Math. Program., Ser. A (2019) 175:241–262
https://doi.org/10.1007/s10107-018-1240-1

FULL LENGTH PAPER

Rates of convergence for inexact Krasnosel'skii–Mann iterations in Banach spaces

Mario Bravo[1] · **Roberto Cominetti**[2] · **Matías Pavez-Signé**[3]

Received: 14 June 2017 / Accepted: 17 January 2018 / Published online: 30 January 2018

Abstract We study the convergence of an inexact version of the classical Krasnosel'skii–Mann iteration for computing fixed points of nonexpansive maps. Our main result establishes a new metric bound for the fixed-point residuals, from which we derive their rate of convergence as well as the convergence of the iterates towards a fixed point. The results are applied to three variants of the basic iteration: infeasible iterations with approximate projections, the Ishikawa iteration, and diagonal Krasnosels'kii–Mann schemes. The results are also extended to continuous time in order to study the asymptotics of nonautonomous evolution equations governed by nonexpansive operators.

This work was partially supported by Núcleo Milenio Información y Coordinación en Redes ICM/FIC RC130003. Mario Bravo was partially funded by FONDECYT 11151003. Roberto Cominetti and Matías Pavez-Signé gratefully acknowledge the support provided by FONDECYT 1130564 and FONDECYT 1171501.

✉ Roberto Cominetti
roberto.cominetti@uai.cl

Mario Bravo
mario.bravo.g@usach.cl

Matías Pavez-Signé
mpavez@dim.uchile.cl

1 Departamento de Matemática y Ciencia de la Computación, Universidad de Santiago de Chile, Alameda Libertador Bernardo O'higgins 3363, Santiago, Chile

2 Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Diagonal Las Torres 2640, Santiago, Chile

3 Departamento de Ingeniería Matemática, Universidad de Chile, Beauchef 851, Santiago, Chile

Keywords Nonexpansive maps · Fixed point iterations · Rates of convergence · Evolution equations

Mathematics Subject Classification 47H09 · 47H10 · 65J08 · 65K15 · 60J10

1 Introduction

Let $T : C \to C$ be a nonexpansive map defined on a closed convex domain C in a Banach space $(X, \|\cdot\|)$. The Krasnosel'skii–Mann iteration approximates a fixed point of T by the sequential averaging process

$$x_{n+1} = (1-\alpha_{n+1})\, x_n + \alpha_{n+1}\, T x_n, \tag{KM}$$

where $x_0 \in C$ is an initial guess and $\alpha_n \in [0, 1]$ is a given sequence of scalars.

This iteration, introduced by Krasnosel'skii [20] and Mann [23], arises frequently in convex optimization as many algorithms can be cast in this framework. This is the case of the gradient method for convex functions with Lipschitz gradient [28], the proximal point method [24,29], as well as different decomposition methods such as the forward-backward splitting method [25,26], the alternating direction method of multipliers ADMM [14], the Douglas–Rachford splitting [13], and the Peaceman–Rachford splitting [27]. For a comprehensive survey of these methods and their numerous applications we refer to [4]. Note that (KM) also arises when discretizing the evolution equation $u'(t) + [I - T]u = 0$ (see e.g. [2,6]), so that the results for (KM) admit natural extensions to continuous time.

A central issue when studying the convergence of the iterates x_n to a fixed point of T is to prove the strong convergence of the residuals $\|x_n - T x_n\| \to 0$, a property known as *asymptotic regularity* [9,10]. For a historical account of results in this area we refer to [2,3]. In the case of a bounded domain C, an explicit estimate for the residual was conjectured in [2] and recently confirmed in [12], namely

$$\|x_n - T x_n\| \le \frac{\operatorname{diam}(C)}{\sqrt{\pi \sum_{k=1}^{n} \alpha_k (1-\alpha_k)}}, \tag{1}$$

where the constant $1/\sqrt{\pi}$ is known to be tight (see [5]). This inequality implies that asymptotic regularity holds as soon as $\sum_{k=1}^{\infty} \alpha_k (1-\alpha_k) = \infty$. The bound can also be used to estimate the number of iterations required to attain any prescribed accuracy, as well as to establish the rate of convergence of the residuals. For instance, if α_n remains away from 0 and 1 the bound yields $\|x_n - T x_n\| = O(1/\sqrt{n})$, whereas for $\alpha_n = 1/n$ one gets an order $O(1/\sqrt{\ln n})$.

When the operator values Tx can only be computed up to some precision, one is naturally led to consider the inexact iteration

$$x_{n+1} = (1-\alpha_{n+1})\, x_n + \alpha_{n+1}\, (T x_n + e_{n+1}), \tag{IKM}$$

where e_{n+1} can be interpreted as an error in the computation of Tx_n, or as a perturbation of the iteration. Note that the evaluation of Tx_n requires $x_n \in C$ so that (IKM) assumes implicitly that the iterates remain in C.

This inexact iteration was used by Liu [22] to study the equation $Sx = f$, restated as a fixed point of $Tx = f + x - Sx$, with S demicontinuous and strongly accretive on a uniformly smooth space. Liu proved the strong convergence of the iterates assuming that Tx_n remains bounded, $\sum_{k\geq 1} \|e_k\| < \infty$, and $\alpha_n \to 0$ with $\sum_{k\geq 1} \alpha_k = \infty$. Weak convergence of (IKM) was also established for T nonexpansive, first on Hilbert spaces [11] and then on uniformly convex spaces [18], provided that $\mathrm{Fix}(T) \neq \emptyset$ and $\sum_{k\geq 1} \alpha_k(1-\alpha_k) = \infty$ with $\sum_{k\geq 1} \alpha_k \|e_n\| < \infty$. The rate of convergence of (IKM) was recently studied in a Hilbert setting by Liang et al. [21], proving that $\|x_n - Tx_n\| = O(1/\sqrt{n})$ under the stronger summability condition $\sum_{k\geq 1} k\|e_k\| < \infty$ and with α_n bounded away from 0 and 1. The proof exploits the Hilbert structure and does not seem to carry over to general Banach spaces.

1.1 Our contribution

The main result in this paper is an extension of the bound (1) which holds for the inexact iteration (IKM) in general normed spaces. From this extended bound we draw a number of consequences on the convergence of the iterates and the rate of convergence of the fixed point residuals, and we derive continuous time analogs for the asymptotics of evolution equations governed by nonexpansive operators.

In all what follows we denote $\epsilon_n \geq \|e_n\|$ a bound for the errors, and we let

$$\tau_n = \sum_{k=1}^{n} \alpha_k (1-\alpha_k).$$

We also consider the function $\sigma : [0, \infty) \to \mathbb{R}$ defined by

$$\sigma(y) = \min\{1, 1/\sqrt{\pi y}\}.$$

With these notations, our main result can be stated as follows.

Theorem 1 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence generated by (IKM) and assume that*

$$\exists\, \kappa \geq 0 \; such\; that\; x_n \in C \; and \; \|Tx_n - x_0\| \leq \kappa \; for\; all \; \forall\, n \in \mathbb{N}. \tag{H$_0$}$$

Then, for all $n \in \mathbb{N}$ we have

$$\|x_n - Tx_n\| \leq \kappa\, \sigma(\tau_n) + \sum_{i=1}^{n} 2\, \alpha_i \epsilon_i\, \sigma(\tau_n - \tau_i) + 2\, \epsilon_{n+1}. \tag{2}$$

In particular, if $\tau_n \to \infty$ and $\epsilon_n \to 0$ with $\sum_{k\geq 1} \alpha_k \epsilon_k < \infty$, then $\|x_n - Tx_n\| \to 0$.

Remark 1 Clearly, in the exact case with $\epsilon_n \equiv 0$ the bound (2) yields (1). As already mentioned, in [5] the bound (1) was shown to be tight, and is attained for constant sequences $\alpha_n \equiv \alpha$ with α close to 1. Therefore, the tightness of the extended bound (2) depends on the specific sequence of errors ϵ_n and stepsizes α_n being considered.

The proof of Theorem 1 is presented in Sect. 2 and uses probabilistic arguments by reducing the analysis to the study of an associated Markov reward process in $\mathbb{Z}^2$. As a first consequence of this result, in Sect. 2.3 we explain how the property $\|x_n - Tx_n\| \to 0$ can be used to show that $\mathrm{Fix}(T) \neq \emptyset$ as well as the convergence of the iterates x_n towards a fixed point.

The assumption (H_0) above imposes two conditions: the iterates must remain in C and the images Tx_n must be bounded. Some situations in which these conditions hold are discussed in Sect. 2.4, including the case when T is defined on the whole space and either it has a bounded range, or $\sum_{k\geq 1} \alpha_k \epsilon_k < \infty$ and $\mathrm{Fix}(T) \neq \emptyset$. Note also that when C is bounded one can take $\kappa = \mathrm{diam}(C)$ so that it suffices to ensure that the iterates remain in C. Alternatively, in Sect. 4.1 we consider an iteration that uses approximate projections to deal with the case when x_n might fall outside of C.

Section 3 exploits the bound (2) in order to establish several results on the rate of convergence of the residuals. Theorem 3 shows that if α_n remains away from 0 and 1 and $\sum_{k\geq 1} k^a \|e_k\| < \infty$ then $\|x_n - Tx_n\| = O(1/n^b)$ with $b = \min\{\frac{1}{2}, a\}$. This extends the main result in [21] that covers only the case $a = 1$ and is restricted to Hilbert spaces. On the other hand, from known properties of the Gauss hypergeometric function ${}_2F_1(a, b; c; z)$, we obtain as a Corollary of Theorem 4 that when $\|e_n\| = O(1/n^a)$ with α_n bounded away from 0 and 1, the residual norm satisfies

$$\|x_n - Tx_n\| = \begin{cases} O(1/n^{a-1/2}) & \text{if } \frac{1}{2} \leq a < 1, \\ O(\log n/\sqrt{n}) & \text{if } a = 1, \\ O(1/\sqrt{n}) & \text{if } a > 1. \end{cases}$$

Note that for $a \leq 1$ the assumption $\|e_n\| = O(1/n^a)$ is very mild and allows for nonsummable errors. Similar rates are obtained for vanishing stepsizes of the form $\alpha_n = 1/n^c$ with $c \leq 1$.

Section 4 explores three variants of the basic iteration (IKM). In Sect. 4.1 we deal with the case in which the iterates might fall outside C by using a suitable approximate projection of x_n onto C. Then, in Sect. 4.2 we analyze the Ishikawa iteration which can be seen as a special case of the inexact scheme (IKM), while in Sect. 4.3 we consider a diagonal version of (IKM) in which the operator T might change at each iteration.

The final Sect. 5 extends the results to continuous time, establishing the rate of convergence for the nonautonomous evolution equation

$$\begin{cases} u'(t) + (I - T)u(t) = f(t), \\ u(0) = x_0. \end{cases} \tag{E}$$

2 Proof of Theorem 1

We begin by noting that $x_n - Tx_n = (x_n - x_{n+1})/\alpha_{n+1} + e_{n+1}$ so that

$$\|x_n - Tx_n\| \leq \frac{\|x_n - x_{n+1}\|}{\alpha_{n+1}} + \|e_{n+1}\|. \tag{3}$$

In order to obtain a sharp estimate for $\|x_n - x_{n+1}\|$ we follow a similar approach as in [3,12] by establishing a recursive bound for the differences $\|x_m - x_n\|$ for all $0 \leq m \leq n$. In what follows we let $\alpha_0 = 1$ and for $0 \leq i \leq n$ we denote

$$\pi_i^n = \alpha_i \prod_{k=i+1}^{n} (1-\alpha_k)$$

with the convention $\pi_n^n = \alpha_n$, so that $\sum_{i=0}^n \pi_i^n = 1$. The following is a slight variant of [12, Proposition 2], which itself extends a similar result by Baillon and Bruck [2,3].

Lemma 1 *Let* $(x_n)_{n\in\mathbb{N}}$ *defined inductively by* $x_{n+1} = (1-\alpha_{n+1})x_n + \alpha_{n+1}y_n$ *with* $x_0 \in X$ *and* $y_n \in X$. *Then, setting* $y_{-1} = x_0$, *we have* $x_n = \sum_{i=0}^n \pi_i^n y_{i-1}$ *for all* $n \geq 0$. *Moreover, for* $0 \leq m \leq n$ *it holds that*

$$x_m - x_n = \sum_{i=0}^{m} \sum_{j=m+1}^{n} \pi_i^m \pi_j^n (y_{i-1} - y_{j-1}). \tag{4}$$

Proof The equality $x_n = \sum_{i=0}^n \pi_i^n y_{i-1}$ follows by a straightforward inductive argument, while (4) follows from this equality and the identities $\sum_{i=0}^n \pi_i^n = 1$ and $\pi_i^m - \pi_i^n = \sum_{j=m+1}^n \pi_i^m \pi_j^n$ for $0 \leq i \leq m \leq n$. □

We note that the sequence generated by (IKM) corresponds to $y_n = Tx_n + e_{n+1}$, from which we deduce the following recursive bound.

Corollary 1 *Let* $(x_n)_{n\in\mathbb{N}}$ *be given by* (IKM). *Assume* (H_0) *and let* $\epsilon_n \geq \|e_n\|$. *For each* $n \in \mathbb{N}$ *define inductively* $w_{m,n}$ *for* $-1 \leq m \leq n$ *by setting* $w_{-1,n} = \kappa$ *and*

$$w_{m,n} = \sum_{i=0}^{m} \sum_{j=m+1}^{n} \pi_i^m \pi_j^n (w_{i-1,j-1} + \epsilon_i + \epsilon_j) \quad \textit{for } m = 0, \ldots, n. \tag{5}$$

Then $\|x_m - x_n\| \leq w_{m,n}$ *for all* $0 \leq m \leq n$.

Proof The proof is by induction on n. The base case $n = 0$ being trivial, let us suppose that $\|x_i - x_j\| \leq w_{i,j}$ holds for all i, j with $0 \leq i \leq j < n$. Applying Lemma 1 with $y_n = Tx_n + e_{n+1}$ and setting by convention $Tx_{-1} = y_{-1} = x_0$ and $e_0 = 0$, we get the inequality

$$\|x_m - x_n\| \leq \sum_{i=0}^{m} \sum_{j=m+1}^{n} \pi_i^m \pi_j^n (\|Tx_{i-1} - Tx_{j-1}\| + \|e_i\| + \|e_j\|).$$

The terms with $i = 0$ can be bounded as

$$\|Tx_{-1} - Tx_{j-1}\| = \|x_0 - Tx_{j-1}\| \le \kappa = w_{-1,j-1},$$

while for the remaining terms the nonexpansivity of T and the induction hypothesis give $\|Tx_{i-1} - Tx_{j-1}\| \le \|x_{i-1} - x_{j-1}\| \le w_{i-1,j-1}$. Hence

$$\|x_m - x_n\| \le \sum_{i=0}^{m} \sum_{j=m+1}^{n} \pi_i^m \pi_j^n (w_{i-1,j-1} + \epsilon_i + \epsilon_j) = w_{m,n}$$

which completes the induction step. □

2.1 Reduction to a Markov reward process

The main step in the proof of Theorem 1 relies on a probabilistic reinterpretation of $w_{m,n}$ in terms of a Markov reward process evolving in $\mathbb{Z}^2$. The process is similar to the Markov chain used in [12], except that we add rewards to account for the presence of errors.

Namely, let $m < n$ be positive integers and consider a race between a fox at position n that is trying to catch a hare located at position m. At each integer $i \in \mathbb{N}$ the fox jumps over a hurdle to reach the position $i - 1$. The jump succeeds with probability $(1 - \alpha_i)$ in which case the process repeats, otherwise the fox falls at $i-1$ where it gets a reward ϵ_i. The fox catches the hare if it jumps successfully down to m or below. Otherwise, the hare gets a chance to run towards the burrow located at -1 by following the same rules and with the same rewards. The process alternates until either the fox catches the hare, or the hare reaches the burrow. In the latter case the hare gets an additional reward κ.

This yields a Markov chain with state space

$$\mathcal{S} = \{(m, n) : 0 \le m < n\} \cup \{f, h\}$$

with f, h two absorbing states that represent respectively the cases in which the fox or the hare win the race (see Fig. 1). Specifically, starting from a transient state (m, n) with $0 \le m < n$, the process moves with probability $\pi_i^m \pi_j^n$ to a new state of the form $(i-1, j-1)$ with $1 \le i \le m < j \le n$, and otherwise it is absorbed in the state f with probability $\sum_{j=0}^{m} \pi_j^n$ and in the state h with probability $\pi_0^m \sum_{j=m+1}^{n} \pi_j^n$.

When the process visits a transient state $(i - 1, j - 1)$ the hare gets a reward ϵ_i and the fox gets ϵ_j, which combined yield a total reward $R_{i-1,j-1} = \epsilon_i + \epsilon_j$. If the process reaches state h the hare gets a reward κ and the fox gets nothing, whereas at the absorbing state f there is no reward. Then the total expected reward, when the process starts at position (m, n) with $0 \le m \le n$, satisfies exactly the recursion (5) with boundary condition $w_{-1,n} = \kappa$ corresponding to the reward collected at the absorbing state h. This provides an alternative way to compute $w_{n,n+1}$ and allows to establish the following bound.

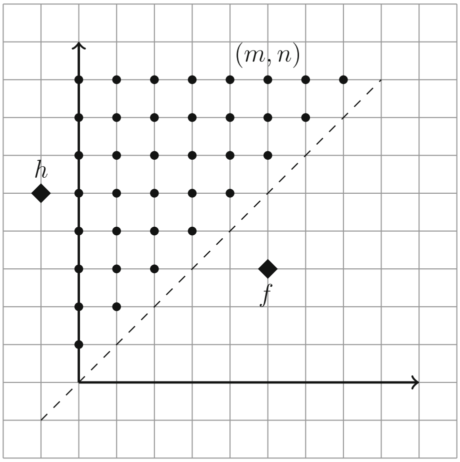

Fig. 1 The state space $\mathcal{S}$

Proposition 1 *Let $w_{m,n}$ be defined recursively by (5) with $w_{-1,n} = \kappa$, where $\epsilon_n \geq 0$ and $\epsilon_0 = 0$. Then, for all $n \in \mathbb{N}$ we have*

$$\frac{w_{n,n+1}}{\alpha_{n+1}} \leq \kappa\, \sigma(\tau_n) + \sum_{i=1}^{n} 2\alpha_i \epsilon_i\, \sigma(\tau_n - \tau_i) + \epsilon_{n+1}. \tag{6}$$

Proof Consider the process starting at state $(n, n+1)$. Let R_i^H denote the event in which the hare collects the reward at the site $i-1$, so that the total expected reward of the hare can be expressed as

$$T^H = \kappa\, \mathbb{P}(R_0^H) + \sum_{i=1}^{n} \epsilon_i\, \mathbb{P}(R_i^H).$$

The event R_i^H occurs iff the process visits a state $(i-1, j-1)$ for some $j > i$, that is to say, if the hare is not captured before the i-th hurdle and it fails this i-th jump. Let F_i and H_i denote independent Bernoulli variables representing the failure of the jump over the i-th hurdle for the fox and hare respectively, with $\mathbb{P}(F_i = 1) = \mathbb{P}(H_i = 1) = \alpha_i$. Denoting S_{i+1} the event that the hare is not captured before the $(i+1)$-th hurdle we have $R_i^H = \{H_i = 1\} \cap S_{i+1}$ so that $\mathbb{P}(R_i^H) = \alpha_i\, \mathbb{P}(S_{i+1})$. Now, the event S_{i+1} can be written as

$$S_{i+1} = \left\{ \sum_{k=j}^{n+1} F_k > \sum_{k=j}^{n} H_k \text{ for all } j = i+1, \ldots, n+1 \right\}$$

which translates the fact that the hare is not captured provided that the fox falls more often than the hare. For $j = n+1$ the condition above amounts to $F_{n+1} = 1$, so that denoting $Z_k = F_k - H_k$ we can write

$$S_{i+1} = \{F_{n+1} = 1\} \cap \left\{ \sum_{k=j}^{n} Z_k \geq 0 \quad \text{for all } j = i+1, \ldots, n \right\}$$

and we may use [12, Proposition 4] to get

$$\mathbb{P}(S_{i+1}) = \alpha_{n+1} \mathbb{P}\left(\sum_{k=j}^{n} Z_k \geq 0 \quad \text{for all } j = i+1, \ldots, n \right) \leq \alpha_{n+1}\, \sigma(\tau_n - \tau_i).$$

From this we obtain $\mathbb{P}(R_i^H) \leq \alpha_i \alpha_{n+1} \sigma(\tau_n - \tau_i)$ and, since $\alpha_0 = 1$, we get

$$T^H \leq \alpha_{n+1} \left[\kappa\, \sigma(\tau_n) + \sum_{i=1}^{n} \alpha_i \epsilon_i\, \sigma(\tau_n - \tau_i) \right]. \tag{7}$$

A similar argument can be used to bound the total reward collected by the fox. Indeed, denoting R_j^F the event in which the hare collects the reward at the site $j-1$, the total expected reward of the fox is

$$T^F = \sum_{j=1}^{n+1} \epsilon_j\, \mathbb{P}(R_j^F).$$

In this case the event R_j^F corresponds to the fact that the process visits a state $(i-1, j-1)$ for some $i \in \{1, \ldots, j-1\}$. This requires that the fox fails the jump of the j-th hurdle, that the hare has not been captured, and that after the fox rests at $j-1$ the hare falls before reaching the burrow. Ignoring the latter condition we get the inclusion

$$R_j^F \subseteq \{F_j = 1\} \cap \{\sum_{k=i}^{n+1} F_k > \sum_{k=i}^{n} H_k \text{ for all } i = j, \ldots, n+1\}.$$

For $j = n+1$ this gives $R_{n+1}^F \subseteq \{F_{n+1} = 1\}$ so that $\mathbb{P}(R_{n+1}^F) \leq \alpha_{n+1}$. Now, for $j = 1, \ldots, n$ the condition $\sum_{k=i}^{n+1} F_k > \sum_{k=i}^{n} H_k$ is superfluous when $i = j$ as it follows from the same condition for $i = j+1$ and the fact that $F_j = 1$. Hence we have $R_j^F \subseteq \{F_j = 1\} \cap S_{j+1}$ which yields as before the upper bound $\mathbb{P}(R_j^F) \leq \alpha_j \alpha_{n+1}\, \sigma(\tau_n - \tau_j)$. From these bounds we get

$$T^F \leq \alpha_{n+1} \left[\sum_{j=1}^{n} \alpha_j \epsilon_j\, \sigma(\tau_n - \tau_j) + \epsilon_{n+1} \right]. \tag{8}$$

Since $w_{n,n+1} = T^H + T^F$, combining (7) and (8) we readily get (6). □

2.2 Proof of Theorem 1

Using (3) and Corollary 1 we get

$$\|x_n - Tx_n\| \leq \frac{w_{n,n+1}}{\alpha_{n+1}} + \epsilon_{n+1}$$

which combined with (6) yields (2). It remains to show that $\|x_n - Tx_n\| \to 0$ when $\tau_n \to \infty$, $\epsilon_n \to 0$, and $\sum_{k\geq 1} \alpha_k \epsilon_k < \infty$. Using (2) and considering a fixed $m \in \mathbb{N}$, we can use the bound $\sigma(\tau_n - \tau_i) \leq 1$ for $i = m+1, \ldots, n$ to get

$$\|x_n - Tx_n\| \leq \kappa\, \sigma(\tau_n) + \sum_{i=1}^{m} 2\alpha_i \epsilon_i\, \sigma(\tau_n - \tau_i) + \sum_{i=m+1}^{n} 2\alpha_i \epsilon_i + 2\epsilon_{n+1}. \qquad (9)$$

Since $\sigma(\tau_n - \tau_i) \to 0$ as $n \to \infty$ we get $\limsup_{n\to\infty} \|x_n - Tx_n\| \leq \sum_{i=m+1}^{\infty} 2\alpha_i \epsilon_i$ so that the conclusion follows by letting $m \to \infty$. □

2.3 Convergence of the iterates

Under some additional conditions, the fact that $\|x_n - Tx_n\| \to 0$ implies the existence of fixed points and the convergence of the (IKM) iteration. The arguments are standard (see e.g. [15]) but for the sake of completeness we sketch the proof. We recall that X is said to have Opial's property if for every weakly convergent sequence $x_n \rightharpoonup x$ we have

$$\liminf_{n\to\infty} \|x_n - x\| < \liminf_{n\to\infty} \|x_n - y\| \quad \forall y \neq x.$$

Theorem 2 *Let $(x_n)_{n\in\mathbb{N}}$ be given by (IKM) and suppose that $\|x_n - Tx_n\| \to 0$ and $\sum_{k\geq 1} \alpha_k \|e_k\| < \infty$.*

a) *If $T(C)$ is relatively compact, x_n converges strongly to a fixed point of T.*
b) *If X is uniformly convex and x_n is bounded then $\mathrm{Fix}(T) \neq \emptyset$. If moreover X satisfies Opial's property then x_n converges weakly to a fixed point of T.*

Proof Let us first note that if $\mathrm{Fix}(T) \neq \phi$ then for each $x \in \mathrm{Fix}(T)$ we have

$$\begin{aligned}\|x_n - x\| &= \|(1-\alpha_n)x_{n-1} + \alpha_n(Tx_{n-1} + e_n) - x\| \\ &\leq (1-\alpha_n)\|x_{n-1} - x\| + \alpha_n\|Tx_{n-1} - Tx\| + \alpha_n\|e_n\| \\ &\leq \|x_{n-1} - x\| + \alpha_n\|e_n\| \end{aligned} \qquad (10)$$

so that the sequence $\|x_n - x\| + \sum_{k>n} \alpha_k \|e_k\|$ decreases with n and hence it converges. Since the tail $\sum_{k>n} \alpha_k \|e_k\|$ tends to 0, it follows that the limit $\ell(x) = \lim_{n\to\infty} \|x_n - x\|$ is well defined for all $x \in \mathrm{Fix}(T)$.

Now, in case a) we may find a strongly convergent subsequence $T(x_{n_k}) \to x$. Since $\|x_n - Tx_n\| \to 0$ we also have $x_{n_k} \to x$ and therefore $x \in \mathrm{Fix}(T)$. Hence,

$\ell(x) = \lim_{n\to\infty} \|x_n - x\| = \lim_{k\to\infty} \|x_{n_k} - x\| = 0$ which proves that $x_n \to x$ in the strong sense.

Similarly, in case b) we can find a weakly convergent subsequence $x_{n_k} \rightharpoonup x$ and since $I - T$ is demiclosed (see [8]) the assumption $\|x_n - Tx_n\| \to 0$ implies that x is a fixed point and $\mathrm{Fix}(T) \neq \emptyset$. Moreover, Opial's property implies that x_n has only one weak cluster point: if $x_{n'_k} \rightharpoonup y$ is another weakly convergent subsequence with $y \neq x$ then

$$\ell(x) = \liminf_{k\to\infty} \|x_{n_k} - x\| < \liminf_{k\to\infty} \|x_{n_k} - y\| = \ell(y)$$
$$\ell(y) = \liminf_{k\to\infty} \|x_{n'_k} - y\| < \liminf_{k\to\infty} \|x_{n'_k} - x\| = \ell(x)$$

which yields a contradiction and therefore $x_n \rightharpoonup x$. □

Remark 2 The existence of fixed points when $T(C)$ is relatively compact goes back to the original work of Krasnosel'skii in 1955 [20], whereas for C bounded in a uniformly convex space this was proved in 1965 independently by Browder [7], Göhde [16], and Kirk [19]. On the other hand, without the summability condition $\sum_{k\geq 1} \alpha_k \|e_k\| < \infty$ the iterates might fail to converge as in the trivial example $Tx = x$ where $x_n = \sum_{k=1}^{n} \alpha_k e_k$.

2.4 The assumption (H_0)

Theorem 1 is based on assumption (H_0) which requires simultaneously that $i)$ the iterates remain in C, and $ii)$ $\|Tx_n - x_0\| \leq \kappa$ for some constant κ.

For a bounded domain C, property $ii)$ holds with $\kappa = \mathrm{diam}(C)$ so that one only needs to check $i)$. This holds automatically for the exact iteration (KM) and more generally when the errors e_{n+1} are such that $Tx_n + e_{n+1} \in C$. Note also that if X is a Hilbert space one could replace T by $\tilde{T} = T \circ P_C$, where P_C is the projection onto the closed convex set C, so that $\tilde{T} : X \to C$ and (H_0) holds with $\kappa = \mathrm{diam}(C)$. When X is not a Hilbert space the projection P_C might not exist and, even if it does, it might fail to be nonexpansive. Moreover, even in a Hilbert setting the projection P_C might be difficult to compute exactly. To deal with these cases, in Sect. 4.1 we will show how (IKM) can be adapted using approximate projections.

When T is defined on the whole space, condition $i)$ is trivial and one only has to check $ii)$. The next result describes two simple cases where this holds.

Proposition 2 *Let $T : X \to X$ be a nonexpansive map.*

a) *If T has a bounded range then (H_0) holds with $\kappa = \sup_{x\in X} \|Tx - x_0\|$.*
b) *If $\mathrm{Fix}(T) \neq \emptyset$ and the sum $S = \sum_{k=1}^{\infty} \alpha_k \|e_k\|$ is finite then (H_0) holds with $\kappa = 2\, dist(x_0, \mathrm{Fix}(T)) + S$.*

Proof Property a) is self evident. In order to prove b) we use (10) inductively to deduce that for each $x \in \mathrm{Fix}(T)$ we have

$$\|x_n - x\| \leq \|x_0 - x\| + \sum_{k=1}^{n} \epsilon_k \|e_k\| \leq \|x_0 - x\| + S \tag{11}$$

and then

$$\|Tx_n - x_0\| \le \|Tx_n - x\| + \|x - x_0\| \le \|x_n - x\| + \|x - x_0\| \le 2\|x_0 - x\| + S.$$

The conclusion follows by taking the infimum over $x \in \mathrm{Fix}(T)$. □

Remark 3 From (11) we get $x_n \in B(x_0, \kappa)$ so in part b) above it suffices that T is defined on a domain C containing this ball.

3 Rates of convergence

In this section we use the bound (2) to estimate the rate of convergence of the fixed point residuals.

Theorem 3 *Assume* (H$_0$). *Suppose that* $\sum_{k\ge 1} \|e_k\| < \infty$ *and that* α_n *is bounded away from* 0 *and* 1. *Then there exists a constant* $\nu \ge 0$ *such that*

$$\|x_n - Tx_n\| \le \frac{\nu}{\sqrt{n}} + \sum_{i \ge \lfloor \frac{n}{2} \rfloor} 2\|e_i\|. \tag{12}$$

Moreover, if $\varphi : [0, \infty) \to [0, \infty)$ *is nondecreasing and* $\mu = \sum_{k\ge 1} \varphi(k)\|e_k\| < \infty$, *then*

$$\|x_n - Tx_n\| \le \frac{\nu}{\sqrt{n}} + \frac{2\mu}{\varphi(\lfloor \frac{n}{2} \rfloor)}. \tag{13}$$

In particular, if $\sum_{k\ge 1} k^a \|e_k\| < \infty$ *for some* $a \ge 0$ *then* $\|x_n - Tx_n\| = O(1/n^b)$ *with* $b = \min\{\frac{1}{2}, a\}$.

Proof Take $\beta > 0$ such that $\alpha_n(1-\alpha_n) \ge \beta$ for all $n \ge 1$ and define

$$\nu = \left(\kappa + 2\sqrt{2} \sum_{k\ge 1} \|e_k\|\right) / \sqrt{\pi\beta}.$$

Since $\sigma(\tau_n - \tau_i) \le 1/\sqrt{\pi\beta(n-i)}$ and $\alpha_i \le 1$, the inequality (12) follows directly from (9) by taking $m = \lfloor \frac{n}{2} \rfloor$, while (13) follows from this and the inequality

$$\varphi(m) \sum_{k\ge m} \|e_k\| \le \sum_{k\ge m} \varphi(k)\|e_k\| \le \mu. \tag{14}$$

The last claim $\|x_n - Tx_n\| = O(1/n^b)$ follows from (13) by taking $\varphi(k) = k^a$. □

Remark 4 Note that in (14) the tail $\mu_m = \sum_{k\ge m} \varphi(k)\|e_k\|$ tends to 0 so that $\sum_{k\ge m} \|e_k\| = o(1/\varphi(m))$. In particular, for $a < 1/2$ the last claim in the previous result can be strengthened to $\|x_n - Tx_n\| = o(1/n^a)$.

The previous result derives a rate of convergence from a control on the sum $\sum_{k\geq 1}\varphi(k)\|e_k\| < \infty$. The next Lemma deals with the case where we control the errors $\|e_n\| \leq \epsilon_n$ rather than their sum.

Lemma 2 *Let* $\eta = \sqrt{1+4/\pi}$. *If* $\epsilon_n \leq (1-\alpha_n)f(\tau_n)$ *with* $f : [0,\infty) \to [0,\infty)$ *nonincreasing, then*

$$\sum_{i=1}^{n} 2\alpha_i\epsilon_i\,\sigma(\tau_n-\tau_i) \leq \eta\int_0^{\tau_n}\frac{f(s)}{\sqrt{\tau_n-s}}\,ds \tag{15}$$

Proof For $s \in [\tau_{i-1},\tau_i]$ we have $\tau_n - s \leq \tau_n - \tau_{i-1} = \tau_n - \tau_i + \alpha_i(1-\alpha_i)$. Since $\tau_n - \tau_i \leq \frac{1}{\pi}\sigma(\tau_n-\tau_i)^{-2}$ and $\alpha_i(1-\alpha_i) \leq \frac{1}{4} \leq \frac{1}{4}\sigma(\tau_n-\tau_i)^{-2}$ it follows that $\tau_n - s \leq (\frac{1}{\pi}+\frac{1}{4})\sigma(\tau_n-\tau_i)^{-2}$. This, combined with the monotonicity of $f(\cdot)$, yields $2f(\tau_i)\sigma(\tau_n-\tau_i) \leq \eta\, f(s)/\sqrt{\tau_n-s}$ so that (15) follows by integrating over the interval $[\tau_{i-1},\tau_i]$ and then summing for $i = 1,\ldots,n$. ◻

Theorem 4 *Assume* (H$_0$). *Suppose that* $\tau_n \to \infty$ *and* $\|e_n\| = O((1-\alpha_n)/\tau_n^a)$.

a) *If* $\frac{1}{2} \leq a < 1$ *then* $\|x_n - Tx_n\| = O(1/\tau_n^{a-1/2})$.
b) *If* $a = 1$ *then* $\|x_n - Tx_n\| = O(\log\tau_n/\sqrt{\tau_n})$.
c) *If* $a > 1$ *then* $\|x_n - Tx_n\| = O(1/\sqrt{\tau_n})$.

Proof Let us consider the three terms in the bound (2). The first term is of order $\kappa\,\sigma(\tau_n) = O(1/\sqrt{\tau_n})$ while the third term is $2\|e_{n+1}\| = O(1/\tau_{n+1}^a)$ so that for $a \geq 1/2$ it is also $O(1/\sqrt{\tau_n})$. To estimate the sum in the middle term we note that $\|e_n\| \leq (1-\alpha_n)f(\tau_n)$ with $f(s) = K/(s+1)^a$ for some constant $K \geq 0$. Hence, denoting $I_a(t) = \int_0^t \frac{1}{(s+1)^a\sqrt{t-s}}\,ds$ and using Lemma 2 we get

$$\sum_{i=1}^{n} 2\alpha_i\,\|e_i\|\,\sigma(\tau_n-\tau_i) \leq \eta\,K\,I_a(\tau_n).$$

For $a = 1$ we have $I_1(t) = \frac{2\mathrm{arcsinh}\sqrt{\mathrm{t}}}{\sqrt{t+1}} = O(\log t/\sqrt{t})$ which yields *b)*. For $a \neq 1$ we may use the change of variables $s = t(1-x)$ to express $I_a(t)$ using the Gauss hypergeometric function ${}_2F_1(a,b;c;z)$, namely

$$I_a(t) = \frac{\sqrt{t}}{(t+1)^a}\int_0^1 \frac{1}{(1-\frac{t}{t+1}x)^a\sqrt{x}}\,dx = \frac{2\sqrt{t}}{(t+1)^a}\,{}_2F_1(a,\frac{1}{2};\frac{3}{2};\frac{t}{t+1}).$$

Now, for $z \sim 0$ we have ${}_2F_1(a,b;c;z) \sim 1$, while setting $d = a+b-c$ we have the following identity (see [1, page 559, equation 15.3.6])

$$\begin{aligned}{}_2F_1(a,b;c;1-z) = &\frac{1}{z^d}\frac{\Gamma(c)\Gamma(d)}{\Gamma(a)\Gamma(b)}{}_2F_1(c-a,c-b;1-d;z)\\ &+\frac{\Gamma(c)\Gamma(-d)}{\Gamma(c-a)\Gamma(c-b)}\,{}_2F_1(a,b;d+1;z),\end{aligned}$$

which is valid when d is not an integer and $|\arg(z)| < \pi$. Taking $z = \frac{1}{t+1}$ with $b = \frac{1}{2}$ and $c = \frac{3}{2}$, it follows that for t large

$$I_a(t) \sim \frac{1}{a-1}\frac{1}{\sqrt{t}} + \frac{\sqrt{\pi}\,\Gamma(1-a)}{\Gamma(\frac{3}{2}-a)}\frac{1}{t^{a-1/2}}.$$

From this we deduce both *a)* and *c)*, except when $a \geq 2$ is an integer since in this case $\Gamma(1-a)$ has a pole. However, for $a \geq 2$ the rate $\|e_n\| = O((1-\alpha_n)/\tau_n^a)$ is stronger than the same condition with $a \in (1,2)$ so that we still get the conclusion $\|x_n - Tx_n\| = O(1/\sqrt{\tau_n})$. □

Remark 5 By adapting the previous proof, if we have a small $o(\cdot)$ estimate of the form $\|\varepsilon_n\| = o((1-\alpha_n)/\tau_n^a)$ then the bounds in parts a) and b) can be improved to $o(\cdot)$. However, one cannot get $o(1/\sqrt{\tau_n})$ in *c*) because the leading term in the bound is $\sigma(\tau_n)$, which is known to be tight (*cf.* Remark 1). A similar remark applies to the next Corollaries.

Corollary 2 *Assume* (H$_0$). *Suppose that* α_n *is bounded away from* 0 *and* 1*, and* $\|e_n\| = O(1/n^a)$.

a) *If* $\frac{1}{2} \leq a < 1$ *then* $\|x_n - Tx_n\| = O(1/n^{a-1/2})$.
b) *If* $a = 1$ *then* $\|x_n - Tx_n\| = O(\log n/\sqrt{n})$.
c) *If* $a > 1$ *then* $\|x_n - Tx_n\| = O(1/\sqrt{n})$.

Proof Since α_n is far from 0 and 1 we have $\tau_n = O(n)$ and the result follows directly from Theorem 4. □

Remark 6 For $\frac{1}{2} \leq a \leq 1$ the condition $\|e_n\| = O(1/n^a)$ is very mild and allows for nonsummable errors. However, this only implies a rate for $\|x_n - Tx_n\|$ and not the convergence of the iterates which in general requires the errors to be summable (see Theorem 2 and Remark 2).

Theorem 4 also gives rates of convergence for vanishing stepsizes of the form $\alpha_n = 1/n^c$ with $c \leq 1$. We record the case $\alpha_n = 1/n$ which is often used.

Corollary 3 *Assume* (H$_0$) *and* $\alpha_n = 1/n$*, and suppose that* $\|e_n\| = O(1/\log^a n)$.

a) *If* $\frac{1}{2} \leq a < 1$ *then* $\|x_n - Tx_n\| = O(1/\log^{a-1/2} n)$.
b) *If* $a = 1$ *then* $\|x_n - Tx_n\| = O(\log\log n/\sqrt{\log n})$.
c) *If* $a > 1$ *then* $\|x_n - Tx_n\| = O(1/\sqrt{\log n})$.

4 Variants of the (IKM) iteration

4.1 Inexact projections

Up to now we assumed (H$_0$) which requires the iterates x_n to remain in C. This is a nontrivial assumption that has to be checked independently. Alternatively one might use the metric projection $P_C : X \to C$, namely $P_C(x) = \arg\min_{z\in C} \|x - z\|$, and consider the iteration

$$x_{n+1} = (1-\alpha_{n+1})x_n + \alpha_{n+1}(T \circ P_C x_n + e_{n+1}). \tag{IKM$_p$}$$

As noted in Sect. 2.4, if P_C is well defined and nonexpansive, which is the case when X is a Hilbert space, the results in the previous sections apply directly by considering the map $T \circ P_C$ instead of T. However, in more general spaces the projection might not exist and even if it exists it might fail to be nonexpansive. On the other hand, even in a Hilbert setting the projection might be hard to compute. To overcome these difficulties one may consider to perform an inexact projection by choosing a sequence $\gamma_n \geq 0$ and, starting from $x_0 \in C$, iterate as follows

$$\begin{cases} \text{take } z_n \in C \text{ with } \|z_n - x_n\| \leq d(x_n, C) + \gamma_n \text{ and} \\ \text{set } x_{n+1} = (1-\alpha_{n+1})x_n + \alpha_{n+1}(Tz_n + e_{n+1}). \end{cases} \tag{IKM$_z$}$$

In general, finding $z_n \in C$ as above requires a specific algorithm. Simple cases where this can be done are when C is a ball or the positive orthant in an L^p space with $1 \leq p \leq \infty$. In these cases the projection might fail to be nonexpansive and might even be nonunique.

Theorem 5 *Let the sequence* (x_n, z_n) *be given by* (IKM$_z$) *with* $\|e_n\| \to 0$ *and* $\sum_{k\geq 1}(\alpha_k\|e_k\|+\gamma_k) < \infty$. *Suppose that* $\sum_{k\geq 1}\alpha_k(1-\alpha_k) = \infty$ *and* $\|x_0 - Tz_n\| \leq \kappa$ *for some* $\kappa \geq 0$. *Then* $\|x_n - z_n\| \to 0$ *and* $\|z_n - Tz_n\| \to 0$.

Proof Let us denote $\delta_n = d(x_n, C) + \gamma_n$. Lemma 3 in Appendix shows that δ_n tends to 0 so that $\|x_n - z_n\| \leq \delta_n \to 0$. On the other hand

$$\|z_n - Tz_n\| \leq \|z_n - x_n\| + \|x_n - Tz_n\| \leq \delta_n + \frac{\|x_{n+1} - x_n\|}{\alpha_{n+1}} + \|e_{n+1}\| \tag{16}$$

so that it remains to show that $\frac{\|x_{n+1}-x_n\|}{\alpha_{n+1}}$ tends to 0. We proceed as before by establishing a recursive bound $\|x_m - x_n\| \leq w_{m,n}$. Taking $y_n = Tz_n + e_{n+1}$ with $y_{-1} = x_0$ and using Lemma 1 we get

$$\|x_m - x_n\| \leq \sum_{i=0}^{m} \sum_{j=m+1}^{n} \pi_i^m \pi_j^m \|y_{i-1} - y_{j-1}\|.$$

Set $w_{-1,n} = \kappa$ for all $n \in \mathbb{N}$ and denote $\epsilon_n = \|e_n\| + \delta_{n-1}$ with $\epsilon_0 = 0$. The terms with $i = 0$ in the previous sum can be bounded as

$$\|y_{-1} - y_{j-1}\| \leq \|x_0 - Tz_{j-1}\| + \|e_j\| \leq w_{-1,j-1} + \epsilon_0 + \epsilon_j.$$

On the other hand, since $\|z_i - z_j\| \leq \|x_i - x_j\| + \delta_i + \delta_j$, the nonexpansivity of T implies that for $j > i \geq 1$

$$\|y_{i-1} - y_{j-1}\| \leq \|Tz_{i-1} - Tz_{j-1}\| + \|e_i\| + \|e_j\| \leq \|x_{i-1} - x_{j-1}\| + \epsilon_i + \epsilon_j.$$

Proceeding as in Corollary 1 we get $\|x_m - x_n\| \leq w_{m,n}$ with $w_{m,n}$ defined recursively by (5), and then Proposition 1 yields

$$\frac{\|x_{n+1} - x_n\|}{\alpha_{n+1}} \leq \frac{w_{n,n+1}}{\alpha_{n+1}} \leq \kappa\, \sigma(\tau_n) + \sum_{i=1}^{n} 2\alpha_i \epsilon_i\, \sigma(\tau_n - \tau_i) + \epsilon_{n+1}. \tag{17}$$

Since Lemma 3 shows that $\epsilon_n \to 0$ and $\sum_{k\geq 1} \alpha_k \epsilon_k < \infty$, by arguing as in the proof of Theorem 1 we deduce that $\frac{\|x_{n+1}-x_n\|}{\alpha_{n+1}}$ converges to 0 as claimed. □

Corollary 4 *Under the same conditions of Theorem 5 the following holds.*

a) *If $T(C)$ is relatively compact, x_n converges strongly to a fixed point of T.*
b) *If x_n remains bounded and X is uniformly convex with Opial's property, then x_n converges weakly to a fixed point of T.*

Proof Let $\epsilon_n = \|e_n\| + \delta_{n-1}$ as in the previous proof so that $\sum_{k\geq 1} \alpha_k \epsilon_k < \infty$. For each $x \in \mathrm{Fix}(T)$ a simple computation yields $\|x_n - x\| \leq \|x_{n-1} - x\| + \alpha_n \epsilon_n$ so that the sequence $\|x_n - x\| + \sum_{k>n} \alpha_k \epsilon_k$ decreases with n, and then $\|x_n - x\|$ converges. Since $\|z_n - x_n\| \to 0$ it follows that $\|z_n - x\|$ converges as well. Then, since $\|z_n - Tz_n\| \to 0$, we may argue as in the proof of Theorem 2 to get the strong/weak convergence of z_n, and hence the convergence of x_n. □

Remark 7 The bound for δ_n in Lemma 3, together with (16) and (17), provide an explicit estimate for $\|z_n - Tz_n\|$ from which one can study its rate of convergence using similar techniques as in Sect. 3.

4.2 Ishikawa iteration

In [17] Ishikawa proposed an alternative method to approximate a fixed point of a nonexpansive $T : C \to C$. Namely, given two sequences $\alpha_n, \beta_n \in (0, 1)$ and starting from $x_0 \in C$, the Ishikawa process generates a sequence by the following two-stage iteration

$$\begin{cases} y_n = (1-\beta_{n+1})x_n + \beta_{n+1} T x_n \\ x_{n+1} = (1-\alpha_{n+1})x_n + \alpha_{n+1} T y_n \end{cases} \tag{I}$$

In this subsection we assume that C is bounded and we denote $\kappa = \mathrm{diam}(C)$.

Corollary 5 *Let the sequence (x_n) be given by the iteration (I) with $\beta_n \to 0$ and $\sum_{k\geq 1} \alpha_k \beta_k < \infty$, and assume $\sum_{k\geq 1} \alpha_k (1-\alpha_k) = \infty$. Then $\|x_n - Tx_n\| \to 0$ and the following estimate holds*

$$\|x_n - Tx_n\| \leq \kappa \left[\sigma(\tau_n) + \sum_{i=1}^{n} \alpha_i \beta_i\, \sigma(\tau_n - \tau_i) + 2\beta_{n+1} \right].$$

Proof We observe that (I) can be written as an (IKM) iteration with errors given by $e_{n+1} = Ty_n - Tx_n$. Since $Ty_n \in C$ the iterates x_n remain in C while by nonexpansivity we have

$$\|e_{n+1}\| \leq \|y_n - x_n\| = \beta_{n+1}\|x_n - Tx_n\| \leq \kappa\, \beta_{n+1}$$

so the result follows directly from Theorem 1. □

Remark 8 Ishikawa proved in [17] that if C is a convex compact subset of a Hilbert space X, the iteration (I) converges strongly to a fixed point as soon as $0 \leq \alpha_n \leq \beta_n \leq 1$ with $\beta_n \to 0$ and $\sum_{k\geq 1} \alpha_k \beta_k = \infty$. Interestingly, Corollary 5 together with Theorem 2 implies the convergence when $\sum_{k\geq 1} \alpha_k \beta_k < \infty$ which is complementary to Ishikawa's condition. Note also that we do not require $\alpha_n \leq \beta_n$. On the other hand, Ishikawa's theorem holds for the larger class of Lipschitzian pseudo-contractive maps, whereas our result is restricted to nonexpansive maps but is valid in more general spaces and it yields the rate of convergence of the fixed-point residual as in Sect. 3.

4.3 Diagonal KM iteration

Let $T_n : C \to C$ be a sequence of nonexpansive maps converging uniformly to T so that $\rho_n = \sup\{\|T_n x - Tx\| : x \in C\}$ tends to 0. Starting from $x_0 \in C$ consider the diagonal iteration

$$x_{n+1} = (1 - \alpha_{n+1})x_n + \alpha_{n+1} T_{n+1} x_n. \tag{DKM}$$

Corollary 6 *Let x_n be a sequence generated by (DKM) with $\rho_n \to 0$ and $\sum_{k\geq 1} \alpha_k \rho_k < \infty$. Suppose that $\sum_{k\geq 1} \alpha_k(1-\alpha_k) = \infty$ and $\|x_0 - Tx_n\| \leq \kappa$ for some $\kappa \geq 0$. Then $\|x_n - Tx_n\| \to 0$ and the following estimate holds*

$$\|x_n - Tx_n\| \leq \kappa\, \sigma(\tau_n) + \sum_{i=1}^{n} 2\alpha_i \rho_i\, \sigma(\tau_n - \tau_i) + 2\rho_{n+1}.$$

Proof Note that (DKM) corresponds to an (IKM) iteration with errors given by $e_{n+1} = T_{n+1}x_n - Tx_n$. Since $T_n x_n \in C$ the iterates remain in C, and $\|e_n\| \leq \rho_n$. Hence the result follows again from Theorem 1. □

Remark 9 The diagonal iteration (DKM) was introduced in [31] in order to compute a solution for the split feasibility problem in Hilbert spaces. Weak convergence of (DKM) was established in [30] for uniformly convex spaces with a differentiable norm, under the same assumptions of Corollary 6. Our result shows that this also holds for uniformly convex spaces with Opial's property, and moreover it yields rates of convergence for the residuals in the same way as in Sect. 3.

5 Application to nonautonomous evolution equations

Let $T : X \to X$ be a nonexpansive map and $f : [0,\infty) \to X$ a continuous function. Let $u : [0,\infty) \to X$ be the unique solution of the evolution equation

$$\begin{cases} u'(t) + (I - T)u(t) = f(t), \\ u(0) = x_0. \end{cases} \tag{E}$$

In the autonomous case $f(t) \equiv 0$, Baillon and Bruck [3] used the Krasnosel'skii–Mann iteration to prove that $\|u'(t)\| = O(1/\sqrt{t})$, assuming that $T : C \to C$ with C a bounded closed convex domain. In the nonautonomous case $u(t)$ could leave the domain C so we assume that T is defined on the whole space.

In order to deal with the unboundedness of the domain, and inspired by Proposition 2, we consider a continuous scalar function $\epsilon(t) \geq \|f(t)\|$ and we assume one of the following alternative conditions

(H_2') *T has a bounded range, and then we let* $\kappa = \sup\{\|Tx - x_0\| : x \in X\}$, *or*
(H_2'') $\mathrm{Fix}(T) \neq \emptyset$ *and $\epsilon(t)$ is decreasing with* $S = \int_0^\infty \epsilon(t)\,dt < \infty$, *and we let* $\kappa = 2\,\mathrm{dist}(x_0, \mathrm{Fix}(T)) + S$.

Under either one of these conditions we have the following analog of Theorem 1.

Theorem 6 *Let $u(t)$ be the solution of* (E) *and assume* (H_2') *or* (H_2''). *Then*

$$\|u'(t)\| \leq \kappa\,\sigma(t) + \int_0^t 2\,\epsilon(s)\sigma(t-s)\,ds + \epsilon(t). \tag{18}$$

Moreover, if $\epsilon(t) \to 0$ and $\int_0^\infty \epsilon(s)\,ds < \infty$ then $\|u'(t)\| \to 0$ as $t \to \infty$.

Proof Fix $t > 0$ and set $\lambda^n = \frac{t}{n}$. Let us consider the sequence $(x_k^n)_{k\geq 0}$ defined by $x_0^n = x_0$ and

$$\frac{x_{k+1}^n - x_k^n}{\lambda^n} = -(I - T)x_k^n + f((k+1)\lambda^n).$$

It is well known that $x_n^n \to u(t)$ and $(x_{n+1}^n - x_n^n)/\lambda^n \to u'(t)$ as $n \to \infty$. On the other hand, x_k^n corresponds to the k-th term of an (IKM) iteration with errors $e_k^n = f(k\lambda^n)$ and constant stepsizes $\alpha_k \equiv \lambda^n$. We claim that ($H_0$) holds with κ defined as in (H_2') or (H_2''). Indeed, in the case (H_2') this follows directly from Proposition 2 a), whereas in the case (H_2'') it follows from Proposition 2 b) and the estimate

$$\sum_{k=1}^\infty \lambda^n \|e_k^n\| \leq \sum_{k=1}^\infty \lambda^n \epsilon(k\lambda^n) \leq \int_0^\infty \epsilon(s)\,ds = S.$$

Hence, letting $\tau_k^n = \sum_{i=1}^k \alpha_i(1-\alpha_i)$ and invoking Proposition 1 we get

$$\frac{\|x_{n+1}^n - x_n^n\|}{\lambda^n} \leq \kappa\,\sigma(\tau_n^n) + 2\sum_{i=1}^n \frac{t}{n}\,\epsilon(i\frac{t}{n})\,\sigma(\tau_n^n - \tau_i^n) + \epsilon((n+1)\frac{t}{n}). \tag{19}$$

Since $\tau_n^n = t(1-\frac{t}{n}) \to t$ as $n \to \infty$ the term $\kappa\,\sigma(\tau_n^n)$ converges to $\kappa\,\sigma(t)$, while for the third term we have $\epsilon((n+1)\frac{t}{n}) \to \epsilon(t)$. Also $\tau_n^n - \tau_i^n = (1-\frac{t}{n})(t-i\frac{t}{n})$ so that the middle term is a Riemann sum for the function $h_n(s) = 2\epsilon(s)\sigma((1-\frac{t}{n})(t-s))$. Since $h_n(s)$ converges uniformly for $s \in [0,t]$ towards $h(s) = 2\epsilon(s)\sigma(t-s)$, this Riemann sum converges as $n \to \infty$ to the integral $\int_0^t 2\epsilon(s)\sigma(t-s)ds$. Therefore, by letting $n \to \infty$ in (19) we obtain (18).

To prove the last claim $\|u'(t)\| \to 0$ we note that $\sigma(t) \to 0$ for $t \to \infty$ while $\epsilon(t) \to 0$ by assumption, so that it suffices to prove that $\int_0^t 2\epsilon(s)\sigma(t-s)ds$ tends to 0 as $t \to \infty$. Denoting $h_t(s) = 2\epsilon(s)\sigma(t-s)\mathbb{1}_{[0,t]}(s)$ this integral is exactly $\int_{\mathbb{R}} h_t(s)ds$. Now, the definition of $\sigma(\cdot)$ implies $h_t(s) \to 0$ pointwise as $t \to \infty$, and since $h_t(s) \leq 2\epsilon(s)$, the conclusion follows from Lebesgue's dominated convergence theorem. □

Clearly, from (18) we also get

$$\|u(t) - Tu(t)\| \leq \kappa\,\sigma(t) + \int_0^t 2\,\epsilon(s)\sigma(t-s)\,ds + 2\epsilon(t)$$

so that $\|u(t) - Tu(t)\| \to 0$ as soon as $\epsilon(t) \to 0$ and $\int_0^\infty \epsilon(s)\,ds < \infty$. As in the discrete setting, from this one can deduce that $\mathrm{Fix}(T) \neq \emptyset$ as well as the convergence of $u(t)$ to a fixed point of T.

Theorem 7 *Let $u(t)$ be the solution of* (E). *Suppose that $\int_0^\infty \epsilon(s)\,ds < \infty$ and $\|u(t) - Tu(t)\| \to 0$.*

a) *If $T(C)$ is relatively compact, $u(t)$ converges strongly to a fixed point of T.*
b) *If X is uniformly convex and $u(t)$ remains bounded then* $\mathrm{Fix}(T) \neq \emptyset$.

Furthermore, if X satisfies Opial's property then $u(t)$ converges weakly to a fixed point of T.

Proof We claim that for all $x \in \mathrm{Fix}(T)$ the limit $\ell(x) = \lim_{t\to\infty}\|u(t) - x\|$ exists. To prove this let $\theta(t) = \frac{1}{2}\|u(t) - x\|^2$ and $g(t) = \sqrt{2\theta(t) + 1} + \int_t^\infty \epsilon(s)\,ds$. In order to establish the existence of the limit $\ell(x)$ it suffices to show that $g(t)$ is decreasing. Let us prove that $g'(t) \leq 0$, that is to say, $\frac{d}{dt}\sqrt{2\theta(t) + 1} \leq \epsilon(t)$. We recall that the duality mapping on X is the subdifferential $J(x) = \partial\psi(x)$ of the convex function $\psi(x) = \frac{1}{2}\|\cdot\|^2$. Choosing $u^*(t) \in J(u(t)-x)$, the subdifferential inequality gives

$$\frac{1}{2}\|u(t) - x\|^2 + \langle u^*(t), v - u(t)\rangle \leq \frac{1}{2}\|v - x\|^2 \qquad (\forall v \in X)$$

so that taking $v = u(t-h)$ with $h > 0$ we get

$$\langle u^*(t), u(t-h) - u(t)\rangle \leq \theta(t-h) - \theta(t).$$

Dividing by h and letting $h \downarrow 0$ it follows that $\theta'(t) \leq \langle u^*(t), u'(t)\rangle$. Then, using (E) and the fact that $x \in \mathrm{Fix}(T)$, the nonexpansivity of T gives

$$\begin{aligned}\theta'(t) &\leq \langle u^*(t), (I-T)x - (I-T)u(t) + f(t)\rangle \\ &= \langle u^*(t), Tu(t) - Tx\rangle - \langle u^*(t), u(t) - x\rangle + \langle u^*(t), f(t)\rangle \\ &\leq \|u^*(t)\|\|u(t)-x\| - \langle u^*(t), u(t)-x\rangle + \|u^*(t)\|\|f(t)\|.\end{aligned}$$

Now, from known properties of the duality mapping we have $\langle u^*(t), u(t) - x\rangle = \|u(t)-x\|^2$ and $\|u^*(t)\| = \|u(t) - x\|$ so that

$$\theta'(t) \leq \|u^*(t)\|\|f(t)\| = \|u(t)-x\|\|f(t)\| \leq \sqrt{2\theta(t)+1}\,\epsilon(t)$$

which proves our claim $\frac{d}{dt}\sqrt{2\theta(t)+1} \leq \epsilon(t)$. This implies the existence of $\ell(x) = \lim_{t\to\infty}\|u(t)-x\|$, from which the rest of the proof follows the same pattern as the proof of Theorem 2. □

The estimate (18) can also be used to derive the following continuous time analogs of the rates of convergence in Theorem 3 and Theorem 4.

Theorem 8 *Let $u(t)$ be the unique solution of* (E). *Assume* (H_2'') *and let $\nu = (\kappa + 2\sqrt{2}\,S)/\sqrt{\pi}$. Then, for all $t \geq 1$ we have*

$$\|u'(t)\| \leq \frac{\nu}{\sqrt{t}} + \int_{t/2}^{\infty} 4\,\epsilon(s)\,ds. \tag{20}$$

Moreover, if $\varphi : [0,\infty) \to [0,\infty)$ is nondecreasing and $\mu = \int_0^\infty \varphi(s)\epsilon(s)\,ds < \infty$, then

$$\|u'(t)\| \leq \frac{\nu}{\sqrt{t}} + \frac{4\mu}{\varphi(t/2)}. \tag{21}$$

In particular, if $\int_0^\infty s^a\,\epsilon(s)\,ds < \infty$ for some $a \geq 0$, then $\|u'(t)\| = O(1/t^b)$ with $b = \min\{\frac{1}{2}, a\}$.

Proof Let us fix $t \geq 1$ and consider the bound (18). Splitting the integral $\int_0^t 2\epsilon(s)\sigma(t-s)\,ds$ into $[0, \frac{t}{2}]$ and $[\frac{t}{2}, t]$, and noting that $\sigma(t-s) \leq \sqrt{2/\pi t}$ on the first interval and $\sigma(t-s) \leq 1$ on the second, we get

$$\|u'(t)\| \leq \frac{\nu}{\sqrt{\pi t}} + \int_{t/2}^{t} 2\,\epsilon(s)\,ds + \epsilon(t). \tag{22}$$

Since $\epsilon(\cdot)$ is nonincreasing and $t \geq 1$, we have $\epsilon(t) \leq \int_{t/2}^{t} 2\,\epsilon(s)\,ds$ which plugged into (22) yields (20). Now, since $\varphi(\cdot)$ is nondecreasing, (21) follows directly from (20) using the inequality

$$\varphi(\frac{t}{2})\int_{t/2}^{\infty} \epsilon(s)\,ds \leq \int_{t/2}^{\infty} \varphi(s)\epsilon(s)\,ds \leq \mu,$$

while the last claim $\|u'(t)\| = O(1/t^b)$ follows from (21) taking $\varphi(s) = s^a$. $\square$

Theorem 9 *Let $u(t)$ be the solution of* (E) *and assume* (H_2') *or* (H_2'')*, and $\epsilon(t) = O(1/t^a)$ with $a \geq \frac{1}{2}$.*

a) *If $\frac{1}{2} \leq a < 1$ then $\|u'(t)\| = O(1/t^{a-1/2})$.*
b) *If $a = 1$ then $\|u'(t)\| = O(\log t/\sqrt{t})$.*
c) *If $a > 1$ then $\|u'(t)\| = O(1/\sqrt{t})$.*

Proof This follows from (18) and from the analysis of the asymptotics of the integral $I_a(t) = \int_0^t \frac{1}{(s+1)^a\sqrt{t-s}}\, ds$ established in the proof of Theorem 4. $\square$

Appendix: Bound for approximate projections

The goal of this Appendix is to establish the next technical Lemma used in the proof of Theorem 5.

Lemma 3 *Let (x_n, z_n) be given by* (IKM_z)*, and denote $\delta_n = d(x_n, C) + \gamma_n$ and $\xi_n = \|e_n\| + \gamma_n/\alpha_n$ with $\delta_0 = \xi_0 = 0$. If $\sum_{k\geq 1} \alpha_k = \infty$ and $\sum_{k\geq 1} \alpha_k \xi_k < \infty$, then $\delta_n \to 0$ and $\sum_{k\geq 1} \alpha_k \delta_{k-1} < \infty$.*

Proof Starting from the identity

$$x_n = (1-\alpha_n)z_{n-1} + \alpha_n T z_{n-1} + (1-\alpha_n)(x_{n-1} - z_{n-1}) + \alpha_n e_n$$

and since $(1-\alpha_n)z_{n-1} + \alpha_n T z_{n-1} \in C$, we get

$$\mathrm{dist}(x_n, C) \leq \|(1-\alpha_n)(x_{n-1} - z_{n-1}) + \alpha_n e_n\| \leq (1-\alpha_n)\delta_{n-1} + \alpha_n \|e_n\|.$$

It follows that $\delta_n \leq (1-\alpha_n)\delta_{n-1} + \alpha_n \xi_n$ and letting $\rho_n = \prod_{j=1}^n (1-\alpha_j)$ with $\rho_0 = 1$ we get

$$\frac{\delta_n}{\rho_n} \leq \frac{\delta_{n-1}}{\rho_{n-1}} + \frac{\alpha_n}{\rho_n}\xi_n.$$

Iterating this inequality we get $\frac{\delta_n}{\rho_n} \leq \sum_{i=0}^n \frac{\alpha_i}{\rho_i}\xi_i$ which yields $\delta_n \leq \sum_{i=0}^n \alpha_i \xi_i \frac{\rho_n}{\rho_i}$.

This inequality can be written as $\delta_n \leq \int_{\mathbb{N}} f_n\, d\mu$ with μ the finite measure on $\mathbb{N}$ defined by $\mu(\{i\}) = \alpha_i \xi_i$, and $f_n : \mathbb{N} \to \mathbb{R}$ given by $f_n(i) = \frac{\rho_n}{\rho_i}$ for $i \leq n$ and $f_n(i) = 0$ for $i > n$. Since $f_n(i) \to 0$ as $n \to \infty$ and $f_n(i) \leq 1$, Lebesgue's dominated convergence theorem implies that $\int_{\mathbb{N}} f_n d\mu$ tends to zero so that $\delta_n \to 0$.

It remains to show that the sum $S = \sum_{k\geq 1} \alpha_k \delta_{k-1}$ is finite. Using the previous bound for δ_{k-1} and exchanging the order of summation we get

$$S \leq \sum_{k=1}^{\infty} \alpha_k \sum_{i=0}^{k-1} \alpha_i \xi_i \frac{\rho_{k-1}}{\rho_i} = \sum_{i=0}^{\infty} \alpha_i \xi_i \sum_{k=i+1}^{\infty} \alpha_k \prod_{j=i+1}^{k-1} (1-\alpha_j).$$

The term $q_{i+1}^k = \alpha_k \prod_{j=i+1}^{k-1}(1-\alpha_j)$ in this last sum can be interpreted as a probability. Namely, suppose that at every integer j we toss a coin that falls head with probability α_j. Then, q_{i+1}^k is the probability that starting at position $i+1$ the first head occurs exactly at position k. Hence $\sum_{k=i+1}^{\infty} q_{i+1}^k = 1$ and therefore $S \leq \sum_{i=0}^{\infty} \alpha_i \xi_i < \infty$. □

References

1. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables. Courier Corporation, North Chelmsford (1964)
2. Baillon, J.B., Bruck, R.E.: Optimal rates of asymptotic regularity for averaged nonexpansive mappings. In: Tan, K.K. (ed.) Proceedings of the Second International Conference on Fixed Point Theory and Applications, pp. 27–66. World Scientific Press, London (1992)
3. Baillon, J.B., Bruck, R.E.: The rate of aymptotic regularity is $O(1/\sqrt{n})$. Lect. Notes Pure Appl. Math. **178**, 51–81 (1996)
4. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New York (2011)
5. Bravo, M., Cominetti, R.: Sharp convergence rates for averaged nonexpansive maps. Isr. J. Math. (2018) **(to appear)**
6. Brézis, H.: Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert. North-Holland, Amsterdam (1973)
7. Browder, F.E.: Nonexpansive nonlinear operators in a Banach space. Proc. Natl. Acad. Sci. USA **54**, 1041–1044 (1965)
8. Browder, F.E.: Semicontractive and semiaccretive nonlinear mappings in Banach spaces. Bull. Am. Math. Soc. **74**, 660–665 (1968)
9. Browder, F.E., Petryshyn, W.V.: The solution by iteration of nonlinear functional equations in Banach spaces. Bull. Am. Math. Soc. **72**, 571–575 (1966)
10. Browder, F.E., Petryshyn, W.V.: Construction of fixed points of nonlinear mappings in Hilbert space. J. Math. Anal. Appl. **20**, 197–228 (1967)
11. Combettes, P.L.: Quasi–Fejérian analysis of some optimization algorithms. Stud. Comput. Math. **8**, 115–152 (2001)
12. Cominetti, R., Soto, J., Vaisman, J.: On the rate of convergence of Krasnosel'skiĭ–Mann iterations and their connection with sums of Bernoullis. Isr. J. Math. **199**, 757–772 (2014)
13. Douglas, J., Rachford, H.H.: On the numerical solution of heat conduction problems in two and three space variables. Trans. Am. Math. Soc. **82**, 421–439 (1956)
14. Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite element approximation. Comput. Math. Appl. **2**, 17–40 (1976)
15. Goebel, K., Kirk, W.A.: Classical theory of nonexpansive mappings. In: Kirk, M., Sims, B. (eds.) Handbook of Metric Fixed Point Theory, pp. 49–91. Kluwer Academic Publishers, Dordrecht (2001)
16. Göhde, D.: Zum prinzip der kontraktiven abbildung. Math. Nachr. **30**, 251–258 (1965)
17. Ishikawa, S.: Fixed points by a new iteration method. Proc. Am. Math. Soc. **44**, 147–150 (1974)
18. Kim, T.H., Xu, H.K.: Robustness of Mann's algorithm for nonexpansive mappings. J. Math. Anal. Appl. **327**, 1105–1115 (2007)
19. Kirk, W.A.: A fixed point theorem for mappings which do not increase distances. Am. Math. Mon. **72**, 1004–1006 (1965)
20. Krasnosel'skii, M.A.: Two remarks on the method of successive approximations. Uspekhi Mat. Nauk **63**, 123–127 (1955)
21. Liang, J., Fadili, J., Peyré, G.: Convergence rates with inexact non-expansive operators. Math. Prog. Ser. A **159**, 1–32 (2016)
22. Liu, L.S.: Ishikawa and Mann iterative process with errors for nonlinear strongly accretive mappings in Banach spaces. J. Math. Anal. Appl. **194**, 114–125 (1995)
23. Mann, W.R.: Mean value methods in iteration. Proc. Am. Math. Soc. **4**, 506–510 (1955)
24. Martinet, B.: Régularisation d'inéquations variationnelles par approximations successives. Rev. Franc. Inform. Rech. Opér. **4**, 154–159 (1970)
25. Mercier, B.: Lectures on Topics in Finite Element Solution of Elliptic Problems, Lectures on Mathematics and Physics, vol. 63. Tata Institute of Fundamental Research, Bombay (1979)

26. Passty, G.B.: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. J. Math. Anal. Appl. **72**, 283–390 (1979)
27. Peaceman, D.W., Rachford, H.H.: The numerical solution of parabolic and elliptic equations. J. Soc. Ind. Appl. Math. **3**, 28–41 (1955)
28. Polyak, B.T.: Gradient methods for the minimisation of functionals. USSR Comput. Math. Math. Phys. **3**, 864–878 (1963)
29. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)
30. Xu, H.K.: A variable Krasnosel'skii–Mann algorithm and the multiple-set split feasibility problem. Inverse Probl. **22**, 2021–2034 (2006)
31. Zhao, J., Yang, Q.: Several solution methods for the split feasibility problem. Inverse Probl. **21**, 1791–1799 (2005)