Bull. Aust. Math. Soc. **96** (2017), 162–170
doi:10.1017/S000497271600109X

ON THE CONVERGENCE RATE OF THE KRASNOSEL'SKIĬ–MANN ITERATION

SHIN-YA MATSUSHITA

(Received 19 October 2016; accepted 23 October 2016; first published online 5 January 2017)

Abstract

The Krasnosel'skiĭ–Mann (KM) iteration is a widely used method to solve fixed point problems. This paper investigates the convergence rate for the KM iteration. We first establish a new convergence rate for the KM iteration which improves the known big-O rate to little-o without any other restrictions. The proof relies on the connection between the KM iteration and a useful technique on the convergence rate of summable sequences. Then we apply the result to give new results on convergence rates for the proximal point algorithm and the Douglas–Rachford method.

2010 *Mathematics subject classification*: primary 47H09; secondary 47H05, 47H10, 47J25.

Keywords and phrases: Krasnosel'skiĭ–Mann iteration, convergence rate, nonexpansive, proximal point algorithm, Douglas–Rachford method.

1. Introduction

Let H be a real Hilbert space endowed with the inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$, and consider the following fixed point problem:

$$\text{Find } u \in H \text{ such that } T(u) = u, \tag{1.1}$$

where T is a nonexpansive mapping on H. Henceforth, the set of fixed points, $\mathrm{Fix}(T)$, of T is always assumed to be nonempty. An iterative procedure for solving (1.1) is the Krasnosel'skiĭ–Mann (KM) iteration, which was first proposed in [14, 17]. Consider the following KM iteration: for any initial point $x_0 \in H$,

$$x_{k+1} = (1-\alpha_k)x_k + \alpha_k T(x_k), \quad k = 0, 1, 2, \ldots, \tag{1.2}$$

where $\{\alpha_k\} \subset [0, 1]$ is a sequence of relaxation parameters. To simplify the notation, we let $\sigma_k := \sum_{j=0}^{k} \alpha_j(1-\alpha_j)$ $(k \in \mathbb{N})$. The KM iteration can be specified as the proximal

This work was supported in part by the Ministry of Education, Culture, Sports, Science, and Technology (grant number 16K05280).

point algorithm [18, 21], the Douglas–Rachford method [8, 16], the alternating direction method of multipliers [10, 16] and a three-operator splitting [6]. The convergence of (1.2) is well studied (see [1, 20]). In particular, under the assumption that $\lim_{k\to\infty}\sigma_k = \infty$, the sequence generated by (1.2) weakly converges to a point in $\mathrm{Fix}(T)$ [20, Theorem 2].

In this paper, we focus on analysing the convergence rate of $\{x_k\}$. Throughout, we use the quantity

$$\|(I-T)(x_k)\| \tag{1.3}$$

as a measure of the convergence rate since $\|(I-T)(x)\| = 0$ if and only if $T(x) = x$ and the property $\lim_{k\to\infty}\|(I-T)(x_k)\| = 0$ always holds when $\mathrm{Fix}(T) \neq \emptyset$. Recently, Cominetti *et al.* [3] showed that (1.3) converges to zero at a rate of $O(1/\sqrt{\sigma_k})$ (big-O) when $\lim_{k\to\infty}\sigma_k = \infty$. Similar big-$O$ results were also considered in [15]. Little-o rates of convergence for (1.3) were established by Davis and Yin [5] when $\epsilon > 0$ and $\{\alpha_k(1-\alpha_k)\} \subset (\epsilon, \infty)$. They showed that (1.3) converges to zero at a rate of $o(1/\sqrt{k+1})$ (little-o), which means that $\lim_{k\to\infty}\sqrt{k+1}\|(I-T)(x_k)\| = 0$. More precisely, $\|(I-T)(x_k)\|^2 = o(1/(k+1))$ [5, Theorem 1]. However, it is not clear whether the big-O rate in [3] can be improved to little-o.

The purpose of this paper is to show that (1.3) converges to zero at a rate of $o(1/\sqrt{\sigma_k})$ when $\lim_{k\to\infty}\sigma_k = \infty$. To achieve this goal, we consider a useful technique on the convergence rates of summable sequences which appeared in [7, Lemma 3.2]. We show that this technique can be applied to the KM iteration and establish that $\|(I-T)(x_k)\| = o(1/\sqrt{\sigma_k})$. This result improves the existing convergence rate [3] without any other restrictions.

The KM iteration generalises several other methods. In particular, we apply our result to analyse the proximal point algorithm and the Douglas–Rachford method. Recently, some results on convergence rates for these methods were established in [4, 11] by using constant relaxation parameters. We establish improved convergence rates for the proximal point algorithm and the Douglas–Rachford method under mild assumptions.

The rest of this paper is organised as follows. In Section 2, some preliminaries are presented. In Section 3, we improve the convergence rate of the KM iteration. Then, we discuss convergence rates for the proximal point algorithm and the Douglas–Rachford method in Sections 4 and 5, respectively.

2. Preliminaries

The following notation will be used in this paper: $\mathbb{R}$ denotes the set of real numbers; $\mathbb{N}$ denotes the set of nonnegative integers; H denotes a real Hilbert space: for any $x, y \in H$, $\langle x, y\rangle$ denotes the inner product of x and y and, for any $z \in H$, $\|z\|$ denotes the norm of z, that is, $\|z\| = \sqrt{\langle z, z\rangle}$; for any $C \subset H$ and mapping $U : C \to C$, $\mathrm{Fix}(U)$ denotes the fixed point set of U, that is, $\mathrm{Fix}(U) = \{x \in C : U(x) = x\}$; for any set-valued mapping $A : H \to 2^H$, $D(A) = \{x \in H : A(x) \neq \emptyset\}$ denotes the domain of A,

https://doi.org/10.1017/S000497271600109X Published online by Cambridge University Press

$R(A) = \bigcup\{A(x) : x \in D(A)\}$ denotes the range of A and $G(A) = \{(x, x^*) : x^* \in A(x)\}$ denotes the graph of A; and the set of zero points of A is denoted by $A^{-1}(0)$, that is, $A^{-1}(0) = \{z \in D(A) : 0 \in A(z)\}$.

A mapping $U : C \to C$ is said to be:

(i) *firmly nonexpansive* if

$$\|U(x) - U(y)\|^2 \leq \langle x - y, U(x) - U(y)\rangle \quad (x, y \in C);$$

(ii) *nonexpansive* if

$$\|U(x) - U(y)\| \leq \|x - y\| \quad (x, y \in C).$$

A set-valued mapping $A : H \to 2^H$ is said to be:

(i) *monotone* if

$$\langle x - y, x^* - y^*\rangle \geq 0 \quad ((x, x^*), (y, y^*) \in G(A));$$

(ii) *maximal monotone* if A is monotone and $A = B$ whenever $B : H \to 2^H$ is a monotone mapping such that $G(A) \subset G(B)$.

The maximal monotonicity of A implies that $R(I + rA) = H$ for all $r > 0$. To simplify the notation in this paper, we let $r := 1$. Then we can define the *resolvent* J_A of A by

$$J_A(x) = \{z \in H : x \in z + A(z)\} = (I + A)^{-1}(x)$$

for all $x \in H$. The *reflected resolvent* R_A of J_A is defined by $2J_A - I$ (see [1, 23]).

Let $A : H \to 2^H$ and $B : H \to 2^H$ be maximal monotone set-valued mappings. Then $J_A : H \to H$ is firmly nonexpansive and $R_A : H \to H$ is nonexpansive and

$$\mathrm{Fix}(J_A) = \mathrm{Fix}(R_A) = A^{-1}(0); \tag{2.1}$$

$$J_A(\mathrm{Fix}(R_B R_A)) = (A + B)^{-1}(0) \subset \mathrm{Fix}(R_B R_A); \tag{2.2}$$

$$\tfrac{1}{2}(I + R_B R_A) = J_B(2J_A - I) + (I - J_A). \tag{2.3}$$

See [1, 2] for more details.

The following result will be the key to deducing convergence rates for the KM iteration.

Lemma 2.1 [7, Lemma 3.2]. *Let $\{b_k\}$, $\{c_k\}$ be sequences of positive numbers. Assume that the sequence $\{b_k\}$ is nonsummable, the sequence $\{c_k\}$ is decreasing and*

$$\sum_{i=0}^{\infty} b_i c_i < \infty.$$

Then

$$c_k = o\left(1 \Big/ \sum_{i=0}^{k} b_i\right),$$

where the o-notation means that $s_k = o(1/t_k)$ if and only if $\lim_{k\to\infty} s_k t_k = 0$.

Remark 2.2. Dong [7] used Lemma 2.1 to analyse the proximal point algorithm.

https://doi.org/10.1017/S000497271600109X Published online by Cambridge University Press

3. Krasnosel'skiĭ–Mann iteration

In this section, we study the convergence rate for the KM iteration in a Hilbert space. Using Lemma 2.1, we prove the following result.

THEOREM 3.1. *Let C be a nonempty closed convex subset of H, let $T : C \to C$ be a nonexpansive mapping such that* $\mathrm{Fix}(T) \neq \emptyset$ *and let* $\{x_k\}$ *be the sequence generated by (1.2), where* $x_0 \in C$, *and* $\{\alpha_k\}$ *is a sequence in* $[0, 1]$ *such that* $\sigma_k := \sum_{j=0}^{k} \alpha_j(1 - \alpha_j)$ *for* $k \in \mathbb{N}$ *and* $\lim_{k\to\infty} \sigma_k = \infty$. *Then the convergence rate estimate*

$$\|(I - T)(x_k)\| = o(1/\sqrt{\sigma_k})$$

holds, that is, $\lim_{k\to\infty} \sqrt{\sigma_k}\,\|(I - T)(x_k)\| = 0$.

PROOF. Let $u \in F(T)$. By virtue of [1, Theorem 5.14], the following properties hold.

(i) For any $k \in \mathbb{N}$,

$$\|x_{k+1} - u\|^2 \leq \|x_k - u\|^2 - \alpha_k(1 - \alpha_k)\|(I - T)(x_k)\|^2. \tag{3.1}$$

(ii) The sequence $\{\|(I - T)(x_k)\|\}$ is decreasing, that is, for any $k \in \mathbb{N}$,

$$\|(I - T)(x_{k+1})\| \leq \|(I - T)(x_k)\|.$$

Rearranging (3.1) as $\alpha_k(1 - \alpha_k)\|(I - T)(x_k)\|^2 \leq \|x_k - u\|^2 - \|x_{k+1} - u\|^2$ and summing from $j = 0$ to l implies that

$$\sum_{j=0}^{l} \alpha_j(1 - \alpha_j)\|(I - T)(x_j)\|^2 \leq \|x_0 - u\|^2.$$

By taking $l \to \infty$, we see that

$$\sum_{j=0}^{\infty} \alpha_j(1 - \alpha_j)\|(I - T)(x_j)\|^2 < \infty.$$

Since $\lim_{k\to\infty} \sigma_k = \infty$, the assumptions of Lemma 2.1 hold with $b_k := \alpha_k(1 - \alpha_k)$ and $c_k := \|(I - T)(x_k)\|^2$ and hence

$$\|(I - T)(x_k)\|^2 = o(1/\sigma_k).$$

We can therefore conclude that $\|(I - T)(x_k)\| = o(1/\sqrt{\sigma_k})$. □

REMARK 3.2.

(a) Theorem 3.1 improves the known big-O rate in [3, Proposition 11] to little-o without any other restrictions.

(b) Let $\epsilon > 0$. The condition $\{\alpha_k(1 - \alpha_k)\} \subset (\epsilon, \infty)$ implies that $\lim_{k\to\infty} \sigma_k = \infty$. But the reverse implication does not hold. An example of $\{\alpha_k\}$ satisfying the conditions $\sum_{j=0}^{\infty} \alpha_j(1 - \alpha_j) = \infty$ and $\inf_k \alpha_k(1 - \alpha_k) = 0$ is $\alpha_k := 1/(k + 1)$.

(c) Under the assumptions that $\epsilon > 0$ and $\{\alpha_k(1 - \alpha_k)\} \subset (\epsilon, \infty)$,

$$\sqrt{\epsilon(k + 1)}\|(I - T)(x_k)\| \leq \sqrt{\sigma_k}\|(I - T)(x_k)\|,$$

so the $o(1/\sqrt{k + 1})$ rate in [5, Theorem 1] follows from Theorem 3.1.

4. Proximal point algorithm

We consider the convergence rates for the proximal point algorithm. Theorem 3.1 can be applied directly to derive new convergence rates.

The proximal point algorithm is an algorithm for solving the inclusion problem, $0 \in A(u)$, where A is a maximal monotone set-valued mapping on H. This algorithm was first introduced by Martinet [18] and further developed by Rockafellar [21]. It is known that the sequence generated by the proximal point algorithm weakly converges to a point in $A^{-1}(0)$ under mild assumptions in the infinite-dimensional Hilbert spaces.

The framework of the generalised proximal point algorithm for a maximal monotone set-valued mapping A is as follows: given $x_0 \in H$, set

$$x_{k+1} = (1 - \beta_k)x_k + \beta_k J_A(x_k), \quad k = 0, 1, 2, \ldots, \tag{4.1}$$

where $\{\beta_k\} \subset [0, 2]$ is a sequence of relaxation parameters and J_A is the resolvent of A. The convergence of (4.1) under some conditions has been discussed in [2, 9, 12, 13, 19, 21]. Using the definition of R_A, we can write (4.1) equivalently as

$$x_{k+1} = \left(1 - \frac{\beta_k}{2}\right)x_k + \frac{\beta_k}{2} R_A(x_k), \quad k = 0, 1, 2, \ldots. \tag{4.2}$$

To simplify the notation, we let

$$\sigma_k := \sum_{j=0}^{k} \frac{\beta_j}{2}\left(1 - \frac{\beta_j}{2}\right) \quad (k \in \mathbb{N}). \tag{4.3}$$

Since R_A is nonexpansive, (4.2) can be viewed as the KM iteration and $\{x_k\}$ weakly converges to a point in $\mathrm{Fix}(R_A)$ $(= A^{-1}(0))$ when $\lim_{k\to\infty} \sigma_k = \infty$ and $\mathrm{Fix}(R_A) \neq \emptyset$.

Remark 4.1. Since J_A is (firmly) nonexpansive, (4.1) can also be viewed as the KM iteration. In order to apply the KM iteration to (4.1), it is necessary to restrict $\{\beta_k\}$ in $[0, 1]$.

Using Theorem 3.1, we obtain new estimates of convergence rates for (4.1).

Theorem 4.2. *Let A be a maximal monotone set-valued mapping on H such that $A^{-1}(0) \neq \emptyset$, let $\{x_k\}$ be the sequence generated by (4.1) and define σ_k by (4.3). If $\lim_{k\to\infty} \sigma_k = \infty$, then*

$$\|(I - R_A)(x_k)\| = o(1/\sqrt{\sigma_k}) \tag{4.4}$$

and

$$\|(I - J_A)(x_k)\| = o(1/\sqrt{\sigma_k}). \tag{4.5}$$

Proof. Together, (2.1) and $(A)^{-1}(0) \neq \emptyset$ imply that $\mathrm{Fix}(R_A) \neq \emptyset$. Using (4.2), (4.4) follows directly from Theorem 3.1. Since $I - R_A = 2(I - J_A)$, (4.5) follows from (4.4). □

https://doi.org/10.1017/S000497271600109X Published online by Cambridge University Press

REMARK 4.3. The estimate (4.5) is better than the corresponding result in [4, Theorem 3.1]. Using constant relaxation parameters, Corman and Yuan have analysed convergence rates for (4.1). Under the assumptions that $\beta \in (0, 2)$ and $\beta_k := \beta$ $(k \in \mathbb{N})$, we have $\lim_{k\to\infty} \sigma_k = \infty$ and

$$\sqrt{\frac{\beta}{2}\Big(1 - \frac{\beta}{2}\Big)(k+1)}\,\|(I - J_A)(x_k)\| = \sqrt{\sigma_k}\|(I - J_A)(x_k)\|.$$

Thus, the $o(1/\sqrt{k+1})$ rate in [4, Theorem 3.1] follows from Theorem 4.2.

5. Douglas–Rachford method

We next consider the Douglas–Rachford (DR) method. Theorem 3.1 can also be applied to improve the convergence rate for the DR method.

The DR method is a fundamental algorithm for solving the inclusion problem $0 \in (A + B)(u)$, where A and B are maximal monotone set-valued mappings on H. This method was first introduced by Douglas and Rachford [8] and further developed by Lions and Mercier [16] and Eckstein and Bertsekas [9].

The framework of the DR method for maximal monotone set-valued mappings A and B is as follows: given $x_0 \in H$, set

$$x_{k+1} = x_k + \gamma_k(J_B R_A(x_k) - J_A(x_k)), \quad k = 0, 1, 2, \ldots, \tag{5.1}$$

where $\{\gamma_k\} \subset [0, 2]$ is a sequence of relaxation parameters. Under appropriate assumptions, the sequence generated by (5.1) weakly converges to a point $x^* \in H$ such that $x^* \in \mathrm{Fix}(R_B R_A)$ and $J_A(x^*) \in (A + B)^{-1}(0)$ (see [1, 2, 9, 16]). Using (2.3), we can write (5.1) in the equivalent form

$$x_{k+1} = \Big(1 - \frac{\gamma_k}{2}\Big)x_k + \frac{\gamma_k}{2} R_B R_A(x_k), \quad k = 0, 1, 2, \ldots. \tag{5.2}$$

To simplify the notation, we let

$$\sigma_k := \sum_{j=0}^{k} \frac{\gamma_j}{2}\Big(1 - \frac{\gamma_j}{2}\Big) \quad (k \in \mathbb{N}). \tag{5.3}$$

Since $R_B R_A$ is nonexpansive, (5.2) can be viewed as the KM iteration and $\{x_k\}$ weakly converges to a point in $\mathrm{Fix}(R_B R_A)$ when $\lim_{k\to\infty} \sigma_k = \infty$ and $\mathrm{Fix}(R_B R_A) \neq \emptyset$.

Note that it is not guaranteed that the sequence $\{x_k\}$ generated by (5.2) weakly converges to a point in $(A + B)^{-1}(0)$. Svaiter [22] showed that the shadow sequence $\{J_A(x_k)\}$ weakly converges to a point in $(A + B)^{-1}(0)$ when $\gamma_k := 1$ $(k \in \mathbb{N})$ and $(A + B)^{-1}(0) \neq \emptyset$. By using the demiclosed principle, Bauschke and Combettes [1, Proposition 25.17] showed weak convergence of $\{J_A(x_k)\}$ when $\sum_{j=0}^{\infty} \gamma_j(2 - \gamma_j) = \infty$.

On the other hand, the worst-case convergence rate of $\{J_A(x_k)\}$ has been recently analysed. He and Yuan [11, Theorem 3.1] showed that $\|J_A(x_{k+1}) - J_A(x_k)\|$ converges to zero at a rate of $O(1/\sqrt{k})$ when H is finite dimensional, $\gamma \in (0, 2)$ and $\gamma_k := \gamma (k \in \mathbb{N})$.

https://doi.org/10.1017/S000497271600109X Published online by Cambridge University Press

They used the quantity $\|J_A(x_{k+1}) - J_A(x_k)\|$ to estimate the convergence rate, since $\{J_A(x_k)\}$ strongly converges to a point in $(A + B)^{-1}(0)$. In infinite-dimensional spaces, however, the strong convergence of $\{J_A(x_k)\}$ is not guaranteed.

In order to estimate the convergence rate of $\{J_A(x_k)\}$, the following result is useful.

LEMMA 5.1. *If* $\|J_A(x) - J_B R_A(x)\| = 0$*, then* $J_A(x) \in (A + B)^{-1}(0)$.

PROOF. Using the definition of the resolvent and $J_A(x) = J_B R_A(x)$,

$$x \in J_A(x) + A(J_A(x)) \quad \text{and} \quad R_A \in J_A(x) + B(J_A(x)).$$

From these two inclusions, $0 \in A(J_A(x)) + B(J_A(x))$ and hence $J_A(x) \in (A + B)^{-1}(0)$. □

REMARK 5.2. Let $\{x_k\}$ be the sequence generated by the DR method. By applying Lemma 5.1, if $\|J_A(x_k) - J_B R_A(x_k)\| = 0$, then $J_A(x_k)$ is in $(A + B)^{-1}(0)$, so the quantity $\|J_A(x_k) - J_B R_A(x_k)\|$ is a convenient estimator for the convergence rate of $\{J_A(x_k)\}$.

Using Theorem 3.1, we obtain new estimates of convergence rates for $\{x_k\}$ and $\{J_A(x_k)\}$.

THEOREM 5.3. *Let A and B be maximal monotone set-valued mappings on H such that $(A + B)^{-1}(0) \neq \emptyset$, let $\{x_k\}$ be the sequence generated by (5.1) and define σ_k by (5.3). If $\lim_{k\to\infty} \sigma_k = \infty$, then*

$$\|(I - R_B R_A)(x_k)\| = o(1/\sqrt{\sigma_k}) \tag{5.4}$$

and

$$\|J_A(x_k) - J_B R_A(x_k)\| = o(1/\sqrt{\sigma_k}). \tag{5.5}$$

PROOF. Together, (2.2) and $(A + B)^{-1}(0) \neq \emptyset$ imply that $\mathrm{Fix}(R_B R_A) \neq \emptyset$. Using (5.2), (5.4) follows directly from Theorem 3.1.

From Lemma 5.1, we can use the quantity $\|J_A(x_k) - J_B R_A(x_k)\|$ to measure the proximity of $\{J_A(x_k)\}$ to a point in $(A + B)^{-1}(0)$. From (2.3),

$$(I - R_B R_A)(x_k) = 2(J_B R_A - J_A)(x_k)$$

and hence

$$\|(I - R_B R_A)(x_k)\| = 2\|(J_B R_A - J_A)(x_k)\|.$$

Therefore, (5.5) follows from (5.4). □

REMARK 5.4.

(a) The estimate (5.4) is better than the corresponding result in [11] even in the special case considered there. Indeed, using constant relaxation parameters (for example, $\gamma \in (0, 2)$ and $\gamma_k := \gamma$ $(k \in \mathbb{N})$), (5.4) is equivalent to

$$\|(I - R_B R_A)(x_k)\| = o(1/\sqrt{k + 1}).$$

This rate improves on the known big-O rate [11, Theorem 3.1].

(b) To the best of our knowledge, (5.5) is a new estimate for the convergence rate of $\{J_A(x_k)\}$.

https://doi.org/10.1017/S000497271600109X Published online by Cambridge University Press

Acknowledgements

The author is grateful to Professors W. Takahashi of Tokyo Institute of Technology, D. Kuroiwa of Shimane University and Li Xu of Akita Prefectural University for their helpful support.

References

[1] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces* (Springer, New York, 2011).
[2] P. L. Combettes, 'Solving monotone inclusions via compositions of nonexpansive averaged operators', *Optimization* **53** (2004), 475–504.
[3] R. Cominetti, J. A. Soto and J. Vaisman, 'On the rate of convergence of Krasnosel'skiĭ–Mann iterations and their connection with sums of Bernoullis', *Israel J. Math.* **199** (2014), 757–772.
[4] E. Corman and X. Yuan, 'A generalized proximal point algorithm and its convergence rate estimate', *SIAM J. Optim.* **24** (2014), 1614–1638.
[5] D. Davis and W. Yin, 'Convergence rate analysis of several splitting schemes', in: *Splitting Methods in Communication and Imaging, Science and Engineering* (eds. R. Glowinski, S. Osher and W. Yin) (Springer, New York), to appear.
[6] D. Davis and W. Yin, 'A three-operator splitting scheme and its optimization applications', Preprint, 2015, arXiv:1504.01032.
[7] Y. Dong, 'Comments on 'the proximal point algorithm revisited'', *J. Optim. Theory Appl.* **116** (2015), 343–349.
[8] J. Douglas and H. H. Rachford, 'On the numerical solution of heat conduction problems in two and three space variables', *Trans. Amer. Math. Soc.* **82** (1956), 421–439.
[9] J. Eckstein and D. P. Bertsekas, 'On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators', *Math. Program.* **55** (1992), 293–318.
[10] R. Glowinski and A. Marroco, 'Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problémes de Dirichlet non linéaires', *Rev. Fr. Autom. Inform. Rech. Oper.* **9** (1975), 41–76.
[11] B. S. He and X. M. Yuan, 'On the convergence rate of Douglas–Rachford operator splitting method', *Math. Program.* **153** (2015), 715–722.
[12] S. Kamimura, F. Kohsaka and W. Takahashi, 'Weak and strong convergence theorems for maximal monotone operators in a Banach space', *Set-Valued Anal.* **12** (2004), 417–429.
[13] S. Kamimura and W. Takahashi, 'Approximating solutions of maximal monotone operators in Hilbert spaces', *J. Approx. Theory* **106** (2000), 226–240.
[14] M. A. Krasnosel'skiĭ, 'Two remarks on the method of successive approximations', *Uspekhi Mat. Nauk* **10** (1955), 123–127.
[15] J. Liang, J. Fadili and G. Peyré, 'Convergence rates with inexact nonexpansive operators', *Math. Program.* **159** (2016), 403–434.
[16] P. L. Lions and B. Mercier, 'Splitting algorithms for the sum of two nonlinear operators', *SIAM J. Numer. Anal.* **16** (1979), 964–979.
[17] W. R. Mann, 'Mean value methods in iteration', *Proc. Amer. Math. Soc.* **4** (1953), 506–510.
[18] B. Martinet, 'Regularisation d'inequations variationnelles par approximations successives', *Rev. Fr. Autom. Inform. Rech. Oper.* **4** (1970), 154–159.
[19] S. Matsushita and W. Takahashi, 'Weak and strong convergence theorems for relatively nonexpansive mappings in Banach spaces', *Fixed Point Theory Appl.* **2004** (2004), 37–47.
[20] S. Reich, 'Weak convergence theorems for nonexpansive mappings in Banach spaces', *J. Math. Anal. Appl.* **67** (1979), 274–276.
[21] R. T. Rockafellar, 'Monotone operators and the proximal point algorithm', *SIAM J. Control Optim.* **14** (1976), 877–898.

[22] B. F. Svaiter, 'On weak convergence of the Douglas–Rachford method', *SIAM J. Control Optim.* **49** (2011), 280–287.

[23] W. Takahashi, *Nonlinear Functional Analysis. Fixed Point Theory and its Applications* (Yokohama Publishers, Yokohama, 2000).

SHIN-YA MATSUSHITA, Department of Electronics and Information Systems,
Akita Prefectural University, Yuri-Honjo, Akita 015-0055, Japan
e-mail: matsushita@akita-pu.ac.jp

https://doi.org/10.1017/S000497271600109X Published online by Cambridge University Press