired.

CASE 2: (s_n) is not eventually decreasing, that is, there is a subsequence (s_{n_j}) of (s_n) such that $s_{n_j} \leq s_{n_{j+1}}$ for all $j \geq 0$. We therefore define an integer sequence $(\tau(n))$ as in Lemma 2.2 so that for all $n \geq n_0$, $s_{\tau(n)} \leq s_{\tau(n)+1}$ holds. In this case, we derive from (3.10) and (3.11)

$$\|v_{\tau(n)+1} - J_\beta v_{\tau(n)+1}\| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

The demiclosedness property of A_β^{-1} yields $\omega_w((v_{\tau(n)+1})) \subset F$. Consequently,

$$\limsup_{n \rightarrow \infty} \langle u - P_F u, v_{\tau(n)+1} - P_F u \rangle \leq 0.$$

Therefore, for $n \geq n_0$, we have from (3.9)

$$s_{\tau(n)+1} \leq \langle u - P_F u, v_{\tau(n)+1} - P_F u \rangle.$$

Passing to the limit in the above inequality, we arrive at $s_{\tau(n)+1} \rightarrow 0$ as $n \rightarrow \infty$. Thus, from (2.7) it follows that $s_n \rightarrow 0$ as $n \rightarrow \infty$. This completes the proof of the theorem. \square

Remark 3.1. *Theorem 3.1 refines [2, Theorem 5] and [3, Theorems 1-2]. Note that when T is the identity operator, then we recover many other results announced recently [1, 5, 6, 12, 13, 14].*

References

- [1] O. A. Boikanyo and G. Moroşanu, Inexact Halpern-type proximal point algorithms, *J. Glob. Optim.*, 51 (2011), no. 1, 1126.
- [2] O. A. Boikanyo and G. Moroşanu, Four parameter proximal point algorithms, *Non-linear Anal.* 74 (2011), no. 2, 544-555.

- [3] O. A. Boikanyo and G. Moroşanu, Multi parameter proximal point algorithms, *J. Nonlinear Conv. Anal.*, 13 (2012), no. 2, 221-231.
- [4] O. Güler, On the convergence of the proximal point algorithm for convex minimization, *SIAM J. Control Optim.* 29 (1991), 403-419.
<http://dx.doi.org/10.1137/0329022>
- [5] S. Kamimura and W. Takahashi, Approximating solutions of maximal monotone operators in Hilbert spaces, *J. Approx. Theory* 106 (2000), 226-240.
<http://dx.doi.org/10.1006/jath.2000.3493>
- [6] N. Lehdili and A. Moudafi, Combining the proximal algorithm and Tikhonov regularization, *Optimization* 37 (1996), 239-252.
<http://dx.doi.org/10.1080/02331939608844217>
- [7] P. E. Maingé, Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization, *Set-Valued Anal.*, 16 (2008) 899-912.
<http://dx.doi.org/10.1007/s11228-008-0102-z>
- [8] B. Martinet, Régularisation d'inéquations variationnelles par approximations successives, *Rev. Française Informat. Recherche Opérationnelle* 4 (1970), Ser. R-3, 154-158.
- [9] G. Moroşanu, *Nonlinear Evolution Equations and Applications*, Reidel, Dordrecht, 1988.
- [10] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.* 14 (1976), 877-898.
<http://dx.doi.org/10.1137/0314056>
- [11] M. V. Solodov and B. F. Svaiter, Forcing strong convergence of proximal point iterations in a Hilbert space, *Math. Program. Ser. A* 14 (2000), 189-202.
- [12] F. Wang and H. Cui, On the contraction-proximal point algorithms with multi-parameters, *J. Glob. Optim.*
<http://dx.doi.org/10.1007/s10898-011-9772-4>
- [13] H. K. Xu, Iterative algorithms for nonlinear operators, *J. London Math. Soc. (2)* 66 (2002), 240-256.
<http://dx.doi.org/10.1112/S0024610702003332>
- [14] H. K. Xu, A regularization method for the proximal point algorithm, *J. Glob. Optim.* 36 (2006), 115-125.
<http://dx.doi.org/10.1007/s10898-006-9002-7>
- [15] Y. Yao and M. A. Noor, On convergence criteria of generalized proximal point algorithms, *J. Comp. Appl. Math.* 217 (2008), 46-55.
<http://dx.doi.org/10.1016/j.cam.2007.06.013>