Journal of Global Optimization (2006) 36: 115–125
DOI 10.1007/s10898-006-9002-7

A Regularization Method for the Proximal Point Algorithm

HONG-KUN XU*
School of Mathematical Sciences, University of KawZulu-Natal, Westville Campus, Private Bag X54001, Durban 4000, South Africa (e-mail: xuhk@ukzn.ac.za)

(Received 27 July 2005; accepted 30 January 2006)

Abstract. A regularization method for the proximal point algorithm of finding a zero for a maximal monotone operator in a Hilbert space is proposed. Strong convergence of this algorithm is proved.

Mathematics Subject Classification (2000). 49J40, 47J20, 65J15

Key words: maximal monotone operator, projection, proximal point algorithm, regularization method, resolvent identity, strong convergence

1. Introduction

Let H be a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$, and let T be an operator with domain $D(T)$ and range $R(T)$ in H. Recall that T is *monotone* if its graph $G(T) := \{(x, y) \in H \times H : x \in D(T),\ y \in Tx\}$ is a monotone set in $H \times H$. That is, T is monotone if and only if,

$$(x, y),\ (x', y') \in G(T) \Rightarrow \langle x - x', y - y'\rangle \geqslant 0. \tag{1.1}$$

A monotone operator T is said to be *maximal* monotone if the graph $G(T)$ is not properly contained in the graph of any other monotone operator on H.

Recall that a zero of a maximal monotone operator T is a point $z \in D(T)$ such that $0 \in Tz$. We denote by S the set of all zeros of T; hence $S = T^{-1}(0)$. It is known that S is closed and convex. In the rest of this paper it is always assumed that S is nonempty so that the metric projection P_S from H onto S is well-defined.

One of the major problems in the theory of maximal operators is to find a point in the zero set S. A variety of problems, including convex programming and variational inequalities, can be formulated as finding a zero of maximal monotone operators. The proximal point algorithm of Rockafellar

*Supported in part by NRF.

[6] is recognized as a powerful and successful algorithm in finding a zero of maximal monotone operators. Starting from any initial guess $x^0 \in H$, this proximal point algorithm generates a sequence $\{x^k\}$ according to the inclusion:

$$x^k + e^k \in x^{k+1} + c_k T(x^{k+1}), \tag{1.2}$$

where $\{e^k\}$ is a sequence of errors and $\{c_k\}$ is a sequence of positive regularization parameters. Note that the algorithm (1.2) can be rewritten as

$$x^{k+1} = J_{c_k}^T (x^k + e^k), \tag{1.3}$$

where for $c > 0$, J_c^T denotes the resolvent of T given by

$$J_c^T = (I + cT)^{-1}$$

with I being the identity map on the space H.

Rockafellar's [6] proved the weak convergence of his algorithm (1.3) provided the regularization sequence $\{c_k\}$ remains bounded away from zero and the error sequence $\{e^k\}$ satisfies the condition

$$\sum_{k=0}^{\infty} \|e^k\| < \infty. \tag{1.4}$$

Güler's example [3] however shows that in an infinite-dimensional Hilbert space, Rochafellar's algorithm (1.3) has only weak convergence. So to have strong convergence, one has to modify the algorithm (1.3). Recently several authors proposed modifications of Rochafellar's proximal point algorithm (1.3) to have strong convergence. Solodov and Svaiter [7] initiated such investigation followed by Kamimura and Takahashi [4] (in which the work of [7] is extended to the framework of uniformly convex and uniformly smooth Banach spaces), and Xu [8]. In the work of [7] and [4], at each step of iteration, an additional projection onto the intersection of two half-spaces is needed.

Lehdili and Moudafi [5] combined the technique of the proximal map and the Tikhonov regularization to introduce the prox-Tikhonov method which generates the sequence $\{x^k\}$ by the algorithm

$$x^{k+1} = J_{\lambda_k}^{T_k} x^k, \quad k \geqslant 0, \tag{1.5}$$

where $T_k = \mu_k I + T$, $\mu_k > 0$, is viewed as a Tikhonov regularization of T. Note that T_k is strongly monotone; i.e., $\langle x - x', y - y' \rangle \geqslant \mu_k \|x - x'\|^2$ for all (x, y), $(x', y') \in G(T_k)$.

Using the technique of variational distance, Lehdili and Moudafi [5] were able to prove convergence theorems for the algorithm (1.5) and its perturbed version, under certain conditions imposed upon the sequences $\{\lambda_k\}$ and $\{\mu_k\}$.

It is the purpose of this paper to look at the algorithm (1.5) from another point of view. We use the technique of nonexpansive mappings to get convergence theorems for (the perturbed version of) the algorithm (1.5) under more relaxed conditions on the sequences $\{\lambda_k\}$ and $\{\mu_k\}$. Moreover, our approaches are simpler and more straightforward than those of Lehdili and Moudafi [5].

2. Preliminaries

Let T be a maximal monotone operator on a real Hilbert space H such that $S := T^{-1}(0) \neq \emptyset$. For $c > 0$, we use J_c^T to denote the resolvent of T; that is,

$$J_c^T = (I + cT)^{-1}.$$

It is known that the resolvent J_c is nonexpansive on H. Namely,

$$\|J_c^T x - J_c^T y\| \leqslant \|x - y\|, \quad \forall x, y \in H.$$

It is also easily seen that S is indeed the set of fixed points of J_c^T for all $c > 0$; that is, $S = \{x \in H : J_c^T x = x\}$ for $c > 0$.

We use P_K to denote the (metric or nearest point) projection from H onto K; that is, for each $x \in H$, $P_K x$ is the only point in K with the property

$$\|x - P_K x\| = \min_{v \in K} \|v - x\|.$$

It is known that P_K is nonexpansive and characterized by the inequality: Given $x \in H$ and $v \in K$; then $v = P_K x$ if and only if

$$\langle x - v, v - y \rangle \geqslant 0, \quad y \in K. \tag{2.1}$$

In order to facilitate our investigation in the next section we list some useful lemmas.

LEMMA 2.1 (cf. [8]). *Assume that* $\{a_k\}_{k=0}^{\infty}$ *is a sequence of nonnegative real numbers such that*

$$a_{k+1} \leqslant (1 - \lambda_k) a_k + \lambda_k \beta_k + \sigma_k, \quad k \geqslant 0,$$

where $\{\lambda_k\}$, $\{\beta_k\}$ *and* $\{\sigma_k\}$ *satisfy the conditions*

(i) $\lim_{k\to\infty}\lambda_k=0$;
(ii) $\sum_{k=0}^{\infty}\lambda_k=\infty$;
(iii) *either* $\limsup_{k\to\infty}\beta_k\leqslant 0$ *or* $\sum_{k=0}^{\infty}|\lambda_k\beta_k|<\infty$;
(iv) $\sigma_k\geqslant 0$ *for all* k *and* $\sum_{k=0}^{\infty}\sigma_k<\infty$.

Then $\lim_{k\to\infty}a_k=0$.

LEMMA 2.2 ([2]). *Let* C *be a closed convex subset of* H *and let* $f:C\to C$ *be a nonexpansive mapping such that* $Fix(f)\neq\emptyset$. *Assume* $\{x^k\}$ *is a sequence in* C *which weakly converges to* $x\in C$ *and for which* $\{(I-f)(x^k)\}$ *converges strongly to* $y\in H$. *Then* $(I-f)(x)=y$.

LEMMA 2.3 [1] (The Resolvent Identity). *For* $\lambda,\ \mu>0$, *there holds the identity:*

$$J_\lambda^T x=J_\mu^T\left(\frac{\mu}{\lambda}x+\left(1-\frac{\mu}{\lambda}\right)J_\lambda^T x\right),\quad x\in H.$$

LEMMA 2.4. *Assume* $c'\geqslant\bar{c}>0$. *Then* $\|J_{\bar{c}}^T x-x\|\leqslant 2\|J_{c'}^T x-x\|$ *for all* $x\in H$.

Proof. By the resolvent identity (Lemma 2.3), we have

$$\begin{aligned}\|J_{c'}x-J_{\bar{c}}x\|&=\left\|J_{\bar{c}}\left(\frac{\bar{c}}{c'}x+\left(1-\frac{\bar{c}}{c'}\right)J_{c'}x\right)-J_{\bar{c}}x\right\|\\&\leqslant\left(1-\frac{\bar{c}}{c'}\right)\|x-J_{c'}x\|\\&\leqslant\|x-J_{c'}x\|.\end{aligned}$$

Hence

$$\|J_{\bar{c}}x-x\|\leqslant\|J_{\bar{c}}x-J_{c'}x\|+\|J_{c'}x-x\|\leqslant 2\|J_{c'}x-x\|.\qquad\square$$

LEMMA 2.5. $\|x+y\|^2\leqslant\|x\|^2+2\langle y,x+y\rangle,\ x,\ y\in H$.

3. The Regularization Method

Nonexpansive mappings can be approximated by contractions. So we use contractions as the Tikhonov regularization of the resolvent J_c^T. Concretely consider the map V_t defined by

$$V_t x:=J_c^T((1-t)x+tu),\quad x\in H,\tag{3.1}$$

where $t\in(0,1)$, $c>0$ is a fixed constant, and $u\in H$ is a fixed point in H. Since J_c^T is nonexpansive, it is easily seen that V_t is a contraction; indeed we have

$$\|V_t x-V_t y\|\leqslant(1-t)\|x-y\|,\quad x,y\in H.$$

By Banach's contraction principle, V_t has a unique fixed point which is denoted by v_t (we suppress the dependence of v_t on c and u). Hence,

$$v_t = J_c^T((1-t)v_t + tu). \tag{3.2}$$

Our first result is a strong convergence result for $\{v_t\}$ which will subsequently be used in proving the strong convergence of another algorithm below.

THEOREM 3.1. *For any $c>0$ and $u \in H$, $\{v_t\}$ converges strongly to the projection of u onto S; that is, $s-\lim_{t\to 0} v_t = P_S u$. Moreover, this limit is attained uniformly for $c>0$.*

Proof. First by (3.2) we obtain

$$\frac{t}{c}(u-v_t) \in Tv_t. \tag{3.3}$$

Take any $p \in S$; hence $0 \in Tp$. By the monotonicity of T, we have

$$\langle u-v_t, v_t-p\rangle \geqslant 0. \tag{3.4}$$

This implies that

$$\|v_t-p\|^2 \leqslant \langle v_t-p, u-p\rangle \tag{3.5}$$

which in turn implies that

$$\|v_t-p\| \leqslant \|u-p\|, \quad p \in S. \tag{3.6}$$

In particular, $\{v_t\}$ is bounded. Now if $\{t_n\}$ is a sequence in (0,1) such that $t_n \to 0$ and $\{v_{t_n}\}$ is weakly convergent to $\tilde{v}$. By (3.3) and the maximal monotonicity of T, we have $0 \in T\tilde{v}$; that is, $\tilde{v} \in S$. Also by (3.6) we get $\|\tilde{v}-p\| \leqslant \|u-p\|$ for all $p \in S$. We therefore must have that $\tilde{v} = P_S u$. The arbitrariness of the subsequence $\{v_{t_n}\}$ of $\{v_t\}$ ensures that $\{v_t\}$ indeed converges weakly as $t \to 0$ to $P_S u$. We next show that the weak convergence is indeed strong. As a matter of fact, replacing p with $P_S u$ in (3.5) yields

$$\|v_t - P_S u\|^2 \leqslant \langle v_t - P_S u, u - P_S u\rangle \to 0 \tag{3.7}$$

as $v_t \to P_S u$ weakly. The uniformity of the limit for $c>0$ is easily obtained by noticing that c does not apparently appear in (3.4)–(3.7). $\square$

Inspired by Theorem 3.1, we now propose the following regularization for the proximal point algorithm (1.3):

$$x^{k+1} = J_{c_k}^T((1-t_k)x^k + t_k u + e^k), \quad k \geqslant 0, \tag{3.8}$$

where $\{c_k\}$ is a sequence of real positive numbers, $\{t_k\}$ is a sequence in $(0, 1)$, and $\{e^k\}$ is a sequence of errors. We will prove that the sequence $\{x^k\}$ converges strongly to $P_S u$ provided certain conditions imposed on $\{c_k\}$, $\{t_k\}$ and $\{e^k\}$ are met. We present two such results below.

THEOREM 3.2. *Assume*

(i) $\lim_{k\to\infty} t_k = 0$,
(ii) $\sum_{k=0}^{\infty} t_k = \infty$,
(iii) $t_{k+1} \leqslant \frac{c_{k+1}}{c_k}$ *for all k and either* $\lim_{k\to\infty} \frac{1}{t_k}\left|\frac{c_{k+1}t_k}{c_k t_{k+1}} - 1\right| = 0$ *or* $\sum_{k=0}^{\infty}\left|\frac{c_{k+1}t_k}{c_k t_{k+1}} - 1\right| < \infty$, *and*
(iv) $\sum_{k=0}^{\infty} \|e^k\| < \infty$.

Then the sequence $\{x^k\}$ *generated by the algorithm (3.8) strongly converges strongly to* $P_S u$.

Proof. For each k, let v^k be the unique fixed point of the contraction

$$x \mapsto J_{c_k}^T((1-t_k)x + t_k u).$$

By Theorem 3.1, we have that $v^k \to P_S u$. If we can show that $\|x^k - v^k\| \to 0$, then $x^k \to P_S u$ and the proof is complete. To see that $\|x^k - v^k\| \to 0$, we estimate $\|x^{k+1} - v^{k+1}\|$ as follows.

$$\|x^{k+1} - v^{k+1}\| \leqslant \|x^{k+1} - v^k\| + \|v^{k+1} - v^k\|. \tag{3.9}$$

Noticing (3.8) and the fact that

$$v^k = J_{c_k}^T((1-t_k)v^k + t_k u), \tag{3.10}$$

and using the nonexpansiveness of the resolvent $J_{c_k}^T$, we deduce that

$$\|x^{k+1} - v^k\| \leqslant (1-t_k)\|x^k - v^k\| + \|e^k\|. \tag{3.11}$$

To estimate $\|v^{k+1}-v^k\|$, we apply the resolvent identity (Lemma 2.3) to get

$$\begin{aligned}
\|v^{k+1}-v^k\| &= \|J^T_{c_{k+1}}((1-t_{k+1})v^{k+1}+t_{k+1}u)-J^T_{c_k}((1-t_k)v^k+t_ku)\| \\
&= \left\| J^T_{c_k}\left(\frac{c_k}{c_{k+1}}((1-t_{k+1})v^{k+1}+t_{k+1}u)\right.\right. \\
&\quad \left. +\left(1-\frac{c_k}{c_{k+1}}\right)J^T_{c_{k+1}}((1-t_{k+1})v^{k+1}+t_{k+1}u)\right) \\
&\quad \left. -J^T_{c_k}((1-t_k)v^k+t_ku)\right\| \\
&\leqslant \left\| \frac{c_k}{c_{k+1}}((1-t_{k+1})v^{k+1}+t_{k+1}u)+\left(1-\frac{c_k}{c_{k+1}}\right)v^{k+1}\right. \\
&\quad \left. -((1-t_k)v^k+t_ku)\right\| \\
&= \left\|\left(1-\frac{t_{k+1}c_k}{c_{k+1}}\right)(v^{k+1}-v^k)+\left(t_k-\frac{t_{k+1}c_k}{c_{k+1}}\right)(v^k-u)\right\|.
\end{aligned}$$

Consequently, we obtain

$$\|v^{k+1}-v^k\| \leqslant M\left|\frac{c_{k+1}t_k}{c_kt_{k+1}}-1\right|, \tag{3.12}$$

where $M>0$ is a constant such that $M \geqslant \|v^k-u\|$ for all $k \geqslant 0$. Substitute (3.11) and (3.12) into (3.9) to get

$$\|x^{k+1}-x^k\| \leqslant (1-t_k)\|x^k-v^k\|+M\left|\frac{c_{k+1}t_k}{c_kt_{k+1}}-1\right|+\|e^k\|. \tag{3.13}$$

By the conditions (i)–(iv) and Lemma 2.1, we conclude that $\|x^k-v^k\| \to 0$ and therefore $x^k \to P_Su$ strongly. □

The next result also gives a strong convergence result on the algorithm (3.8) but with different conditions on the sequences of $\{t_k\}$ and $\{c_k\}$.

THEOREM 3.3. *Assume*

(i) $\lim_{k\to\infty} t_k = 0$*,*
(ii) $\sum_{k=0}^{\infty} t_k = \infty$*,*
(iii) $\sum_{k=0}^{\infty} |t_{k+1}-t_k| < \infty$*,*
(iv) there are constants $0<\underline{c}\leqslant\overline{c}$ *such that* $\underline{c}\leqslant c_k\leqslant\overline{c}$ *for all* $k\geqslant 0$*, and* $\sum_{k=1}^{\infty}|c_{k+1}-c_k|<\infty$*,*
(v) $\sum_{k=0}^{\infty}\|e^k\|<\infty$*.*

Then the sequence $\{x^k\}$ *generated by the algorithm (3.8) strongly converges strongly to* $P_S u$.

Proof. First we show that $\{x^k\}$ is bounded. Indeed, take a $p \in S$ to get

$$\begin{aligned}\|x^{k+1} - p\| &= \|J_{c_k}^T((1-t_k)x^k + t_k u + e^k) - p\| \\ &\leqslant \|(1-t_k)x^k + t_k u + e^k - p\| \\ &\leqslant (1-t_k)\|x^k - p\| + t_k\|u - p\| + \|e^k\|.\end{aligned}$$

An induction gives that for all $k \geqslant 0$,

$$\|x^k - p\| \leqslant \max\{\|x^0 - p\|, \|u - p\|\} + \sum_{j=0}^{k-1} \|e^j\|.$$

In particular, $\{x^k\}$ is bounded.

We next estimate $\|x^{k+1} - x^k\|$. By (3.8) and Lemma 2.3 we have

$$\begin{aligned}\|x^{k+1} - x^k\| &= \Big\| J_{c_k}^T((1-t_k)x^k + t_k u + e^k) \\ &\quad - J_{c_{k-1}}^T((1-t_{k-1})x^{k-1} + t_{k-1}u + e^{k-1}) \Big\| \\ &= \Big\| J_{c_{k-1}}^T\left(\frac{c_{k-1}}{c_k}((1-t_k)x^k + t_k u + e^k)\right. \\ &\quad \left. + \left(1 - \frac{c_{k-1}}{c_k}\right) J_{c_k}^T((1-t_k)x^k + t_k u + e^k)\right) \\ &\quad - J_{c_{k-1}}^T((1-t_{k-1})x^{k-1} + t_{k-1}u + e^{k-1}) \Big\| \\ &\leqslant \Big\| \frac{c_{k-1}}{c_k}((1-t_k)x^k + t_k u + e^k) + \left(1 - \frac{c_{k-1}}{c_k}\right) x^{k+1} \\ &\quad - ((1-t_{k-1})x^{k-1} + t_{k-1}u + e^{k-1}) \Big\| \\ &\leqslant \frac{c_{k-1}}{c_k}\|(1-t_k)x^k + t_k u + e^k - ((1-t_{k-1})x^{k-1} + t_{k-1}u + e^{k-1})\| \\ &\quad + \left(1 - \frac{c_{k-1}}{c_k}\right)\|x^{k+1} - ((1-t_{k-1})x^{k-1} + t_{k-1}u + e^{k-1})\| \\ &= \frac{c_{k-1}}{c_k}\|(1-t_k)(x^k - x^{k-1}) + (t_k - t_{k-1})(u - x^{k-1}) + e^k - e^{k-1}\| \\ &\quad + \left(1 - \frac{c_{k-1}}{c_k}\right)\|x^{k+1} - ((1-t_{k-1})x^{k-1} + t_{k-1}u + e^{k-1})\|.\end{aligned}$$

It follows that, for an appropriate constant $M>0$,

$$\begin{aligned} c_k\|x^{k+1}-x^k\| \leqslant (1-t_k)c_{k-1}\|x^k-x^{k-1}\| + c_{k-1}M|t_k-t_{k-1}| \\ +\|e^k\|+\|e^{k-1}\|+M|c_k-c_{k-1}|. \end{aligned} \tag{3.14}$$

By Lemma 2.1 we conclude that

$$\|x^{k+1}-x^k\| \to 0. \tag{3.15}$$

Now by (3.8) and using the nonexpansivity of J_{c_k} we obtain

$$\begin{aligned} \|x^k-J_{c_k}^T x^k\| &\leqslant \|x^{k+1}-x^k\| + \|J_{c_k}^T((1-t_k)x^k+t_k u+e^k)-J_{c_k}^T x^k\| \\ &\leqslant \|x^{k+1}-x^k\|+t_k\|u-x^k\|+\|e^k\|. \end{aligned}$$

It follows from (3.15) that

$$\|x^k-J_{c_k}^T x^k\| \to 0. \tag{3.16}$$

Since, by Lemma 2.4, $\|x_k-J_{\underline{c}}^T x_k\| \leqslant 2\|x^k-J_{c_k}^T x^k\| \to 0$, we see that $\omega_w(z^k) \subset Fix(J_{\underline{c}}^T)=S$ by Lemma 2.2. To prove strong convergence, we take a subsequence $\{x^{k_j}\}$ of $\{x^k\}$ in such a way that

$$\limsup_{k\to\infty}\langle u-P_S u, x^k-P_S u\rangle = \lim_{j\to\infty}\langle u-P_S u, x^{k_j}-P_S u\rangle.$$

We may assume that $x^{k_j}\to q$ weakly. Now since $q\in S$ we deduce from the last equation and the characteristic inequality (2.1) that

$$\limsup_{k\to\infty}\langle u-P_S u, x^k-P_S u\rangle = \langle u-P_S u, q-P_S u\rangle \leqslant 0. \tag{3.17}$$

Finally we have by Lemma 2.5, for an appropriate constant $\gamma>0$,

$$\begin{aligned} \|x^{k+1}-P_S u\|^2 &= \|J_{c_k}^T((1-t_k)x^k+t_k u+e^k)-P_S u\|^2 \\ &\leqslant \|(1-t_k)(x^k-P_S u)+t_k(u-P_S u)+e^k\|^2 \\ &\leqslant \|(1-t_k)(x^k-P_S u)+t_k(u-P_S u)\|^2+\gamma\|e^k\| \\ &\leqslant (1-t_k)\|x^k-P_S u\|^2+2t_k\langle u-P_S u, x^{k+1}-P_S u\rangle+\gamma\|e^k\|. \end{aligned}$$

Noticing (3.17) and by Lemma 2.1, we obtain that $\|x^k-P_S u\|\to 0$. □

Remark 3.4. The prox-Tikhonov algorithm (1.5) of Lehdili and Moudafi [5] deals essentially with a special case of the algorithm (3.8). Indeed, if we set $u=0$ and $e^k=0$ for all k, then (3.8) becomes

$$x^{k+1}=J_{c_k}^T((1-t_k)x^k), \quad k\geqslant 0. \tag{3.18}$$

Setting

$$\lambda_k = \frac{c_k}{1-t_k} \quad \text{and} \quad \mu_k = \frac{t_k}{c_k}, \tag{3.19}$$

we can rewrite (3.18) as

$$x^{k+1} = J_{\lambda_k}^{T_k}(x^k), \quad k \geqslant 0, \tag{3.20}$$

where $T_k = \mu_k I + T$. (3.20) is the prox-Tikhonov algorithm (1.5) considered in [5]. Note that the approaches given in [5] are different from ours. The argument given in [5] depends heavily on the concept of the variational distance between two maximal monotone operators T and T' which is defined as

$$\delta_{\lambda,\rho}(T, T') := \sup_{\|x\| \leqslant \rho} \|J_\lambda^T x - J_\lambda^{T'} x\|$$

where $\rho \geqslant 0$ and $\lambda > 0$. (The distance $\delta_{\lambda,\rho}(T, T')$ may not be easily manipulated due to its involvement with the resolvent.)

Our argument is simpler and more straightforward. Indeed we use the technique of nonexpansive mappings (observe that each resolvent J_c^T is nonexpansive). We use contractions as the Tikhonov regularization of the resolvent J_c^T.

Moreover, the main convergence theorem in [5] requires the assumption on the sequence $\{\mu_k\}$ (in the case of a bounded $\{\lambda_k\}$):

$$\lim_{k\to\infty} \mu_k = 0, \quad \sum_{k=1}^{\infty} \mu_k = \infty, \quad \lim_{k\to\infty} \left| \frac{1}{\mu_{k+1}} - \frac{1}{\mu_k} \right| = 0. \tag{3.21}$$

It is easily seen that the natural choice of $\{\mu_k\}$,

$$\mu_k = \frac{1}{k}, \quad k \geqslant 1,$$

fails to satisfy (3.21). However, this choice of $\{\mu_k\}$ does satisfy the requirements (i)–(iii) on the sequence $\{t_k\}$ in our Theorem 3.3. In fact, if $\mu_k = 1/k$, then from the relation (3.19), we have

$$\begin{aligned} |t_{k+1} - t_k| &= \left| \frac{c_{k+1}}{k+1} - \frac{c_k}{k} \right| \\ &\leqslant \frac{|c_{k+1} - c_k|}{k+1} + \frac{c_k}{k(k+1)}. \end{aligned}$$

Hence condition (iii) of Theorem 3.3 holds provided $\{c_k\}$ satisfies some condition (e.g., $\{c_k\}$ is bounded and monotone, or $\{c_k\}$ is such that

$\sum_{k=1}^{\infty} |c_{k+1} - c_k| < \infty$; in particular, $c_k \equiv c$ is constant). In this sense, Theorem 3.3 improves the convergence result presented in [5].

References

1. Brezis, H. (1973), *Operateurs Maximaux Monotones et Semi-Groups de Contractions dans les Espaces de Hilbert*. North-Holland, Amsterdam.
2. Goebel, K. and Kirk, W.A. (1990), *Topics in Metric Fixed Point Theory*, *Cambridge Studies in Advanced Mathematics*, Vol. 28, Cambridge University Press.
3. Güler, O. (1991), On the convergence of the proximal point algorithm for convex optimization, *SIAM Journal of Control Optimization* 29, 403–419.
4. Kamimura, S. and Takahashi, W. (2003), Strong convergence of a proximal-type algorithm in a Banach space, *SIAM Journal of Optimization* 13(3), 938–945.
5. Lehdili, N. and Moudafi, A. (1996), Combining the proximal algorithm and Tikhonov regularization, *Optimization* 37, 239–252.
6. Rockafellar, R.T. (1976), Monotone operators and the proximal point algorithm, *SIAM Journal of Control Optimization* 14, 877–898.
7. Solodov, M.V. and Svaiter, B.F. (2000), Forcing strong convergence of proximal point iterations in a Hilbert space, *Mathematical Programming, Series A*, 87, 189–202.
8. Xu, H.K. (2002), Iterative algorithms for nonlinear operators, *Journal of the London Mathematical Society* 66, 240–256.