ISRAEL JOURNAL OF MATHEMATICS **199** (2014), 757–772
DOI: 10.1007/s11856-013-0045-4

ON THE RATE OF CONVERGENCE OF KRASNOSEL'SKIĬ–MANN ITERATIONS AND THEIR CONNECTION WITH SUMS OF BERNOULLIS

BY

R. Cominetti*
Departamento Ingeniería Industrial, Universidad de Chile
República 701, Santiago, Chile
e-mail: rccc@dii.uchile.cl

AND

J. A. Soto**
Departamento Ingeniería Matemática and Centro de Modelamiento Matemático (UMI 2807 CNRS), Universidad de Chile, Blanco Encalada 2120, Santiago, Chile
e-mail: jsoto@dim.uchile.cl

AND

J. Vaisman
Departamento de Ingeniería Matemática, Universidad de Chile
Blanco Encalada 2120, Santiago, Chile
e-mail: hellovaisman@gmail.com

ABSTRACT

In this paper we establish an estimate for the rate of convergence of the Krasnosel'skiĭ–Mann iteration for computing fixed points of non-expansive maps. Our main result settles the Baillon–Bruck conjecture [3] on the asymptotic regularity of this iteration. The proof proceeds by establishing a connection between these iterates and a stochastic process involving sums of non-homogeneous Bernoulli trials. We also exploit a new Hoeffding-type inequality to majorize the expected value of a convex function of these sums using Poisson distributions.

* Supported by Fondecyt 1100046 and Núcleo Milenio Información y Coordinación en Redes ICM/FIC P10-024F.
** Supported by Basal-Conicyt project and Núcleo Milenio Información y Coordinación en Redes ICM/FIC P10-024F.
Received July 6, 2012 and in revised form November 27, 2012

1. Introduction

Let $T : C \to C$ be a non-expansive map defined on a convex subset $C \subseteq X$ of a normed space $(X, \|\cdot\|)$. The Krasnosel'skiĭ–Mann iteration for computing a fixed-point of T is defined by (cf. [22, 23])

$$x_k = (1-\alpha_k)x_{k-1} + \alpha_k T x_{k-1} \tag{1}$$

with $x_0 \in C$ given and $\alpha_k \in [0,1]$.

Strong convergence of x_k to a fixed point was proved in [22] for $\alpha_k \equiv \frac{1}{2}$, when X is a uniformly convex Banach space and $T(C)$ is contained in a compact subset of C. This result was extended to $\alpha_k \equiv \alpha$ [28] and X strictly convex [9], while [17] proved it for general Banach spaces with α_k bounded away from 1 and $\sum \alpha_k = \infty$. The Banach case with $\alpha_k \equiv \alpha$ was also considered in [10]. Without the compactness assumption, weak convergence was established in [25] assuming $\sum \alpha_k(1-\alpha_k) = \infty$ and $\mathrm{Fix}(T) \neq \phi$, for X uniformly convex with a Fréchet differentiable norm. Although strong convergence does not hold in general (see [12] and [5]), it does occur for most operators in the sense of Baire's categories (see [27]).

The crucial step in proving the convergence of the iterates in all these results is to show that $\|x_n - Tx_n\|$ tends to 0, a property which is now known as **asymptotic regularity** [4, 6, 8, 26]. Under various assumptions, asymptotic regularity was also proved in [15] and [13]. The latter noted a certain uniformity in the convergence, namely, for each $\epsilon > 0$ we have $\|x_n - Tx_n\| \le \epsilon$ for all $n \ge n_0$, with n_0 depending on ϵ and C but independent of the initial point x_0 and the map T. More recently, using proof mining techniques, Kohlenbach [20, 21] showed that n_0 could be chosen to depend on C only through its diameter. An explicit metric estimate which readily implies all these results was stated in [3], namely, they conjectured the existence of a universal constant κ such that

$$\|x_n - Tx_n\| \le \kappa \frac{\mathrm{diam}(C)}{\sqrt{\sum_{i=1}^{n} \alpha_i(1-\alpha_i)}} \tag{2}$$

and proved it for the case $\alpha_i \equiv \alpha$ with $\kappa = 1/\sqrt{\pi}$.

In this paper we settle this conjecture by proving that the bound holds in general with $\kappa = 1/\sqrt{\pi}$ for any sequence α_k and each non-expansive $T : C \to C$. Although we do not know whether this is the smallest possible κ, we provide an example which shows that it cannot be improved by more than 17%. We also discuss how the result can be used to analyze the convergence of (1), and how it applies when C is unbounded but $\mathrm{Fix}(T) \neq \phi$.

Our proof is based on a recursive bound for the distances between the iterates $\|x_m - x_n\| \leq c_{mn}$, where c_{mn} admits a nice probabilistic interpretation in terms of a random walk on $\mathbb{Z}$. In proving the theorem we exploit some properties of the hypergeometric and modified Bessel functions, as well as a known identity for Catalan numbers. We also use the following Hoeffding-type inequality which might be of interest on its own: *if $S = X_1 + \cdots + X_m$ is a sum of independent Bernoullis and Z is a Poisson with the same mean $\mathbb{E}(Z) = \mathbb{E}(S)$, then $\mathbb{E}[g(S)] \leq \mathbb{E}[g(Z)]$ for every convex function $g : \mathbb{N} \to \mathbb{R}$.*

2. Main result

THEOREM 1: *The Krasnosel'skiĭ–Mann iterates generated by* (1) *satisfy*

$$\|x_n - Tx_n\| \leq \frac{\mathrm{diam}(C)}{\sqrt{\pi \sum_{i=1}^n \alpha_i(1-\alpha_i)}}. \tag{3}$$

The proof is split into several intermediate steps. Note that by rescaling the norm, we may assume $\mathrm{diam}(C) = 1$.

2.1. A recursive bound.

Let

$$\rho_k = \Pi_{j=1}^k (1-\alpha_j) \quad \text{and} \quad \pi_k^n = \rho_n \frac{\alpha_k}{\rho_k} = \alpha_k \Pi_{j=k+1}^n (1-\alpha_j).$$

By convention we also set $\rho_0 = \alpha_0 = 1$, while the term Tx_{-1} is interpreted as x_0.

PROPOSITION 2: *For $n \geq 0$ we have $x_n = \sum_{k=0}^n \pi_k^n Tx_{k-1}$ and*

$$x_m - x_n = \sum_{j=0}^m \sum_{k=m+1}^n \pi_j^m \pi_k^n \,[Tx_{j-1} - Tx_{k-1}] \quad \text{for } 0 \leq m \leq n. \tag{4}$$

Proof. Dividing (1) by ρ_k we have $\frac{x_k}{\rho_k} = \frac{x_{k-1}}{\rho_{k-1}} + \frac{\alpha_k}{\rho_k} Tx_{k-1}$ which, when iterated, yields $\frac{x_n}{\rho_n} = x_0 + \sum_{k=1}^n \frac{\alpha_k}{\rho_k} Tx_{k-1}$. Using the conventions $\rho_0 = \alpha_0 = 1$ and $x_0 = Tx_{-1}$ we get precisely $x_n = \sum_{k=0}^n \pi_k^n Tx_{k-1}$. This equality, combined with the identities $\sum_{j=0}^{m}\pi_j^m = 1$ and $\pi_k^m - \pi_k^n = \sum_{j=m+1}^{n}\pi_j^n\pi_k^m$, yields

$$\begin{aligned} x_m - x_n &= \sum_{k=0}^{m}(\pi_k^m - \pi_k^n)Tx_{k-1} - \sum_{k=m+1}^{n}\pi_k^n Tx_{k-1} \\ &= \sum_{k=0}^{m}\sum_{j=m+1}^{n}\pi_j^n\pi_k^m Tx_{k-1} - \sum_{j=0}^{m}\sum_{k=m+1}^{n}\pi_j^m\pi_k^n Tx_{k-1} \end{aligned}$$

so that exchanging j and k in the first double sum we obtain (4). ∎

Corollary 3: *Define c_{mn} recursively by setting $c_{-1,n} = 1$ for all $n \geq 0$ and*

$$(R) \qquad c_{mn} = \sum_{j=0}^{m}\sum_{k=m+1}^{n}\pi_j^m\,\pi_k^n\,c_{j-1,k-1} \quad \text{for } 0 \leq m \leq n.$$

Then $\|x_m - x_n\| \leq c_{mn}$ for all $0 \leq m \leq n$.

Proof. The proof is by induction on n. Suppose that $\|x_j - x_k\| \leq c_{jk}$ holds for all $0 \leq j \leq k \leq n-1$. Using the triangle inequality in (4) we get

$$(5) \qquad \|x_m - x_n\| \leq \sum_{j=0}^{m}\sum_{k=m+1}^{n}\pi_j^m\pi_k^n\|Tx_{j-1} - Tx_{k-1}\|.$$

The induction hypothesis gives $\|Tx_{j-1} - Tx_{k-1}\| \leq \|x_{j-1} - x_{k-1}\| \leq c_{j-1,k-1}$ for $1 \leq j < k$, while for $j = 0$ we have $\|Tx_{-1} - Tx_{k-1}\| = \|x_0 - Tx_{k-1}\| \leq \operatorname{diam}(C) = 1 = c_{-1,k-1}$. Plugging these bounds into (5) and using (R) we deduce $\|x_m - x_n\| \leq c_{mn}$ completing the induction step. ∎

Note that for $m = n$ we have $c_{nn} = 0$ and the inequality $\|x_n - x_n\| \leq c_{nn}$ holds trivially. More interestingly, since $\|x_n - x_{n+1}\| = \alpha_{n+1}\|x_n - Tx_n\|$ we have $\|x_n - Tx_n\| \leq \frac{c_{n,n+1}}{\alpha_{n+1}} \triangleq P^n$ so that Theorem 1 will follow by showing

$$(6) \qquad \sqrt{\sum_{i=1}^{n}\alpha_i(1-\alpha_i)}\;P^n \leq 1/\sqrt{\pi}.$$

Our analysis proves that this bound is sharp, so that $\frac{1}{\sqrt{\pi}}$ is the best constant one can get from Corollary 3. This does not exclude the possibility that other techniques might lead to sharper bounds in Theorem 1 (cf. [2]).

2.2. Fox-and-Hare race and a random walk. The recurrence (R) has a probabilistic interpretation. Consider a fox at position n trying to catch a hare located at $m < n$. At each integer $i \in \mathbb{N}$ the fox must jump over a hurdle to reach $i-1$. The jump succeeds with probability $(1-\alpha_i)$ in which case the process repeats, otherwise the fox falls at $i-1$ where it rests to recover from injuries. Thus, starting from n the probability of landing at $k-1$ is precisely π_k^n. The fox catches the hare if it jumps successfully down to m or below. Otherwise, the hare runs toward the burrow located at -1 by following the same rules. The process alternates until either the fox catches the hare, or the hare reaches the burrow.

The recurrence (R) satisfied by c_{mn} characterizes precisely the probability for the hare to reach the burrow safely when the process starts at (m,n). This is also consistent with the boundary cases $c_{-1,n} = 1$ and $c_{nn} = 0$. Note that $\alpha_0 = 1$ so at $i = 0$ both the fox and hare fall with certainty, landing at -1. From this interpretation we get the following expression for c_{mn}.

Proposition 4: *Let $(F_i)_{i\in\mathbb{N}}$ and $(H_i)_{i\in\mathbb{N}}$ denote independent Bernoulli trials representing respectively the events that the fox and hare fail at the i-th hurdle, so that $\mathbb{P}(F_i = 1) = \mathbb{P}(H_i = 1) = \alpha_i$. Then*

$$c_{mn} = \mathbb{P}\Big(\sum_{i=k}^{n} F_i > \sum_{i=k}^{m} H_i \text{ for all } k = m+1,\dots,1\Big). \tag{7}$$

In particular, denoting $Z_i = F_i - H_i$ we have

$$P^n = \frac{c_{n,n+1}}{\alpha_{n+1}} = \mathbb{P}\Big(\sum_{i=k}^{n} Z_i \geq 0 \text{ for } k = n,\dots,1\Big). \tag{8}$$

Proof. Formula (7) is just a restatement of the fact that the hare wins iff the number of times the fox falls in any interval $\{k,\dots,n\}$ is strictly larger than the number of falls of the hare in $\{k,\dots,m\}$. The expression for P^n follows by noting that the event corresponding to $c_{n,n+1}$ in (7) requires $F_{n+1} = 1$ (take $k = n+1$). ∎

Formula (8) has an alternative interpretation. Let $p_i = 2\alpha_i(1-\alpha_i)$ so that Z_i takes values in $\{-1,0,1\}$ with probabilities $p_i/2, 1-p_i, p_i/2$. The sums $\sum_{i=k}^{n} Z_i$ taken in reverse order $k = n,\dots,1$ define a random walk on $\mathbb{Z}$ where at each stage the process stays at the current position with some probability, and otherwise moves left or right with equal probability as in a standard random walk. Hence, P^n is the probability that the walk remains non-negative over

n stages. Conditioning on the total number of stages at which the process effectively moves, this is also the probability that a standard random walk stays non-negative over a random number of stages. Using this interpretation we get the following more explicit formula.

PROPOSITION 5: *Let $M = M_1 + \cdots + M_n$ be a sum of independent Bernoullis with success probabilities $\mathbb{P}(M_i = 1) = p_i = 2\alpha_i(1-\alpha_i)$ and consider the integer function $F(m) = \binom{m}{\lfloor m/2 \rfloor} 2^{-m}$. Then $P^n = \mathbb{E}[F(M)]$.*

Proof. The variable M_i can be interpreted as move/stay and Z_i can be expressed as $Z_i = M_i D_i$ with D_i independent variables representing the direction of the movement: $\mathbb{P}(D_i = -1) = \mathbb{P}(D_i = 1) = \frac{1}{2}$. Conditioning on the sum M and using the exchangeability of the variables D_i we obtain

$$\begin{aligned} P^n &= \sum_{m=0}^{n} \mathbb{P}\Big(\sum_{i=k}^{n} M_i D_i \geq 0 \text{ for } k = n, \ldots, 1 | M = m \Big) \mathbb{P}(M = m) \\ &= \sum_{m=0}^{n} \mathbb{P}\Big(\sum_{j=1}^{\ell} D_j \geq 0 \text{ for } \ell = 1, \ldots, m \Big) \mathbb{P}(M = m). \end{aligned}$$

The expression $\mathbb{P}(\sum_{j=1}^{\ell} D_j \geq 0 \text{ for } \ell = 1, \ldots, m)$ is the probability that a standard random walk started from 0 remains non-negative over m stages. Its value is precisely $F(m)$ [11, Ch. III.3] so the conclusion follows. ∎

The next result establishes an alternative recursion satisfied by c_{mn}. This is not used in our proof, but we state it in case someone could use it to find a simpler proof of Theorem 1.

PROPOSITION 6: *Denoting $\bar{\alpha}_k = 1 - \alpha_k$, we have the recurrence*

$$c_{mn} = \bar{\alpha}_m c_{m-1,n} + \bar{\alpha}_n c_{m,n-1} + (\alpha_n \alpha_m - \bar{\alpha}_n \bar{\alpha}_m) c_{m-1,n-1}. \tag{9}$$

Proof. Denote $w_{jk} = \pi_j^m \pi_k^n c_{j-1,k-1}$ and let $S = A + B - C - D$ with

$$\begin{aligned} A &= c_{mn} &&= \sum_{j=0}^{m} \sum_{k=m+1}^{n} w_{jk}, \\ B &= \bar{\alpha}_m \bar{\alpha}_n c_{m-1,n-1} &&= \sum_{j=0}^{m-1} \sum_{k=m}^{n-1} w_{jk}, \\ C &= \bar{\alpha}_m c_{m-1,n} &&= \sum_{j=0}^{m-1} \sum_{k=m}^{n} w_{jk}, \\ D &= \bar{\alpha}_n c_{m,n-1} &&= \sum_{j=0}^{m} \sum_{k=m+1}^{n-1} w_{jk}. \end{aligned}$$

Canceling out the common terms we get $S = w_{mn} = \alpha_m \alpha_n c_{m-1,n-1}$ which is exactly (9). ∎

2.3. A sharp upper bound. From Proposition 5, the bound (6) is equivalent to showing that

$$R^n(p) \triangleq \sqrt{p_1 + \cdots + p_n}\, \mathbb{E}[F(M_1 + \cdots + M_n)] \le \sqrt{\frac{2}{\pi}}$$

for all n and $0 \le p_i \le \frac{1}{2}$. The function $R^n(p)$ is strictly concave in each variable p_i separately, so the maximum is attained at the extreme values $0, \frac{1}{2}$ or at a unique point in $(0, \frac{1}{2})$. Interestingly, all non-extreme coordinates may be taken equal.

Lemma 7: *$R^n(p)$ is maximal when $p_i \in \{0, u, \frac{1}{2}\}$ for some $0 < u < \frac{1}{2}$.*

Proof. Let p maximize $R^n(p)$ and suppose $p_j = x$ and $p_k = y$ with $x, y \in (0, \frac{1}{2})$ and $x \ne y$. Let $h(k) = \mathbb{E}[F(k + S)]$ where $S = \sum_{i \ne j,k} M_i$ so that

$$\begin{aligned} P^n =& (1-x)(1-y)h(0) + [x(1-y) + y(1-x)]h(1) + xyh(2) \\ =& a + b(x+y) + cxy \end{aligned}$$

with $a = h(0)$, $b = h(1) - h(0)$ and $c = h(0) + h(2) - 2h(1)$. Setting $m = \sum_{i \ne j,k} p_i$ it follows that $x, y \in (0, \frac{1}{2})$ maximize the expression

$$\sqrt{m + x + y}\, [a + b(x+y) + cxy].$$

Setting the partial derivatives to 0 we get $cx = cy$ and since $x \ne y$ it follows that $c = 0$. But then, the function depends only on the sum $x + y$ and we may change these coordinates to $x + \epsilon, y - \epsilon$ keeping the same value, until one of them hits an extreme value: either $x + \epsilon = \frac{1}{2}$ or $y - \epsilon = 0$. This yields a new optimal p with one coordinate less in $(0, \frac{1}{2})$. Repeating this process we get an optimal p whose coordinates take at most one value in $(0, \frac{1}{2})$. ∎

According to this Lemma, in order to bound $R^n(p)$ it suffices to consider the case $p_i \in \{0, u, \frac{1}{2}\}$ with $0 < u < \frac{1}{2}$. Moreover, by changing n we may ignore the deterministic variables with $p_i = 0$. We distinguish two cases.

2.3.1. *All coordinates $p_i = u$.* In this case $R^n(p) = \sqrt{nu}\ \mathbb{E}[F(S)]$ with $S \sim B(n, u)$ Binomial. This case follows from the results in [3] which were obtained using a computer generated proof. Here we provide a direct proof based on a known identity for Catalan numbers.

Proposition 8: *Let $S \sim B(n,u)$ with $0 < u < \frac{1}{2}$. Then*

$$\mathbb{E}[F(S)] = \sum_{k=0}^{n} \frac{(-1)^k}{k+1} \binom{2k}{k} \binom{n}{k} \left(\frac{u}{2}\right)^k \tag{10}$$

and $R^n(p) = \sqrt{nu}\, \mathbb{E}[F(S)]$ increases with n towards $\sqrt{\frac{2}{\pi}}$.

Proof. Using the Binomial theorem, a straightforward computation gives

$$\begin{aligned} \mathbb{E}[F(S)] &= \sum_{j=0}^{n} F(j) \binom{n}{j} u^j (1-u)^{n-j} \\ &= \sum_{j=0}^{n} F(j) \binom{n}{j} u^j \sum_{i=0}^{n-j} \binom{n-j}{i} (-u)^i \\ &= \sum_{j=0}^{n} \sum_{k=j}^{n} (-1)^j F(j) \binom{n}{j} \binom{n-j}{k-j} (-u)^k \\ &= \sum_{k=0}^{n} \binom{n}{k} (-u)^k \sum_{j=0}^{k} (-1)^j \binom{k}{j} F(j), \end{aligned} \tag{11}$$

where the last equality follows from the identity $\binom{n}{j}\binom{n-j}{k-j} = \binom{n}{k}\binom{k}{j}$ and exchanging the order of the sums. The last inner sum may be computed from a known identity for Catalan numbers $C_k = \frac{1}{k+1}\binom{2k}{k}$, namely[1]

$$C_k = \sum_{j=0}^{k} (-1)^j 2^{k-j} \binom{k}{j} \binom{j}{\lfloor j/2 \rfloor} = 2^k \sum_{j=0}^{k} (-1)^j \binom{k}{j} F(j)$$

which when substituted into (11) yields (10).

By direct verification, the expression on the right of (10) is the hypergeometric function ${}_2F_1(-n, \frac{1}{2}; 2; 2u)$, whose Euler integral representation gives

$$\mathbb{E}[F(S)] = \frac{2}{\pi} \int_0^1 t^{-1/2} (1-t)^{1/2} (1-2ut)^n dt.$$

Multiplying by $\sqrt{nu}$ and using the change of variables $s = 2nut$ we get

$$R^n(p) = \sqrt{nu}\, \mathbb{E}[F(S)] = \frac{\sqrt{2}}{\pi} \int_0^{2nu} \sqrt{\frac{1}{s} - \frac{1}{2nu}}\, (1 - \frac{s}{n})^n ds,$$

which increases with n towards the limit $\frac{\sqrt{2}}{\pi} \int_0^\infty \frac{1}{\sqrt{s}} e^{-s} ds = \frac{\sqrt{2}}{\pi} \Gamma(\frac{1}{2}) = \sqrt{\frac{2}{\pi}}$. ∎

[1] See `http://mathworld.wolfram.com/CatalanNumber.html`. A proof is also given in §4.2.

2.3.2. *At least one coordinate* $p_i = \frac{1}{2}$. With no loss of generality assume $p_1 = \frac{1}{2}$ and denote $S = M_2 + \cdots + M_n$. Conditioning on M_1 and setting $g(k) \triangleq \frac{1}{2}[F(k) + F(k+1)]$ we get

$$\mathbb{E}[F(M_1 + \cdots + M_n)] = \mathbb{E}[g(S)].$$

A direct calculation shows that $g : \mathbb{N} \to \mathbb{R}$ is convex, namely

$$g(k) \leq \frac{1}{2}[g(k-1) + g(k+1)] \quad \text{for all } k \geq 1,$$

so we may use the Hoeffding-type inequality in Proposition 12 to obtain $\mathbb{E}[g(S)] \leq \mathbb{E}[g(Z)]$ with $Z \sim P(z)$ a Poisson variable with $z = p_2 + \cdots + p_n$. From this it follows that

$$\begin{aligned} R^n(p) \leq& \sqrt{z + \frac{1}{2}}\ \mathbb{E}[g(Z)] \\ =& \frac{1}{2}\sqrt{z + \frac{1}{2}}\ \sum_{k=0}^{\infty}[F(k) + F(1+k)]\exp(-z)\frac{z^k}{k!}. \\ =& \sqrt{z + \frac{1}{2}}\ \exp(-z)[I_0(z) + (1 - \frac{1}{2z})I_1(z)], \end{aligned} \tag{12}$$

where $I_0(z) = \sum_{k=0}^{\infty} \frac{1}{(k!)^2}(\frac{z}{2})^{2k}$ and $I_1(z) = \sum_{k=0}^{\infty} \frac{1}{k!(k+1)!}(\frac{z}{2})^{2k+1}$ are modified Bessel functions.

Proposition 9: *Let* $h(z)$ *denote the expression in* (12)*. Then* $h(z)$ *is increasing with* $h(z) \leq \lim_{z\to\infty} h(z) = \sqrt{\frac{2}{\pi}}$.

Proof. The identities $I_0'(z) = I_1(z)$ and $I_1'(z) = I_0(z) - \frac{1}{z}I_1(z)$ imply

$$h'(z) = \frac{\exp(-z)}{4z^2\sqrt{z + \frac{1}{2}}}[2(1+z)I_1(z) - zI_0(z)]$$

so that proving that h is increasing reduces to $zI_0(z) \leq 2(1+z)I_1(z)$. Letting $x = z/2$ and rearranging terms, this is equivalent to

$$\sum_{k=1}^{\infty} \frac{x^{2k+1}}{(k-1)!(k+1)!} \leq 2\sum_{k=0}^{\infty} \frac{x^{2k+2}}{k!(k+1)!}.$$

This latter inequality follows easily by noting that each term on the left can be bounded from above by two consecutive terms on the right, namely

$$\frac{x^{2k+1}}{(k-1)!(k+1)!} \leq \frac{x^{2k}}{(k-1)!k!} + \frac{x^{2k+2}}{k!(k+1)!},$$

which results from the trivial inequality $kx \leq k(k+1)+x^2$.

Thus $h(z)$ is increasing and therefore it is bounded from above by its limit $\ell = \lim_{z\to\infty} h(z)$. To prove that $\ell = \sqrt{\frac{2}{\pi}}$ one may use the known asymptotics $\exp(-z)\sqrt{z}\, I_\alpha(z) \to \frac{1}{\sqrt{2\pi}}$ (see [1, Chapter 9]). Alternatively, one may use the integral representation $I_n(z) = \frac{1}{\pi}\int_0^\pi \cos(n\theta)e^{z\cos\theta}d\theta$ to write

$$\ell = \lim_{z\to\infty} \frac{1}{\pi}\sqrt{z+\frac{1}{2}}\int_0^\pi [1+(1-\frac{1}{2z})\cos\theta]e^{-z(1-\cos\theta)}d\theta.$$

Since $\frac{1}{2z}\sqrt{z+\frac{1}{2}} \to 0$ the relevant term for the limit is $\int_0^\pi[1+\cos\theta]e^{-z(1-\cos\theta)}d\theta$, which is transformed by the change of variables $z(1-\cos\theta) = x^2/2$ into

$$\begin{aligned}\ell &= \lim_{z\to\infty} \frac{2}{\pi}\sqrt{1+\frac{1}{2z}}\int_0^{\sqrt{4z}} (1-\frac{x^2}{4z})^{1/2}\, e^{-x^2/2}\, dx \\ &= \frac{2}{\pi}\int_0^\infty e^{-x^2/2}dx = \sqrt{\frac{2}{\pi}}. \qquad \blacksquare\end{aligned}$$

Remark: An alternative proof of the monotonicity of $h(z)$ is obtained by substituting the well-known recurrence $I_{n+1} = I_{n-1} - \frac{2n}{z}I_n$ into the Turan-type inequality $I_{n-1}I_{n+1} \leq I_n^2$ (see [29]) which gives $I_{n-1}^2 - \frac{2n}{z}I_{n-1}I_n \leq I_n^2$. Denoting $x = I_{n-1}/I_n$ we have $x^2 - \frac{2n}{z}x \leq 1$, and solving the quadratic we get $x \leq \frac{n}{z} + \sqrt{1+(\frac{n}{z})^2}$. For $n=1$ this last expression is smaller than $2(z+1)/z$ which gives $zI_0(z) \leq 2(z+1)I_1(z)$ so that $h'(z) \geq 0$.

2.4. Conclusion. The bounds in §2.3 establish (6) and prove Theorem 1. Moreover, the bound (6) is sharp and cannot be improved. Indeed, for $\alpha_i \equiv \alpha$ constant, setting $u = 2\alpha(1-\alpha)$ and $S \sim B(n,u)$ we have

$$\sqrt{\sum_{i=1}^n \alpha_i(1-\alpha_i)}\; P^n = \sqrt{\frac{nu}{2}}\; \mathbb{E}[F(S)]$$

and by Proposition 8 this quantity converges to $1/\sqrt{\pi}$ as $n\to\infty$. This does not mean that (3) is itself sharp since we only have $\|x_n - Tx_n\| \leq P^n$. Thus, a natural question is to find the smallest constant κ for which (2) holds. Although we do not know whether (3) is sharp or not, the following example shows that this bound cannot be improved by more than 17%.

Example: Take $X = \ell^1(\mathbb{N})$ and let C be the set of all sequences $x = (x^i)_{i\in\mathbb{N}}$ with $x^i \geq 0$ and $\sum_{i=0}^\infty x^i \leq 1$, so that $\mathrm{diam}(C) = 2$. Let $T: C\to C$ be the right-shift isometry $T(x^0, x^1, x^2, \ldots) = (0, x^0, x^1, x^2, \ldots)$. Then, the iteration (KM) started from $x_0 = (1, 0, 0, \ldots)$ generates a sequence of the form $x_n = (p_n^0, p_n^1, \ldots, p_n^n, 0, 0, \ldots)$ with

$$p_n^i = \mathbb{P}(X_1 + \cdots + X_n = i)$$

where X_i are independent Bernoullis with $\mathbb{P}(X_i = 1) = \alpha_i$. It follows that

$$\begin{aligned}\|x^n - Tx^n\|_1 =& p_n^0 + |p_n^1 - p_n^0| + |p_n^2 - p_n^1| + \cdots + |p_n^n - p_n^{n-1}| + p_n^n \\ =& 2\ \max\{p_n^i : 0 \leq i \leq n\}.\end{aligned}$$

Now, consider $n = 2m$ Bernoullis trials, half of them with success probability $\alpha_i = \frac{u}{m}$ and the other half with $\alpha_i = 1 - \frac{u}{m}$. Then

$$\max\{p_n^i : 0 \leq i \leq n\} \geq p_{2m}^m = \mathbb{P}(X = Y)$$

with X, Y independent Binomials $B(m, \frac{u}{m})$. When $m \to \infty$ these Binomials converge to Poissons so that p_{2m}^m tends to $\sum_{k=0}^{\infty} (\frac{\exp(-u)u^k}{k!})^2 = \exp(-2u) I_0(2u)$. Since $\sqrt{\sum_{i=1}^{2m} \alpha_i(1-\alpha_i)}$ tends to $\sqrt{2u}$, it follows that $p_{2m}^m \sqrt{\sum_{i=1}^{2m} \alpha_i(1-\alpha_i)}$ can be made as close as desired to the value $\eta = \max_{x \geq 0} \sqrt{x} \exp(-x) I_0(x)$. Hence the optimal κ lies in the interval $[\eta, \frac{1}{\sqrt{\pi}}] \sim [0.4688, 0.5642]$ which leaves a margin of at most 17%.

3. Two direct applications of Theorem 1

3.1. Convergence of the iterates. The following result, which is basically known (cf. [7, 14, 15, 17, 18, 25]), shows how Theorem 1 can be used to obtain the convergence of the iterates, proving at the same time the existence of fixed points.

Proposition 10: *Suppose $\sum \alpha_k(1 - \alpha_k) = \infty$ and x_k bounded.*

(a) *If x_k is relatively compact then $x_k \to \bar{x}$ for some $\bar{x} \in \mathrm{Fix}(T)$.*
(b) *If X is a Hilbert space then $x_k \rightharpoonup \bar{x}$ for some $\bar{x} \in \mathrm{Fix}(T)$.*

Proof. (a) Choose a convergent subsequence $x_{k_n} \to \bar{x}$. From (3) we obtain $x_k - Tx_k \to 0$ so that $\bar{x}$ must be a fixed point. Since

$$\|x_k - \bar{x}\| = \|(1 - \alpha_k)(x_{k-1} - \bar{x}) + \alpha_k(Tx_{k-1} - T\bar{x})\| \leq \|x_{k-1} - \bar{x}\|$$

we conclude that $\|x_k - \bar{x}\|$ decreases to 0.

(b) Since $I - T$ is maximal monotone and $x_k - Tx_k \to 0$, all weak cluster points of x_k belong to $\mathrm{Fix}(T)$. As before $\|x_k - \bar{x}\|$ converges for all $\bar{x} \in \mathrm{Fix}(T)$ so that weak convergence follows from Opial's lemma. ∎

3.2. Unbounded domains. When C is unbounded (2) says nothing. However, if $\mathrm{Fix}(T) \neq \phi$ is nonempty [2], then for each $y \in \mathrm{Fix}(T)$ we may still apply (2) on the bounded subset $\tilde{C} = C \cap B(y, \|y - x_0\|)$ which satisfies $T(\tilde{C}) \subseteq \tilde{C}$ and $\mathrm{diam}(\tilde{C}) \leq 2\|y - x_0\|$. Hence, setting $\tilde{\kappa} = 2\kappa$ and taking the infimum over $y \in \mathrm{Fix}(T)$ we obtain

$$\|x_n - Tx_n\| \leq \tilde{\kappa} \frac{\mathrm{dist}(x_0, \mathrm{Fix}(T))}{\sqrt{\sum_{i=1}^{n} \alpha_i(1-\alpha_i)}}. \tag{13}$$

In particular, Theorem 1 implies that (13) holds with $\tilde{\kappa} = 2/\sqrt{\pi} \sim 1.1284$. In Hilbert spaces, [30] established a sharper bound with $\tilde{\kappa} = 1$. We present this result which exploits the well-known identity

$$\|(1-\alpha)u + \alpha v\|^2 = (1-\alpha)\|u\|^2 + \alpha\|v\|^2 - \alpha(1-\alpha)\|u - v\|^2. \tag{14}$$

Proposition 11: *Let $T : C \to C$ be non-expansive on a convex $C \subset E$ with E a Hilbert space and* $\mathrm{Fix}(T)$ *nonempty. Then* (13) *holds with* $\tilde{\kappa} = 1$.

Proof. It is known that $\|x_k - Tx_k\|$ decreases with k. Indeed,

$$\begin{aligned}
\|x_k - Tx_k\| &= \|(1-\alpha_k)x_{k-1} + \alpha_k Tx_{k-1} - Tx_k\| \\
&\leq (1-\alpha_k)\|x_{k-1} - Tx_{k-1}\| + \|Tx_{k-1} - Tx_k\| \\
&\leq (1-\alpha_k)\|x_{k-1} - Tx_{k-1}\| + \|x_{k-1} - x_k\| \\
&= (1-\alpha_k)\|x_{k-1} - Tx_{k-1}\| + \alpha_k\|x_{k-1} - Tx_{k-1}\| \\
&= \|x_{k-1} - Tx_{k-1}\|.
\end{aligned}$$

Now, using (14), for each $y \in \mathrm{Fix}(T)$ we get

$$\begin{aligned}
\|x_i - y\|^2 &= \|(1-\alpha_i)(x_{i-1} - y) + \alpha_i(Tx_{i-1} - Ty)\|^2 \\
&= (1-\alpha_i)\|x_{i-1} - y\|^2 + \alpha_i\|Tx_{i-1} - Ty\|^2 - \alpha_i(1-\alpha_i)\|x_{i-1} - Tx_{i-1}\|^2 \\
&\leq \|x_{i-1} - y\|^2 - \alpha_i(1-\alpha_i)\|x_{i-1} - Tx_{i-1}\|^2.
\end{aligned}$$

[2] A necessary and sufficient condition to have $\mathrm{Fix}(T) \neq \phi$ is that the iterate sequence $\{x_k\}$ remains bounded (cf. [24]).

Summing these inequalities we see that

$$\sum_{i=1}^{n} \alpha_i(1-\alpha_i)\|x_{i-1} - Tx_{i-1}\|^2 \le \|x_0 - y\|^2 - \|x_n - y\|^2$$

and the monotonicity of $\|x_k - Tx_k\|$ yields

$$\|x_n - Tx_n\| \sqrt{\sum_{i=1}^{n} \alpha_i(1-\alpha_i)} \le \|x_0 - y\|.$$

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

Remark: The previous proof yields a slightly sharper estimate

$$\|x_{n-1} - Tx_{n-1}\| \le \frac{\mathrm{dist}(x_0, \mathrm{Fix}(T))}{\sqrt{\sum_{i=1}^{n} \alpha_i(1-\alpha_i)}}$$

with x_{n-1} in place of x_n on the left.

4. Auxiliary results

4.1. A Hoeffding-type inequality. In this short section we establish a Hoeffding-type inequality for sums of Bernoullis and Poisson variables. We consider an integer function $g : \mathbb{N} \to \mathbb{R}$ satisfying the convexity inequalities $g(k) \le \frac{1}{2}[g(k-1) + g(k+1)]$ for all $k \ge 1$.

Proposition 12: *Let $S = X_1 + \cdots + X_m$ be a sum of independent Bernoulli trials with success probabilities $\mathbb{P}(X_i = 1) = p_i$, and let $z = \mathbb{E}(S) = p_1 + \cdots + p_n$. Then $\mathbb{E}[g(S)] \le \mathbb{E}[g(Z)]$ where $Z \sim P(z)$ is a Poisson with the same mean.*

Proof. Let us first note that the expected value $\mathbb{E}[g(S)]$ increases if we replace any variable X_i by a sum $X_i' + X_i''$ of independent Bernoullis with

$$\mathbb{P}(X_i' = 1) = \mathbb{P}(X_i'' = 1) = \frac{p_i}{2}.$$

Indeed, for $k \in \mathbb{N}$ let $A(k) = \mathbb{E}[g(k + X_i)]$ and $B(k) = \mathbb{E}[g(k + X_i' + X_i'')]$ so that

$$A(k) = (1-p_i)g(k) + p_i g(k+1),$$
$$B(k) = (1-\frac{p_i}{2})^2 g(k) + p_i(1-\frac{p_i}{2})g(k+1) + (\frac{p_i}{2})^2 g(k+2).$$

Taking their difference we have

$$B(k) - A(k) = (\frac{p_i}{2})^2[g(k) - 2g(k+1) + g(k+2)] \geq 0$$

so that replacing k by the random variable $\sum_{j \neq i} X_j$ and taking expectation we obtain the asserted monotonicity.

Now, a well-known result by Hoeffding [16, Theorem 3] proves that[3] $\mathbb{E}[g(S)] \leq \mathbb{E}[g(S_1)]$ with $S_1 \sim B(n, p)$ a binomial with $p = \frac{1}{n}(p_1 + \cdots + p_n)$. Writing S_1 as a sum of n Bernoullis $B(p)$ and sequentially replacing each term by two Bernoullis $B(p/2)$, the expected value increases in each step and we get $\mathbb{E}[g(S)] \leq \mathbb{E}[g(S_2)]$ with $S_2 \sim B(2n, p/2)$. Iterating this doubling argument we obtain $\mathbb{E}[g(S)] \leq \mathbb{E}[g(S_k)]$ where $S_k \sim B(2^k n, p/2^k)$. Since $\mathbb{E}(S_k) = z$ for all k, the result follows by letting $k \to \infty$ and noting that S_k converges to a Poisson variable $Z \sim P(z)$. ∎

4.2. An identity for Catalan numbers. In proving Proposition 8 we used the identity

$$C_k = \sum_{j=0}^{k} (-1)^j 2^{k-j} \binom{k}{j} \binom{j}{\lfloor j/2 \rfloor}.$$

Since this is not found in standard textbooks, for completeness we provide a proof. For each $a \in \mathbb{Z}$ and $P(x)$ a Laurent polynomial (i.e., a function whose Laurent series has finitely many terms) we denote by $[x^a]P(x)$ the coefficient of x^a in $P(x)$. We observe that for each non-negative integer j we have

$$[x^0](x^2 + x^{-2})^j = \begin{cases} \binom{j}{\frac{j}{2}} & \text{for } j \text{ even}, \\ 0 & \text{for } j \text{ odd}; \end{cases}$$

$$[x^2](x^2 + x^{-2})^j = \begin{cases} 0 & \text{for } j \text{ even}, \\ \binom{j}{\frac{j-1}{2}} & \text{for } j \text{ odd}, \end{cases}$$

[3] As a matter of fact, Hoeffding assumes g strictly convex but the general case follows by applying his result to $g(x) + \epsilon x^2$ with $\epsilon \downarrow 0$.

so we can write $\binom{j}{\lfloor j/2\rfloor} = ([x^0]+[x^2])(x^2+x^{-2})^j$ and therefore

$$\begin{aligned}\sum_{j=0}^{k}(-1)^j 2^{k-j}\binom{k}{j}\binom{j}{\lfloor j/2\rfloor} &= ([x^0]+[x^2])\sum_{j=0}^{k}\binom{k}{j}2^{k-j}(-x^2-x^{-2})^j\\ &= ([x^0]+[x^2])\left(2-x^2-x^{-2}\right)^k\\ &= ([x^0]+[x^2])\left(-(x^1-x^{-1})^2\right)^k\\ &= ([x^0]+[x^2])\,(-1)^k\left(x^1-x^{-1}\right)^{2k}\\ &= \binom{2k}{k}-\binom{2k}{k+1} = C_k.\end{aligned}$$

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Dover, New York, 1965.

[2] J. B. Baillon and R. E. Bruck, *Optimal rates of asymptotic regularity for averaged nonexpansive mappings*, in *Proceedings of the Second International Conference on Fixed Point Theory and Applications (K. K. Tan, ed.)*, World Scientific Press, London, 1992, pp. 27–66.

[3] J. B. Baillon and R. E. Bruck, *The rate of asymptotic regularity is* $O(1/\sqrt{n})$, in *Theory and Applications of Nonlinear Operators of Accretive and Monotone Types*, Lecture Notes in Pure and Applied Mathematics, Vol. 178, Dekker, New York, 1996, pp. 51–81.

[4] J. B. Baillon, R. E. Bruck and S. Reich, *On the asymptotic behavior of non-expansive mappings and semigroups in Banach spaces*, Houston Journal of Mathematics **4** (1978), 1–9.

[5] H. Bauschke, E. Matoušková and S. Reich, *Projection and proximal point methods: convergence results and counterexamples*, Nonlinear Analysis **56** (2004), 715–738.

[6] J. Borwein, S. Reich and I. Shafrir, *Krasnosel'skiĭ–Mann iterations in normed spaces*, Canadian Mathematical Bulletin **35** (1992), 21–28.

[7] F. E. Browder, *Nonexpansive nonlinear operators in a Banach space*, Proceedings of the National Academy of Sciences of the United States of America **54** (1965), 1041–1044.

[8] F. E. Browder and W. V. Petryshyn, *The solution by iteration of nonlinear functional equations in Banach spaces*, Bulletin of the American Mathematical Society **72** (1966), 571–575.

[9] M. Edelstein, *A remark on a theorem of M. A. Krasnosel'skiĭ*, The American Mathematical Monthly **73** (1966), 509–501.

[10] M. Edelstein and R. C. O'Brien, *Nonexpansive mappings, asymptotic regularity and successive approximations*, Journal of the London Mathematical Society **17** (1978), 547–554.

[11] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd edn., John Wiley & Sons, New York, 1950.

[12] A. Genel and J. Lindenstrauss, *An example concerning fixed points*, Israel Journal of Mathematics **22** (1975), 81–86.

[13] K. Goebel and W. A. Kirk, *Iteration processes for nonexpansive mappings*, in *Topological Methods in Nonlinear Functional Analysis*, Contemporary Mathematics, Vol. 21, American Mathematical Society, Providence, RI, 1983, pp. 115–123.

[14] D. Göhde, *Zum prinzip der kontraktiven Abbildung*, Mathematische Nachrichten **30** (1965), 251–258.

[15] C. W. Groetsch, *A note on segmenting Mann iterates*, Journal of Mathematical Analysis and Applications **40** (1972), 369–372.

[16] W. Hoeffding, *On the distribution of the number of successes in independent trials*, Annals of Mathematical Statistics **27** (1956), 713–721.

[17] S. Ishikawa, *Fixed points and iterations of a nonexpansive mapping in a Banach space*, Proceedings of the American Mathematical Society **59** (1976), 65–71.

[18] W. A. Kirk, *A fixed point theorem for mappings which do not increase distances*, American Mathematical Monthly **72** (1965), 1004–1006.

[19] W. A. Kirk, *Nonexpansive mappings and asymptotic regularity*, Nonlinear Analysis **40** (2000), 323–332.

[20] U. Kohlenbach, *A quantitative version of a theorem due to Borwein–Reich–Shafrir*, Numerical Functional Analysis and Optimization **22** (2001), 641–656.

[21] U. Kohlenbach, *Uniform asymptotic regularity for Mann iterates*, Journal of Mathematical Analysis and Applications **279** (2003), 531–544.

[22] M. A. Krasnosel'skiĭ, *Two remarks on the method of successive approximations*, Akademiya Nauk SSSR i Moskovskoe Matematicheskoe Obshchestvo. Uspekhi Matematicheskikh Nauk **10** (1955), 123–127.

[23] W. R. Mann, *Mean value methods in iteration*, Proceedings of the American Mathematical Society **4** (1953), 506–510.

[24] S. Reich, *Fixed point iterations of non expansive mappings*, Pacific Journal of Mathematics **60** (1975), 195–198.

[25] S. Reich, *Weak convergence theorems for nonexpansive mappings in Banach spaces*, Journal of Mathematical Analysis and Applications **67** (1979), 274–276.

[26] S. Reich and I. Shafrir, *Nonexpansive iterations in hyperbolic spaces*, Nonlinear Analysis **15** (1990), 537–558.

[27] S. Reich and A. J. Zaslavski, *Convergence of Krasnosel'skiĭ–Mann iterations of nonexpansive operators*, Mathematical and Computer Modelling **32** (2000), 1423–1431.

[28] H. Schaefer, *Über die Methode sukzessiver Approximationen*, Jahresbericht der Deutschen Mathematiker-Vereinigung **59** (1957), 131–140.

[29] V. K. Thiruvenkatachar and T. S. Nagundiah, *Inequalities concerning Bessel functions and orthogonal polynomials*, Proceedings of the Indian Academy of Sciences. Section A **33** (1951), 373–384.

[30] J. Vaisman, *Convergencia fuerte del método de medias sucesivas para operadores lineales no-expansivos*, Memoria de Ingeniería Civil Matemática, Universidad de Chile, 2005.