ELSEVIER

Nonlinear Analysis 56 (2004) 715 – 738

Nonlinear Analysis

www.elsevier.com/locate/na

Projection and proximal point methods: convergence results and counterexamples

Heinz H. Bauschke[a], Eva Matoušková[b], Simeon Reich[c,*]

[a] *Department of Mathematics and Statistics, University of Guelph, Guelph, Ont., Canada N1G 2W1*

[b] *Mathematical Institute, Academy of Sciences of the Czech Republic, Žitná 25, Prague CZ-11567, Czech Republic*

[c] *Department of Mathematics, The Technion—Israel Institute of Technology, Haifa 32000, Israel*

Received 17 June 2003; accepted 17 October 2003

Abstract

Recently, Hundal has constructed a hyperplane H, a cone K, and a starting point y_0 in ℓ_2 such that the sequence of alternating projections $((P_K P_H)^n y_0)_{n\in\mathbb{N}}$ converges weakly to some point in $H \cap K$, but not in norm. We show how this construction results in a counterexample to norm convergence for iterates of averaged projections; hence, we give an affirmative answer to a question raised by Reich two decades ago. Furthermore, new counterexamples to norm convergence for iterates of firmly nonexpansive maps (à la Genel and Lindenstrauss) and for the proximal point algorithm (à la Güler) are provided. We also present a counterexample, along with some weak and norm convergence results, for the new framework of string-averaging projection methods introduced by Censor et al. Extensions to Banach spaces and the situation for the Hilbert ball are discussed as well.

MSC: 47H09; 47J25; 90C25

Keywords: Alternating projections; Averaged projections; Hilbert space; Nonexpansive; Proximal point algorithm; Weak convergence

* Corresponding author.

E-mail addresses: hbauschk@uoguelph.ca (H.H. Bauschke), matouse@math.cas.cz (E. Matoušková), sreich@tx.technion.ac.il (S. Reich).

0362-546X/$ - see front matter
doi:10.1016/j.na.2003.10.010

1. Introduction

Throughout (most of) this paper, we assume that

$$X \text{ is a real Hilbert space with inner product } \langle\cdot,\cdot\rangle \text{ and induced norm } \|\cdot\|. \quad (1)$$

Let A and B be closed convex sets in X with corresponding *projectors* (also known as *projection operators* or *nearest point mappings*, see Fact 2.4) P_A and P_B, respectively. The *convex feasibility problem* asks to

$$\text{find a point in } A\cap B, \quad (2)$$

assuming this intersection is nonempty. This problem is of considerable importance in mathematics and the physical sciences; see [8,25,20,23], and the references therein.

Perhaps the oldest algorithmic approach to solve (2) is to generate the sequence of *alternating projections*, which is defined by

$$x_0 \mapsto x_1 = P_A x_0 \mapsto x_2 = P_B x_1 \mapsto x_3 = P_A x_2 \mapsto \cdots$$
$$\text{for some } \textit{starting point } x_0 \in X. \quad (3)$$

If A and B are *subspaces*, then the sequence $(x_n)_{n\in\mathbb{N}}$ converges *in norm* to the point in the intersection that is nearest to the starting point—this basic result was proved by von Neumann [51] in 1933. Thirty-two years later, Bregman [15] proved that the sequence $(x_n)_{n\in\mathbb{N}}$ converges at least *weakly* to some point in $A\cap B$.

Ever since, there has been a nagging gap between von Neumann's and Bregman's result: *Is it possible that norm convergence fails?* In 2000, Hundal announced his affirmative answer to this question; full details of his ingenious construction became available two years later, see [34].

In the broader setting of fixed point theory, Hundal's counterexample to norm convergence of the iterates of *compositions of projectors* is similar to the counterexample by Genel and Lindenstrauss [29] (see also [14, pp. 72–74]) to norm convergence of the iterates of a *firmly nonexpansive map*, as well as to Güler's counterexample [32] to norm convergence of the *proximal point algorithm*. In passing, we note that certain modifications of these algorithms are able to always generate norm convergent sequences; see, for instance, [11,53].

Another classical algorithmic approach to solve (2) is to employ (*midpoint*) *averages* of projectors, rather than compositions. This amounts to constructing a sequence $(x_n)_{n\in\mathbb{N}}$ via

$$(\forall n\in\mathbb{N}) \quad x_{n+1} = \left(\tfrac{1}{2}P_A + \tfrac{1}{2}P_B\right)x_n, \quad \text{where } x_0\in X. \quad (4)$$

In his 1969 thesis, Auslender [2] established weak convergence of $(x_n)_{n\in\mathbb{N}}$ to some point in $A\cap B$; this result also follows from [45, Corollary 2.6]. Closely related are Cimmino's method of averaged reflectors for solving linear equations [24] and Merzlyakov's method of extrapolated averaged projectors for solving linear inequalities [39]. On the other hand, a more general result by Reich (see [45, Theorems 1.7 and 2.3]) implies that if A and B are *subspaces*, then the sequence $(x_n)_{n\in\mathbb{N}}$ generated by (4) does converge *in norm* to some point in $A\cap B$. (Alternatively, one can apply von Neumann's result to the subspaces $A\times B$ and $\{(x,x): x\in X\}$ in the product space $X\times X$.)

Analogously to the above discussion of alternating projections, Auslender's weak convergence result for general sets and Reich's norm convergence for subspaces lead to the question—originally raised by Reich at the end of [45, Section 2] in 1983—on whether norm convergence of a sequence generated by (4) may actually fail.

The main objective in this paper is to show that Hundal's recent ingenious counterexample can also be used to provide an affirmative answer to Reich's question. We then explain how this construction leads to new counterexamples to norm convergence for the classical proximal point algorithm (see [38,49]) and for the string-averaging projection method recently introduced by Censor et al. [21]. The notion of a strongly nonexpansive map allows us to also comment on the situation in Banach spaces and for the Hilbert ball.

The remainder of this paper is organized as follows. Section 2 contains auxiliary results on projectors and reflectors as well as a review of the useful properties of strongly nonexpansive maps. The classical results by von Neumann and by Bregman are reviewed in Section 3, where we also include a new elementary proof of von Neumann's result. In Section 4, we describe Hundal's counterexample and show that the composition of the two projectors is not firmly nonexpansive. Our solution to Reich's question is presented in Section 5; it also gives rise to a new counterexample à la Genel and Lindenstrauss. Section 6 contains a self-contained and somewhat more explicit proof of Moreau's result that the proximal maps form a convex set. This is used in Section 7, where it leads to a new counterexample à la Güler. Convergence results on string-averaging methods as well as a counterexample are given in Section 8. In the final Section 9, we discuss the situation in the Hilbert ball.

Notation employed is standard in convex analysis: I denotes the identity map and $\mathbb{N}=\{0,1,2,\ldots\}$ are the nonnegative integers. Also, $S^{\perp}=\{x^*\in X:(\forall s\in S)\langle x^*,s\rangle=0\}$ (respectively, $S^{\ominus}=\{x^*\in X:(\forall s\in S)\langle x^*,s\rangle\leqslant 0\}$, $\overline{\operatorname{span}}\,S$, $\overline{\operatorname{cone}}\,S$, $\operatorname{conv}S$, $\operatorname{int}S$, $\imath_S$) is the orthogonal complement (respectively, polar cone, closed linear span, closed convex conical hull, convex hull, interior, and indicator function) of a set S in X. The subdifferential map (respectively, gradient map, Fenchel conjugate) of a function f is denoted by ∂f (respectively, ∇f, f^*), and $f_1\Box f_2$ stands for the infimal convolution of the functions f_1 and f_2. If T is a map defined on S, then its fixed point set is $\operatorname{Fix}T=\{x\in S:x=Tx\}$. Finally, if $r\in\mathbb{R}$, then $\lfloor r\rfloor$ denotes the largest integer less than or equal to r.

2. Projectors, reflectors, and strongly nonexpansive maps

Definition 2.1 (Nonexpansive and firmly nonexpansive). Let C be a set in X and $T:C\to X$ be a map. Define a family of functions by

$$(\forall x\in C)(\forall y\in C)$$

$$\Phi_{x,y}:[0,1]\to[0,+\infty[\,:\lambda\mapsto\|((1-\lambda)x+\lambda Tx)-((1-\lambda)y+\lambda Ty)\|. \tag{5}$$

Then T is *firmly nonexpansive* (respectively, *nonexpansive*), if $\Phi_{x,y}$ is decreasing (respectively, $\Phi_{x,y}(0)\geqslant\Phi_{x,y}(1)$), for all x and y in C.

Clearly, if T is firmly nonexpansive, then it is nonexpansive, which in turn is equivalent to

$$(\forall x \in C)(\forall y \in C) \quad \|Tx - Ty\| \leqslant \|x - y\|. \tag{6}$$

Lemma 2.2. *Let C be a set in X and let $T: C \to X$. Then the following properties are equivalent*:
(i) *T is firmly nonexpansive.*
(ii) $(\forall x \in C)(\forall y \in C)\ \|Tx - Ty\|^2 \leqslant \langle Tx - Ty, x - y\rangle$.
(iii) *$2T - I$ is nonexpansive.*

Proof. See [31, Lemma 1.11.1]. □

Remark 2.3 (Extensions to Banach spaces). In Hilbert space, a firmly nonexpansive map is usually defined as in Lemma 2.2(ii); nonetheless, we begin with Definition 2.1 because the latter is more useful in Banach spaces (see [31, Section 1.11]). Using the duality map, the characterization in Lemma 2.2(ii) holds true in Banach space (see [31, Lemma 1.11.1]).

Fact 2.4 (Projector and reflector). *Suppose that C is a nonempty closed convex set in X. Then, for every point $x \in X$, there exists a unique point $P_C x \in C$ such that $\|x - P_C x\| = \inf_{y \in C} \|x - y\|$. The point $P_C x$ is the* projection *of x onto C; it is characterized by*

$$P_C x \in C \quad \text{and} \quad (\forall c \in C)\ \langle c - P_C x, x - P_C x\rangle \leqslant 0. \tag{7}$$

The corresponding map $P_C: X \to C$ is the projector (*or* projection operator) *onto C. It is firmly nonexpansive and, hence, the associated* reflector *$R_C = 2P_C - I$ is nonexpansive.*

Proof. See [30, Chapter 12, 31, Propositions 1.3.5 and 1.111.2], or [55, Lemma 1.1]. □

Corollary 2.5. *Let C be a nonempty closed convex set in X. Suppose that $U: X \to X$ is* unitary, *i.e., a surjective linear isometry. Then $U(C)$ is a nonempty closed convex set, and $P_{U(C)} = UP_C U^*$.*

Proof. This follows easily from Fact 2.4. □

Corollary 2.6. *Let H be a hyperplane in X and suppose $0 \in H$. Then the corresponding reflector R_H is unitary and it satisfies $R_H = R_H^* = R_H^{-1}$. Let C be a nonempty closed convex set in X and set $D = R_H(C)$. Then the following hold true*:
(i) *D is closed and convex*;
(ii) $P_D = R_H P_C R_H$;
(iii) $(P_H P_C)|_H = \frac{1}{2}(P_C + P_D)|_H$.

Proof. Write $H = \{a\}^{\perp}$, where $\|a\| = 1$. Fix $x \in X$ arbitrarily. Then $P_H x = x - \langle a,x\rangle a$ and hence $R_H x = x - 2\langle a,x\rangle a$. It follows that R_H is one-to-one. Also, one verifies easily that $R_H^2 = I$ and that R_H is an isometry. Hence R_H is unitary and (i) and (ii) follow from Corollary 2.5. Now pick $h \in H$. Then $R_H h = 2P_H h - h = h$ and so, using (ii), $P_D h = R_H P_C R_H h = R_H P_C h = 2P_H P_C h - P_C h$. Hence $P_C h + P_D h = 2P_H P_C h$, as claimed. □

We conclude this section with a discussion of the class of strongly nonexpansive maps, introduced by Bruck and Reich [18] in 1977.

Definition 2.7 (Strongly nonexpansive). Let C be a set in X and let $T : C \to X$ be a map. Then T is *strongly nonexpansive* if it is nonexpansive and $(x_n - y_n) - (Tx_n - Ty_n) \to 0$ whenever $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are sequences in X such that $(x_n - y_n)_{n\in\mathbb{N}}$ is bounded and $\|x_n - y_n\| - \|Tx_n - Ty_n\| \to 0$.

While the class of firmly nonexpansive maps is convex (use Lemma 2.2(iii)), it is not closed under composition (in the Euclidean plane, consider projectors corresponding to two distinct nonorthogonal intersecting lines). This serious limitation does not occur for strongly nonexpansive maps.

Fact 2.8. *Let $T_1, T_2, \ldots, T_m$ be strongly nonexpansive maps defined on a set in X. Suppose T is given in one of the following ways*:

(i) $T = T_m \cdots T_2 T_1$.
(ii) $T = \sum_{i=1}^m \lambda_i T_i$, *where* $\{\lambda_1, \lambda_2, \ldots, \lambda_m\} \subset\,]0,1[$ *and* $\sum_{i=1}^m \lambda_i = 1$.

Then T is strongly nonexpansive. If $F = \bigcap_{i=1}^m \operatorname{Fix} T_i \neq \emptyset$, then $\operatorname{Fix} T = F$.

Proof. [18, Propositions 1.1 and 1.3] imply that T is strongly nonexpansive. To obtain the identity for F, which we now assume to be nonempty, use [18, Lemma 2.1] (for (i)) and [45, Lemma 1.4] (for (ii)). □

Definition 2.9 ((Sunny) retraction). A map $T : X \to C$ is a (*sunny*) *retraction* onto C if it is continuous with $\operatorname{Fix} T = C$ (and $Tx = c$ implies $T(c + \mu(x - c)) = c$, for all $x \in X$, $c \in C$, and $\mu \geqslant 0$). Note that C is necessarily closed.

Remark 2.10. Some comments on retractions are in order. See also [31, Section 1.13] for further information.

(i) If C is a nonempty closed convex set in X, then it not hard to show that the projector P_C is a sunny nonexpansive retraction onto C [31, (3.7) on p.17].
(ii) Outside Hilbert space, projectors are still sunny retractions, but they fail to be nonexpansive.
(iii) Every sunny nonexpansive retraction is firmly nonexpansive [45, Lemma 2.1].

The class of strongly nonexpansive maps is quite rich.

Fact 2.11. *Let T be a map defined on a set in X. Then T is strongly nonexpansive provided that one of the following conditions holds*:

(i) *T is firmly nonexpansive.*

(ii) *$T = (1-\lambda)T_1 + \lambda T_2$, where T_1 is strongly nonexpansive, T_2 is nonexpansive, and $\lambda \in [0, 1[$.*

(iv) *$T : X \to X$ is a sunny nonexpansive retraction onto* $\operatorname{Fix} T$.

Proof. For (i) and (ii), see [18, Propositions 1.3 and 2.1]. Item (iii) follows from (i) and Remark 2.10(iii). □

Remark 2.12. In any smooth Banach space, there is at most one sunny nonexpansive retraction onto a given subset [31, Lemma 1.13.1]; consequently, if T is as in Fact 2.11(iii), then T must coincide with the projector onto $\operatorname{Fix} T$.

The next result shows that the iterations of strongly nonexpansive maps are well understood.

Fact 2.13. *Let C be a closed convex nonempty set in X, and let $T : C \to C$ be strongly nonexpansive. Set $F = \operatorname{Fix} T$ and pick $x \in C$. Then the following hold true*:

(i) *If $F = \emptyset$, then $\lim_n \|T^n x\| = +\infty$.*

(ii) *If $F \neq \emptyset$, then $(T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in F.*

(iii) *If $C = -C$ and T is odd, then $(T^n x)_{n\in\mathbb{N}}$ converges in norm to some point in F.*

(iv) $\|T^n x - T^{n+1} x\| \to \inf_{c\in C} \|c - Tc\|$

Proof. See [18, Corollaries 1.4, 1.3, 1.2, and Proposition 1.2], respectively. □

Remark 2.14 (Extensions to Banach spaces). The results cited above hold true in considerably more general settings: indeed, Fact 2.8 is valid for a general Banach space X, while Fact 2.11 holds true when X is uniformly convex. We now provide some sufficient conditions for the items of Fact 2.13.

(i) C is boundedly weakly compact and each weakly compact convex subset of C has the fixed point property for nonexpansive mappings [18, Corollary 1.4].

(ii) X and X^* are uniformly convex [45, Proposition 2.4].

(iii) X is uniformly convex (combine either [18, Corollary 1.2] or [45, Proposition 1.5] with [3, Theorem 1.1]).

(iv) does not require any additional assumption (use [18, Proposition 1.2], or combine [45, Proposition 1.5] with [43, Proposition 4.3]). Furthermore, norm convergence of $(T^n x - T^{n+1} x)_{n\in\mathbb{N}}$ to the unique element of minimum norm in the closure of the range of $I - T$ is guaranteed either when X is uniformly convex [43, Theorem 3.7(b)], or when the norm of X is Gâteaux differentiable and the norm of its dual X^* is Fréchet differentiable [44, Corollary 5.3(b)].

All these conditions are satisfied provided both X and X^* are uniformly convex—this holds, of course, for the classical L_p and ℓ_p spaces, where $1 < p < +\infty$.

3. von Neumann's and Bregman's classical results

We now present an elementary geometric proof of von Neumann's original norm convergence result.

Theorem 3.1 (von Neumann). *Let A and B be closed linear subspaces in X. Define the sequence of alternating projections by*

$$x_0 \in X, \quad (\forall n \in \mathbb{N})\ x_{2n+1} = P_A x_{2n} \ \text{and}\ x_{2n+2} = P_B x_{2n+1}. \tag{8}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges in norm to $P_{A\cap B}x_0$.

Proof. Set $C = A \cap B$. By Pythagoras, we have

$$(\forall n \in \mathbb{N}) \quad \|x_n\|^2 = \|x_{n+1}\|^2 + \|x_n - x_{n+1}\|^2. \tag{9}$$

In particular,

$$(\|x_n\|)_{n\in\mathbb{N}} \text{ is decreasing and nonnegative, hence convergent.} \tag{10}$$

By (strong) induction on n, we now establish the related statement

$$\begin{aligned}&(\forall n \in \mathbb{N}) \quad (\forall k \in \mathbb{N})(\forall l \in \mathbb{N}) \quad 1 \leqslant k = l - n \\ &\quad \Rightarrow \quad \|x_k - x_l\|^2 \leqslant \|x_k\|^2 - \|x_l\|^2.\end{aligned} \tag{11}$$

Clearly, (11) is true for $n = 0$ and also for $n = 1$ (by (9)). So assume (11) holds true for some $n \geqslant 1$, and take k, l in $\mathbb{N}$ such that $1 \leqslant k = l - (n+1)$.

Case 1: *n is even.*

Then $n+1 = l-k$ is odd. If l is odd, then both $x_{k+1} = P_A x_k$ and $x_l = P_A x_{l-1}$ belong to A, whereas $x_k - x_{k+1} = (I - P_A)x_k \in A^\perp$; hence altogether

$$\langle x_k - x_{k+1}, x_{k+1} - x_l\rangle = 0. \tag{12}$$

If l is even, we argue similarly with A replaced by B and we derive (12) once again. Using (12), (9), and the induction hypothesis, we now obtain

$$\begin{aligned}\|x_k - x_l\|^2 &= \|x_k - x_{k+1}\|^2 + \|x_{k+1} - x_l\|^2 \\ &= \|x_k\|^2 - \|x_{k+1}\|^2 + \|x_{k+1} - x_l\|^2 \\ &\leqslant \|x_k\|^2 - \|x_{k+1}\|^2 + \|x_{k+1}\|^2 - \|x_l\|^2 \\ &= \|x_k\|^2 - \|x_l\|^2.\end{aligned}$$

Case 2: *n is odd.*

Then $n+1 = l-k$ is even, which—similarly to the derivation of (12)—implies that

$$\langle x_k - x_l, x_l - x_{l-1}\rangle = 0. \tag{13}$$

Thus $\|x_k - x_{l-1}\|^2 = \|x_k - x_l\|^2 + \|x_l - x_{l-1}\|^2$. Using this, (9), the induction hypothesis, and (10), we conclude

$$\begin{aligned}\|x_k - x_l\|^2 &= \|x_k - x_{l-1}\|^2 - \|x_l - x_{l-1}\|^2\\ &= \|x_k - x_{l-1}\|^2 - \|x_{l-1}\|^2 + \|x_l\|^2\\ &\leqslant \|x_k\|^2 - \|x_{l-1}\|^2 - \|x_{l-1}\|^2 + \|x_l\|^2\\ &\leqslant \|x_k\|^2 - \|x_l\|^2.\end{aligned}$$

Altogether, statement (11) is verified.

Now, by (10) and (11), the sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy and hence convergent, say to

$$x_\infty = \lim_{n\in\mathbb{N}} x_n. \tag{14}$$

Since $(x_{2n+1})_{n\in\mathbb{N}}$ lies in A and $(x_{2n+2})_{n\in\mathbb{N}}$ lies in B, we conclude that $x_\infty \in A\cap B = C$. Therefore,

$$P_C x_n \to P_C x_\infty = x_\infty. \tag{15}$$

Now fix $n\in\mathbb{N}$ and $t\in\mathbb{R}$, and set $c=(1-t)P_C x_n + tP_C x_{n+1}$. Then $c\in C = A\cap B = \operatorname{Fix}(P_A)\cap \operatorname{Fix}(P_B)$ and so $P_A c = P_B c = c$. Also, $x_{n+1}\in\{P_A x_n, P_B x_n\}$. Since projectors are nonexpansive (Fact 2.4), we obtain $\|x_{n+1}-c\| \leqslant \|x_n - c\|$. After squaring and simplifying, this inequality turns into

$$(1-2t)\|P_C x_{n+1} - P_C x_n\|^2 + \|P_{C^\perp} x_{n+1}\|^2 \leqslant \|P_{C^\perp} x_n\|^2. \tag{16}$$

Since n and t were chosen arbitrarily, we conclude

$$(\forall n\in\mathbb{N}) \quad P_C x_n = P_C x_{n+1}. \tag{17}$$

Combining (14), (15) and (17), we obtain $\lim_{n\in\mathbb{N}} x_n = P_C x_0$. □

Remark 3.2. See [50–52] for classical proofs of Theorem 3.1, and also [27, Chapter 9] for recent information and further pointers to the literature. The convergence part in the proof given above is a modification of the proof of [28, Proposition 1, p. 105]. The part determining the limit is borrowed from [12, Fact 2.2(v)]; see also [4, Theorem 6.2.2], [13, Theorem 2.2]. Let us sketch the following different approach. Define $L: X \to A\cap B$ by $x_0 \mapsto \lim_n x_n$. Then L is nonexpansive and $\operatorname{Fix} L = A\cap B$. Hence $A\cap B$ is a nonexpansive retract of $A\cap B$. Therefore, using either [42, bottom of p. 162] or [45, Lemma 3.2], $L = P_{A\cap B}$. (Alternatively, one can check directly that $L^2 = L$, which implies $x_0 - Lx_0 \in (A\cap B)^\perp$ and thus $Lx_0 = P_{A\cap B}x_0$.)

Theorem 3.1 immediately raises the following question: what can be said about the case where the sets A and B are merely two closed convex sets with nonempty intersection?

Let us recall Bregman's basic weak convergence result from 1965.

Theorem 3.3 (Bregman). *Let A and B be closed convex sets in X such that $A\cap B \neq \emptyset$. Define the sequence of alternating projections by*

$$x_0 \in X, \quad (\forall n \in \mathbb{N})\ x_{2n+1} = P_A x_{2n} \text{ and } x_{2n+2} = P_B x_{2n+1}. \tag{18}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to some point in $A \cap B$.

Proof. The original proof can be found in [15]. We include the following short proof (for another approach via *Fejér monotonicity*, see [5, Theorem 2.10(i)]). The projectors P_A and P_B are firmly nonexpansive (Fact 2.4), and hence strongly nonexpansive (Fact 2.11(i)). By Fact 2.8(i), P_BP_A is strongly nonexpansive with $\mathrm{Fix}\,(P_BP_A)=A\cap B$. Using Fact 2.13(ii), we see that $(x_{2n})_{n\in\mathbb{N}}$ converges weakly to some point $c \in A\cap B$. Now $\|x_{2n}-c\| \geqslant \|x_{2n+1}-c\| \geqslant \|x_{2(n+1)}-c\|$, for all $n\in\mathbb{N}$; consequently, $\|x_{2n}-c\|-\|P_Ax_{2n}-c\| \to 0$. Since P_A is strongly nonexpansive, it follows that $x_{2n}-x_{2n+1}=x_{2n}-P_Ax_{2n} \to 0$. Hence $(x_{2n+1})_{n\in\mathbb{N}}$ converges weakly to c, and so does the entire sequence $(x_n)_{n\in\mathbb{N}}$. □

Remark 3.4 (Extensions to the inconsistent case). For further results on the behavior of the sequence of alternating projections in the inconsistent case (i.e., when $A\cap B=\emptyset$), see [6,7,9], and the references therein.

Remark 3.5 (Extensions to Banach spaces). A closer inspection of the proof of Theorem 3.3 reveals that the following generalization holds true. Suppose X is a Banach space such that both X and X^* are uniformly convex, and C is a closed convex nonempty set in X. Let T_1, T_2 be two strongly nonexpansive maps from C to C with $F = \mathrm{Fix}\,(T_1)\cap \mathrm{Fix}\,(T_2) \neq \emptyset$. Then for every $x_0 \in C$, the sequence $(x_n)_{n\in\mathbb{N}}$ generated by

$$(\forall n \in \mathbb{N})\ x_{2n+1} = T_1 x_{2n} \quad \text{and} \quad x_{2n+2} = T_2 x_{2n+1} \tag{19}$$

converges weakly to some point in F.

4. Sequential projectors and Hundal's counterexample

The gap of knowledge between Theorems 3.1 and 3.3 was closed after nearly four decades. In 2000, during the Haifa workshop on "Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications" [19], Hein Hundal outlined his construction of two sets such that the corresponding sequence of alternating projections converges weakly, but not in norm. Moreover, since the sets are a hyperplane and a cone, his counterexample shows that there is no hope of extending Theorem 3.1 even to cones.

Let us write $\ell_2 = \mathrm{span}\,\{\mathbf{e}_1, \mathbf{e}_2, \ldots,\}$, where the nth standard unit vector $\mathbf{e}_n$ has a one at position n, and zeros elsewhere. We now describe Hundal's construction.

Fact 4.1 (Hundal's counterexample). *Let $X = \ell_2 = \mathrm{span}\,\{\mathbf{e}_1, \mathbf{e}_2, \ldots,\}$ and define $\mathbf{v}$ by*

$$\mathbf{v}: [0,+\infty[\to \ell_2$$

$$r \mapsto \exp(-100r^3)\mathbf{e}_1 + \cos((r-\lfloor r\rfloor)\pi/2)\mathbf{e}_{\lfloor r\rfloor+2} + \sin((r-\lfloor r\rfloor)\pi/2)\mathbf{e}_{\lfloor r\rfloor+3}. \tag{20}$$

Further, define Hundal's hyperplane H, cone K, *and* starting point y_0 *by*

$$H = \{\mathbf{e}_1\}^{\perp}, \quad K = \overline{\operatorname{cone}}\,\{\mathbf{v}(r) : r \geqslant 0\}, \ \textit{and} \ \ y_0 = \mathbf{v}(1). \tag{21}$$

Then the closed convex cone K *satisfies* $\sup\langle \mathbf{e}_1, K\rangle = 0$ *and* $H \cap K = \{0\}$. *Moreover,* Hundal's sequence of alternating projections $(y_n)_{n\in\mathbb{N}}$, *given by*

$$(\forall n \in \mathbb{N}) \quad y_{n+1} = P_K P_H y_n, \tag{22}$$

converges weakly to 0, *but not in norm.*

Proof. See [34]. □

Remark 4.2. Some comments regarding Fact 4.1 are in order.

(i) While lengthy, Hundal's construction is self-contained and elementary (in the sense that no external advanced results are utilized).
(ii) Hundal formulated his example with the cone K defined in (21) and with H replaced by the *halfspace* $\{\mathbf{e}_1\}^{\ominus} = \{x \in \ell_2 : \langle \mathbf{e}_1, x\rangle \leqslant 0\}$. This is fully equivalent to how we stated his example because $K \cap \operatorname{int}(\{\mathbf{e}_1\}^{\ominus}) = \emptyset$. However, for our purposes, it is more convenient to work with the hyperplane H instead.
(iii) For future use, we point out now that

$$(P_H y_n)_{n\in\mathbb{N}} \text{ converges weakly to } 0, \text{ but not in norm.} \tag{23}$$

Indeed, weak convergence is implied by the fact that $H \cap K = \{0\}$ and Theorem 3.3 (or, since H is a closed linear subspace, by the weak continuity of P_H). The lack of norm convergence is seen as follows: since $0 \in H \cap K$ and projectors are nonexpansive (Fact 2.4), we have $\|y_n\| = \|y_n - 0\| \geqslant \|P_H y_n - P_H 0\| = \|P_H y_n - 0\| \geqslant \|P_K P_H y_n - P_K 0\| = \|y_{n+1}\|$, for all $n \in \mathbb{N}$. Hence $0 < \inf_{n\in\mathbb{N}} \|y_n\| = \inf_{n\in\mathbb{N}} \|P_H y_n\|$, and thus $(P_H y_n)_{n\in\mathbb{N}}$ fails to converge to 0 in norm.

We now show that neither $P_H P_K$ nor $P_K P_H$ is firmly nonexpansive; thus, Fact 4.1 does not contain an obvious counterexample to norm convergence of iterates of firmly nonexpansive maps à la Genel and Lindenstrauss [29]. However, it is conceivable—but it seems unlikely—that some powers of these compositions are firmly nonexpansive. Conversely, the firmly nonexpansive map in the counterexample by Genel and Lindenstrauss [29] does not appear to be the product of projectors.

Lemma 4.3. *Let* H *and* K *be Hundal's hyperplane and Hundal's cone, respectively. Then neither* $P_K P_H$ *nor* $P_H P_K$ *is firmly nonexpansive.*

Proof. "$P_K P_H$": We require some notation and results from [34]. Let p, h, and P_1 be as in [34], and let $\eta \in \mathbb{R}$ satisfy [34, Eq. (14)]. For $\xi \in \mathbb{R}$, set $x(\xi) = p(\eta) + \xi \mathbf{e}_1$. For all ξ sufficiently negative, we have $\langle \mathbf{e}_1, x(\xi)\rangle \leqslant 0$. Then $\langle \mathbf{e}_1, P_H x(\xi)\rangle = 0$ for all ξ sufficiently negative and hence $P_H x(\xi) = P_1 p(\eta)$. By [34, Theorem 3.14], there exists $\hat{\eta} \in]\eta, \lfloor \eta \rfloor + 1[$ and $\hat{\alpha} \in]0,1[$ such that $P_K P_H x(\xi) = P_K P_1 p(\eta) = \hat{\alpha} p(\hat{\eta})$. Note that the last term is *independent* of ξ and that $\beta = \langle \mathbf{e}_1, \hat{\alpha} p(\hat{\eta})\rangle = \hat{\alpha} + h(\eta) > 0$. Thus

$$\begin{aligned}\langle P_K P_H x(\xi), x(\xi)\rangle &= \langle P_K P_H x(\xi), P_H x(\xi)\rangle + \langle P_K P_H x(\xi), P_{H^\perp} x(\xi)\rangle \\ &= \langle P_K P_H x(\xi), P_H x(\xi)\rangle + \langle \hat{\alpha} p(\hat{\eta}), (h(\eta)+\xi)\mathbf{e}_1\rangle \\ &= \langle P_K P_H x(\xi), P_H x(\xi)\rangle + (h(\eta)+\xi)\beta. \end{aligned} \tag{24}$$

The last term, $(h(\eta)+\xi)\beta$, can be made arbitrarily negative by choosing ξ sufficiently negative. In particular, for all ξ sufficiently negative, we have

$$\begin{aligned}\|P_K P_H x(\xi) - P_K P_H 0\|^2 = \|P_K P_H x(\xi)\|^2 &> \langle P_K P_H x(\xi), x(\xi)\rangle \\ &= \langle P_K P_H x(\xi) - P_K P_H 0, x(\xi) - 0\rangle. \end{aligned} \tag{25}$$

In view of Lemma 2.2(ii), $P_K P_H$ is not firmly nonexpansive.

"$P_H P_K$": Assume to the contrary that $P_H P_K$ is firmly nonexpansive. On the one hand, since $P_H P_K 0 = 0$, it follows that $\|P_H P_K x\|^2 \leqslant \langle P_H P_K x, x\rangle$, for all $x \in X$. On the other hand, using the Moreau decomposition (see [41, Corollaire 4.b]) $I = P_K + P_{K^\ominus}$ and the fact that P_H is a self-adjoint idempotent, we have for all $x \in X$: $\langle P_H P_K x, x\rangle = \langle P_H P_K x, P_K x + P_{K^\ominus} x\rangle = \|P_H P_K x\|^2 + \langle P_H P_K x, P_{K^\ominus} x\rangle$. Altogether, $\langle P_H P_K x, P_{K^\ominus} x\rangle \geqslant 0$ and so

$$\langle P_H(K), K^\ominus\rangle \geqslant 0. \tag{26}$$

Now set $X^+ = \{(x_n)_{n\in\mathbb{N}} \in \ell_2 : (\forall n \in \mathbb{N}) x_n \geqslant 0\}$. Then $K \subset X^+$ and $(X^+)^\ominus = -X^+$; hence, $-X^+ \subset K^\ominus$. Thus, using (26), we see that

$$\langle P_H(K), -X^+\rangle \geqslant 0. \tag{27}$$

Further, $P_H(K) \subset P_H(X^+) \subset X^+$ and hence $-P_H(K) \subset -X^+$. Now (27) yields $-\|P_H(K)\|^2 \geqslant 0$ so that $P_H(K) = \{0\}$, which is the desired contradiction. □

5. Averaged projectors

We now show that Hundal's sequence can be viewed as the iterates of an *average* of two projectors. Consequently, since Hundal's sequence fails to converge in norm, we have obtained a counterexample to the norm convergence of iterates of averaged projectors. This provides not only an answer to a question posed by Reich (see the last paragraph in [45, Section 2]) that was unresolved for two decades, but also a new counterexample à la Genel and Lindenstrauss [29].

Theorem 5.1 (An iteration of averaged projectors that fails to converge in norm). *Let H, K, and $(y_n)_{n\in\mathbb{N}}$ be Hundal's hyperplane, cone, and sequence, respectively (see Fact 4.1). Define $L = R_H(K)$, $z_0 = P_H(y_0)$, and*

$$(\forall n \in \mathbb{N}) \quad z_{n+1} = \left(\tfrac{1}{2} P_K + \tfrac{1}{2} P_L\right) z_n. \tag{28}$$

Then $(z_n)_{n\in\mathbb{N}} = (P_H y_n)_{n\in\mathbb{N}}$ and this sequence converges weakly to 0, but not in norm.

Proof. Suppose that, $z_n = P_H y_n$, for some $n \in \mathbb{N}$. Then $z_n \in H$; hence, using (22) and Corollary 2.6(iii), we deduce $y_{n+1} = P_H P_K y_n = P_H P_K z_n = \left(\frac{1}{2} P_K + \frac{1}{2} P_L\right) z_n = z_{n+1}$. By induction, $(z_n)_{n\in\mathbb{N}} = (P_H y_n)_{n\in\mathbb{N}}$. The result now follows from Remark 4.2(iii). □

Corollary 5.2 (A new counterexample à la Genel and Lindenstrauss). *Let K, L, and z_0 be as in Theorem* 5.1, *and set $T = \frac{1}{2} P_K + \frac{1}{2} P_L$. Then T is firmly nonexpansive and $(T^n z_0)_{n\in\mathbb{N}}$ converges weakly to $0 \in \operatorname{Fix} T$, but not in norm.*

Proof. By Fact 2.4, the reflectors R_K and R_L are nonexpansive, hence so is $R = \frac{1}{2} R_K + \frac{1}{2} R_L$. Now Lemma 2.2 implies that $T = \frac{1}{2} P_K + \frac{1}{2} P_L = \frac{1}{2} R + \frac{1}{2} I$ is firmly nonexpansive. Since $0 \in K$, it follows that $0 \in R_H(K) = L$ and hence $0 = T0$. The statement regarding the convergence of $(z_n)_{n\in\mathbb{N}}$ has already been observed in Theorem 5.1. □

Remark 5.3. In 1975, Genel and Lindenstrauss provided the first example of an iteration of a firmly nonexpansive map that fails to converge in norm; see [29] and also [14, pp. 72–74]. Corollary 5.2 is a new example of this kind. It is somewhat more explicit because the original construction in [29] relies upon the Kirszbraun–Valentine theorem (see [35] and also [30, Theorem 12.4]).

Remark 5.4 (Extensions to Banach spaces). The following results are drawn from [45]; their proofs also depend on the useful properties of strongly nonexpansive maps sampled in Section 2. Let $T = \sum_{i=1}^{m} \lambda_i R_i$, where $\sum_{i=1}^{m} \lambda_i = 1$, each $\lambda_i > 0$, and each R_i is a retraction onto some closed convex set C_i in X (see Definition 2.9) such that $C = \bigcap_{i=1}^{m} C_i \neq \emptyset$. Fix $x_0 \in X$. Then $(T^n x_0)_{n\in\mathbb{N}}$ converges *in norm* to some point in C if (i) X is a uniformly convex Banach space and each R_i is a linear projection of norm one onto a subspace C_i [45, Theorem 1.7]; or if (ii) X is smooth and uniformly convex, each C_i is symmetric, and each R_i is a sunny nonexpansive retraction onto C_i [45, Theorem 2.3]. Moreover, the sequence $(T^n x_0)_{n\in\mathbb{N}}$ converges at least *weakly*, provided that (iii) X is both uniformly convex and uniformly smooth, and each R_i is a sunny nonexpansive retraction onto C_i [45, Theorem 2.5].

In particular, all these results are applicable to projectors P_{C_i} in Hilbert space. Returning to our usual setting, we note that the sufficient conditions (i)—(iii) from above become the following: (i) each C_i is a closed linear subspace; (ii) each set C_i is symmetric; and (iii) is always satisfied.

6. Proximal maps

We denote the *proximal map* (see [16,41], or [49]) of a given proper lower semicontinuous convex function $f : X \to]-\infty, +\infty]$ by $\operatorname{Prox}(f)$:

$$\operatorname{Prox}(f) = (I + \partial f)^{-1} = \nabla\left(\tfrac{1}{2} \|\cdot\|^2 \Box f^*\right). \tag{29}$$

For example, if C is a nonempty closed convex set in X, then $\operatorname{Prox}(\imath_C) = P_C$. Rockafellar proved that ∂f is *maximal monotone* [48]. The *resolvent* of a maximal monotone operator $A : X \to 2^X$ is defined by $(I + A)^{-1}$. Minty characterized resolvents as firmly nonexpansive maps with full domain [40]. In particular, every proximal map is firmly nonexpansive with full domain, but the converse is false—consider, for instance, $\mathbb{R}^2 \to \mathbb{R}^2 : (\xi, \eta) \mapsto \frac{4}{5}(2\xi - \eta, 2\eta - \xi)$.

In this section, we show that the proximal maps form a convex set, a result first observed by Moreau [41] (see also Remark 6.2 below). Consequently, if H and K denote Hundal's hyperplane and cone, respectively (see Fact 4.1), then $\frac{1}{2}P_H + \frac{1}{2}P_K$, the midpoint average of the projectors P_H and P_K, is not only firmly nonexpansive (Corollary 5.2), but also a proximal map. This will be exploited in the next section, where we provide a new counterexample to norm convergence for sequences generated by the proximal point algorithm.

Theorem 6.1. *Let f_1 and f_2 be functions from X to $]-\infty, +\infty]$ that are convex, lower semicontinuous, and proper. Suppose λ_1 and λ_2 belong to $]0,1[$, $\lambda_1 + \lambda_2 = 1$, and set*

$$f = \left(\lambda_1 \left(f_1^* \Box \tfrac{1}{2}\|\cdot\|^2\right) + \lambda_2 \left(f_2^* \Box \tfrac{1}{2}\|\cdot\|^2\right)\right)^* - \tfrac{1}{2}\|\cdot\|^2. \tag{30}$$

Then f is a proper lower semicontinuous convex function such that

$$f^* = \left(\lambda_1 \left(f_1 \Box \tfrac{1}{2}\|\cdot\|^2\right) + \lambda_2 \left(f_2 \Box \tfrac{1}{2}\|\cdot\|^2\right)\right)^* - \tfrac{1}{2}\|\cdot\|^2 \tag{31}$$

and $\operatorname{Prox}(f) = \lambda_1 \operatorname{Prox}(f_1) + \lambda_2 \operatorname{Prox}(f_2)$. *If each f_i is the indicator function of a nonempty closed convex set C_i in X, then*

$$\begin{aligned} f(x) &= \lambda_1\lambda_2 \inf\left\{\tfrac{1}{2}\|c_1 - c_2\|^2 : \lambda_1 c_1 + \lambda_2 c_2 = x \text{ and each } c_i \text{ belongs to } C_i\right\} \\ &= \frac{1}{\lambda_1\lambda_2}\frac{1}{2} d^2(0, (\lambda_1(C_1 - x)) \cap (\lambda_2(x - C_2))) \end{aligned} \tag{32}$$

for every $x \in X$.

Proof. For convenience, we set $j = \frac{1}{2}\|\cdot\|^2$. It is well-known that $j = j^*$ and $\operatorname{dom} j = X$. For $i \in \{1, 2\}$, we let

$$h_i = j \Box f_i^* \quad \text{and} \quad g_i = \lambda_i h_i. \tag{33}$$

Since each f_i^* is proper, lower semicontinuous, and convex, so is each $h_i = j \Box f_i^*$; moreover, each h_i has full domain. It follows that $g_1 + g_2$ is lower semicontinuous and convex, with full domain. Thus $(g_1 + g_2)^*$ is proper, which shows that

$$f = (g_1 + g_2)^* - j \tag{34}$$

is proper as well. Later, we require

$$(g_1 + g_2)^* = g_1^* \Box g_2^*, \tag{35}$$

an identity proven by Attouch and Brézis [1]. Now fix $x^* \in X^*$. Then

$$\begin{aligned} f^*(x^*) &= \sup_{z^*} (g_1 + g_2)^{**}(x^* + z^*) - j(z^*) \\ &= \sup_{z^*} (g_1 + g_2)(x^* + z^*) - j(z^*) \\ &= \sup_{y^*} (g_1 + g_2)(y^*) - j(y^* - x^*) \\ &= \sup_{y^*} \lambda_1(j\Box f_1^*)(y^*) + \lambda_2(j\Box f_2^*)(y^*) - (j(y^*) + j(x^*) - \langle x^*, y^*\rangle) \\ &= \sup_{y^*} \lambda_1(j(y^*) - (j\Box f_1)(y^*)) + \lambda_2(j(y^*) - (j\Box f_2)(y^*)) \\ &\quad - (j(y^*) + j(x^*) - \langle y^*, x^*\rangle) \\ &= -j(x^*) + \sup_{y^*} \langle x^*, y^*\rangle - (\lambda_1(j\Box f_1) + \lambda_2(j\Box f_2))(y^*) \\ &= (\lambda_1(j\Box f_1) + \lambda_2(j\Box f_2))^*(x^*) - j(x^*), \end{aligned} \tag{36}$$

where • in the 1st equality, we used the fact that $(g_1+g_2)^*$ is proper (since g_1+g_2 is), $j=j^*$, $\operatorname{dom} j^*=X$, and [33, Theorem 2.2]; • the 2nd equality is obtained by noting that g_1+g_2 is proper, lower semicontinuous and convex, so that the Biconjugate Theorem [54, Theorem 2.3.3] holds; • we changed variables in the 3rd equality; the 4th equality follows from the definitions; • Moreau's result that $(w\Box j)+(w^*\Box j)=j$ for every proper, lower semicontinuous, and convex function w (see [41, Eq. (9.1)] or prove it directly via Fenchel duality) implies the 5th equality; • the two remaining equalities follow from the definitions. Since x^* was chosen arbitrarily, we have proved the announced identity for f^*, namely

$$((\lambda_1(j\Box f_1^*) + \lambda_2(j\Box f_2^*))^* - j)^* = (\lambda_1(j\Box f_1) + \lambda_2(j\Box f_2))^* - j. \tag{37}$$

Now this identity holds true for *any* two proper lower semicontinuous convex functions f_1, f_2 from X to $]-\infty, +\infty]$. Applying (37) to f_1^*, f_2^*, we obtain

$$((\lambda_1(j\Box f_1) + \lambda_2(j\Box f_2))^* - j)^* = (\lambda_1(j\Box f_1^*) + \lambda_2(j\Box f_2^*))^* - j. \tag{38}$$

On the one hand, the left side of (38) is a conjugate function, hence it is convex and lower semicontinuous. On the other hand, the right side of (38) is just f. Altogether, f is convex, lower semicontinuous, and (as observed in (34)) proper.

It remains to establish the identity regarding the proximal map of f. The definitions and a further dose of convex calculus now yield

$$\begin{aligned} f = (g_1 + g_2)^* - j &\Leftrightarrow j + f = (g_1 + g_2)^* \\ &\Leftrightarrow (j + f)^* = g_1 + g_2 \\ &\Leftrightarrow j\Box f^* = g_1 + g_2 \\ &\Leftrightarrow j\Box f^* = \lambda_1 h_1 + \lambda_2 h_2 \end{aligned} \tag{39}$$

Therefore, $\operatorname{Prox}(f)=\nabla(j\Box f^*)=\nabla(\lambda_1 h_1+\lambda_2 h_2)=\lambda_1\nabla(h_1)+\lambda_2\nabla(h_2)=\lambda_1\nabla(j\Box f_1^*)+\lambda_2\nabla(j\Box f_2^*)=\lambda_1\operatorname{Prox}(f_1)+\lambda_2\operatorname{Prox}(f_2)$.

Now suppose $f_i=\imath_{C_i}$, for each $i\in\{1,2\}$, and fix $x\in X$. Using (35), we obtain

$$\begin{aligned}
f(x) &= (g_1^*\Box g_2^*)(x)-j(x)\\
&= \inf_{x_1+x_2=x} g_1^*(x_1)+g_2^*(x_2)-j(x)\\
&= \inf_{x_1+x_2=x} \lambda_1\left(\tfrac{1}{2}\left\|\frac{x_1}{\lambda_1}\right\|^2+\imath_{C_1}\left(\frac{x_1}{\lambda_1}\right)\right)\\
&\quad +\lambda_2\left(\tfrac{1}{2}\left\|\frac{x_2}{\lambda_2}\right\|^2+\imath_{C_2}\left(\frac{x_2}{\lambda_2}\right)\right)-\tfrac{1}{2}\|x_1+x_2\|^2\\
&= \inf_{x_1+x_2=x \text{ and each } x_i/\lambda_i\in C_i} \lambda_1\tfrac{1}{2}\left\|\frac{x_1}{\lambda_1}\right\|^2+\lambda_2\tfrac{1}{2}\left\|\frac{x_2}{\lambda_2}\right\|^2-\tfrac{1}{2}\|x_1+x_2\|^2\\
&= \inf_{\lambda_1c_1+\lambda_2c_2=x \text{ and each } c_i\in C_i} \lambda_1\tfrac{1}{2}\|c_1\|^2+\lambda_2\tfrac{1}{2}\|c_2\|^2-\tfrac{1}{2}\|\lambda_1c_1+\lambda_2c_2\|^2\\
&= \inf_{\lambda_1c_1+\lambda_2c_2=x \text{ and each } c_i\in C_i} \tfrac{1}{2}\lambda_1\lambda_2\|c_1-c_2\|^2, \qquad (40)
\end{aligned}$$

which establishes the first equality in (32). If each c_i belongs to C_i and $x=\lambda_1c_1+\lambda_2c_2$, then $c_2\in(x-\lambda_1C_1)/\lambda_2$, $c_1\in(x-\lambda_2C_2)/\lambda_1$, and $\|c_2-c_1\|=\|x-c_1\|/\lambda_2=\|x-c_2\|/\lambda_1$. This implies that

$$\begin{aligned}
f(x) &= \frac{1}{2}\lambda_1\lambda_2\frac{1}{\lambda_2^2}d^2(x,C_1\cap((x-\lambda_2C_2)/\lambda_1))\\
&= \frac{1}{2}\frac{\lambda_1}{\lambda_2}d^2(0,(C_1-x)\cap((x-\lambda_1x-\lambda_2C_2)/\lambda_1))\\
&= \frac{1}{2}\frac{\lambda_1}{\lambda_2}d^2(0,(C_1-x)\cap((\lambda_2/\lambda_1)(x-C_2)))\\
&= \frac{1}{2}\frac{\lambda_1}{\lambda_2}\frac{1}{\lambda_1^2}d^2(0,(\lambda_1(C_1-x))\cap(\lambda_2(x-C_2))). \qquad (41)
\end{aligned}$$

Therefore, the second equality of (32) is verified. □

Remark 6.2. Theorem 6.1 implies that the proximal maps form a convex set, an observation originally due to Moreau [41, Proposition 9.d]. The present proof appears to be more straightforward and it also provides an explicit formula of the function corresponding to the convex combination of the proximal maps. This will aid us in the derivation of the function f in Corollary 7.1 below.

7. Proximal point algorithm

We are now in a position to derive a new counterexample to the norm convergence of the proximal point algorithm à la Güler.

Corollary 7.1 (A new proximal point iteration that fails to converge in norm). *Let* K, L, *and* z_0 *be as in Theorem* 5.1, *and set* $f(x) = \frac{1}{2}d^2(0,(K-x)\cap(x-L))$, *for every* $x \in X$. *Then* f *is convex*, *lower semicontinuous*, *proper*, *and its proximal map is* $\mathrm{Prox}(f) = \frac{1}{2}P_K + \frac{1}{2}P_L$. *Moreover*, *the sequence* $((\mathrm{Prox}(f))^n z_0)_{n\in\mathbb{N}}$ *converges weakly to* 0, *but not in norm.*

Proof. Combine Corollary 5.2 and Theorem 6.1 (where $\lambda_1 = \lambda_2 = \frac{1}{2}$). □

Remark 7.2 (Proximal point algorithm). Let $f: X \to]-\infty,+\infty]$ be convex, lower semicontinuous, and proper. Recall that (see the classical papers by Martinet [38] and by Rockafellar [49]) given a sequence of strictly positive parameters $(\gamma_n)_{n\in\mathbb{N}}$ and a starting point $x_0 \in X$, the inductive update rule of the *proximal point algorithm* for minimizing f is

$$\begin{aligned} x_{n+1} &= \underset{x\in X}{\mathrm{argmin}} \left(f(x) + \frac{1}{2\gamma_n}\|x - x_n\|^2 \right) \\ &= (I + \gamma_n \partial f)^{-1}(x_n) \\ &= \left(\nabla\left(\tfrac{1}{2}\|\cdot\|^2 + \gamma_n f\right)^*\right)(x_n) \\ &= (\mathrm{Prox}(\gamma_n f))(x_n) \end{aligned} \tag{42}$$

for all $n \in \mathbb{N}$. Brézis and Lions [17] show that if f has minimizers and $\sum_{n\in\mathbb{N}} \gamma_n = +\infty$, then the sequence generated by (42) converges *weakly* to a minimizer of f.

The question whether norm convergence always holds remained open until 1991, when Güler [32] constructed a whole family of counterexamples for all strictly positive parameter sequences $(\gamma_n)_{n\in\mathbb{N}}$ such that $\sum_n \gamma_n = +\infty$.

In [10], Fact 4.1 has been interpreted as a simple counterexample to norm convergence of the proximal point algorithm with $\gamma_n \to 0$. We now realize that Corollary 7.1 provides another counterexample of this type, for the case when $\gamma_n \equiv 1$.

8. String-averaging projection methods

Let C_1, C_2, and C_3 be closed convex sets in X with corresponding projectors P_1, P_2, and P_3, respectively. Suppose further that $C = C_1 \cap C_2 \cap C_3 \neq \emptyset$ and define

$$T = \tfrac{1}{2}P_1P_2 + \tfrac{1}{4}P_2 + \tfrac{1}{4}P_3. \tag{43}$$

Note that T is neither a composition nor an average of the given projectors. However, the iterates of the operator T can be analyzed within the very flexible *string-averaging algorithmic structure* proposed recently by Censor et al. [21] (see also [20,22,26]).

Roughly speaking, a *string* corresponds to a point obtained through the sequential application of operators to the current point, and the update step consists of averaging the resulting strings. This framework is clearly inspired by parallel computing architectures, where each processor can work independently on its string. Thus, the application of the operator T can be viewed as a string-averaging algorithm, where P_1P_2, P_2, P_3 are three strings that are averaged by the weights $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{4}$, respectively.

We now provide a prototypical convergence result for the iterates of the operator T.

Theorem 8.1. *Let T be defined by* (43) *and* fix $x_0 \in X$. *Then the sequence $(T^n x_0)_{n\in\mathbb{N}}$ converges weakly to some point in C. If each set C_i is* symmetric (*i.e.*, $C_i = -C_i$), *then $(T^n x_0)_{n\in\mathbb{N}}$ converges in norm.*

Proof. Each projector is firmly nonexpansive hence strongly nonexpansive (Fact 2.11(i)). Repeated use of Fact 2.8 shows that the map T is strongly nonexpansive as well, and its fixed point set is C. Fact 2.13(ii) now implies that $(T^n x_0)_{n\in\mathbb{N}}$ converges weakly to a point in C.

Assume that each set C_i is symmetric. Then C is symmetric as well. Also, [45, Lemma 2.2] implies that each P_i is *odd* ($P_i(-x) = -P_i x$, for all $x \in X$). Thus T is odd. The norm convergence of $(T^n x_0)_{n\in\mathbb{N}}$ now follows from Fact 2.13(iii). □

Remark 8.2. It is clear that the convergence proof of Theorem 8.1 will also work for iterations of various other maps as long as they are assembled from projectors by averaging and by composition. (See [8, Examples 2.14 and 2.20] for another example of a string-averaging method.) For convergence results in finite-dimensional spaces, see [20–22,26].

Remark 8.3 (Extensions to Banach spaces). A closer inspection of its proof reveals that Theorem 8.1 holds true under the following more general conditions: X is a Banach space such that both X and its dual X^* are uniformly convex, and each C_i is the fixed point set of a strongly nonexpansive map P_i. The norm convergence requires additionally that each P_i be odd.

We now combine Theorem 8.1 with Theorem 5.1 to obtain a counterexample to norm convergence for string-averaging projection methods.

Corollary 8.4. *Let H, K, y_0 be Hundal's hyperplane, cone, and starting point, respectively. Further, define $L = R_H(K)$ and*

$$T = \tfrac{1}{2}P_H P_K + \tfrac{1}{4}P_K + \tfrac{1}{4}P_L. \tag{44}$$

Then the sequence $(T^n P_H y_0)_{n\in\mathbb{N}}$ converges weakly to 0, *but not in norm.*

Proof. By Fact 4.1, $H \cap K \cap L = \{0\}$. Hence, Theorem 8.1 implies that $(T^n P_H y_0)_{n\in\mathbb{N}}$ converges weakly to 0. Suppose $x \in H$. By Corollary 2.6(iii), $P_H P_K x = \frac{1}{2}P_K x + \frac{1}{2}P_L x$ and thus $Tx = P_H P_K$. Using induction, it follows that $(T^n P_H y_0)_{n\in\mathbb{N}}$ is equal to the sequence $(z_n)_{n\in\mathbb{N}}$ of Theorem 5.1; consequently, it does not converge in norm. □

Remark 8.5. In fact, a second glance at the proof of Corollary 8.4 reveals that the conclusion of this result holds whenever

$$T = \omega_1(P_H P_K) + \omega_2(P_H P_L) + \omega_3\left(\tfrac{1}{2}P_K + \tfrac{1}{2}P_L\right), \tag{45}$$

where $\omega_1, \omega_2, \omega_3$ belong to $[0,1]$ and $\omega_1 + \omega_2 + \omega_3 = 1$.

9. The Hilbert ball

Throughout this last section, we let Y be a complex Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. Suppose that $X = \{x \in Y : \|x\| < 1\}$ is the open unit ball of Y and define

$$(\forall x \in X)(\forall y \in X) \quad \sigma(x,y) = \frac{(1-\|x\|^2)(1-\|y\|^2)}{|1-\langle x,y\rangle|^2}$$

and

$$\rho(x,y) = \operatorname{arctanh}\sqrt{1-\sigma(x,y)}. \tag{46}$$

Then (X,ρ) is a complete metric space, commonly referred to as the *Hilbert ball* (see [31, p. 91 and Section 2.15]; for further background material on this and on what follows, see [31, Chapter 2].

Given x and y in X, and $\lambda \in [0,1]$, the results in [31, Section 2.17] show that there exists a unique point $z \in X$ such that $\rho(x,z) = \lambda\rho(x,y)$ and $\rho(z,y) = (1-\lambda)\rho(x,y)$; this point is denoted by

$$(1-\lambda)x \oplus \lambda y \tag{47}$$

and it is a *ρ-convex combination* of x and y. A set S in X is called *ρ-convex* (respectively, *ρ-closed*), if it contains all its ρ-convex combinations (respectively, if it is closed in (X,ρ)).

Given $T : X \to X$, let us define

$$(\forall x \in X)(\forall y \in X)$$

$$\Phi_{x,y} : [0,1] \to [0,+\infty[: \lambda \mapsto \rho((1-\lambda)x \oplus \lambda Tx, (1-\lambda)y \oplus \lambda Ty). \tag{48}$$

Analogously to (5), one says that T is *firmly ρ-nonexpansive of the first kind* (respectively, *ρ-nonexpansive*), if $\Phi_{x,y}$ is decreasing (respectively, $\Phi_{x,y}(0) \geqslant \Phi_{x,y}(1)$) for all x and y in C (see [31, pp. 73, 124]). As in the Hilbert space case (6), $T : X \to X$ is ρ-nonexpansive if and only if

$$(\forall x \in X)(\forall y \in X) \quad \rho(Tx,Ty) \leqslant \rho(x,y), \tag{49}$$

in which case $\operatorname{Fix} T$ is ρ-closed and ρ-convex [31, Theorem 2.23.2].

Fact 9.1 (Projectors in the Hilbert ball). *Suppose C is a ρ-closed and ρ-convex set in X. Then for every $x \in X$, there exists a unique point in C, denoted $\tilde{P}_C x$, such that $\rho(x,\tilde{P}_C x) = \inf_{y \in C}\rho(x,y)$. The induced map $\tilde{P}_C : X \to C$ is firmly ρ-nonexpansive of the first kind.*

Proof. See [31, Section 2.19]. □

Given $y \in X$ and $s > 0$, the set $\{x \in X : \rho(x, y) < s\}$ is the open *ρ-ball* of radius s centered at y (see also [31, Section 15] for characterizations). A set S in X is *ρ-bounded*, if it is contained in some open ρ-ball. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is ρ-bounded, if its orbit $\{x_n : n \in \mathbb{N}\}$ is.

A map $T : X \to X$ is called *para-strongly ρ-nonexpansive* (see [47], where this was called strongly nonexpansive for brevity), if T is ρ-nonexpansive, $\mathrm{Fix}\, T \neq \emptyset$, and for every ρ-bounded sequence $(x_n)_{n\in\mathbb{N}}$ and every $y \in \mathrm{Fix}\, T$, the condition $\rho(x_n, y) - \rho(Tx_n, y) \to 0$ implies $\rho(x_n, Tx_n) \to 0$.

Analogously to Definition 2.9, a map $T : X \to C \subset X$ is called a *ρ-retraction* onto C, if T is continuous in (X, ρ) and $\mathrm{Fix}\, T = C$.

The following basic results on para-strongly ρ-nonexpansive maps are due to Reich [47].

Fact 9.2. *Let T, T_1, T_2 be maps from X to X with fixed point sets F, F_1, F_2, respectively.*

(i) *If T is firmly ρ-nonexpansive of the first kind and $F \neq \emptyset$, then T is para-strongly ρ-nonexpansive.*

(ii) *If T is para-strongly ρ-nonexpansive, then the map*

$$X \to F : x \mapsto \text{ weak } \lim_n T^n x \tag{50}$$

is a well-defined ρ-nonexpansive ρ-retraction onto F.

(iii) *If each T_i is para-strongly ρ-nonexpansive and $F_1 \cap F_2 \neq \emptyset$, then T_2T_1 is para-strongly ρ-nonexpansive and $F = F_1 \cap F_2$.*

Proof. (i): [47, Lemma 2]. (ii): [47, Lemmata 5 and 6]. (iii): [47, Lemmata 3 and 4]. □

Corollary 9.3. *Suppose A and B are ρ-closed ρ-convex sets in X such that $A \cap B \neq \emptyset$. Then for every $x \in X$, the sequence $((\tilde{P}_B\tilde{P}_A)^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $A \cap B$.*

Proof. (See also [47].) The projectors $\tilde{P}_A$ and $\tilde{P}_B$ are firmly ρ-nonexpansive of the first kind (Fact 9.1), hence para-strongly ρ-nonexpansive (Fact 9.2(i)). By Fact 9.2(iii), $T = \tilde{P}_B\tilde{P}_A$ is para-strongly ρ-nonexpansive and $\mathrm{Fix}\, T = A \cap B$. The weak convergence statement now follows from Fact 9.2(ii). □

Remark 9.4. Reich asks at the end of [47] whether the convergence in Corollary 9.3 is actually strong. While Hundal settled the corresponding question in Hilbert space (Fact 4.1), the problem posed by Reich remains open in the Hilbert ball.

The next result exhibits two analogues of Fact 2.8(ii).

Theorem 9.5. *Let T_1, T_2 be para-strongly ρ-nonexpansive maps from X to X such that* $\operatorname{Fix}(T_1) \cap \operatorname{Fix}(T_2) \neq \emptyset$. *Suppose λ_1, λ_2 are in $]0,1[$ and $\lambda_1 + \lambda_2 = 1$. Now define T in one of the following two ways.*

(i) $T : X \to X : x \mapsto \lambda_1 T_1 x \oplus \lambda_2 T_2 x$;
(ii) $T : X \to X : x \mapsto \lambda_1 T_1 x + \lambda_2 T_2 x$.

Then T is para-strongly ρ-nonexpansive with $\operatorname{Fix} T = \operatorname{Fix}(T_1) \cap \operatorname{Fix}(T_2)$.

Proof. (i). Fact 2.8(ii) states that the corresponding Hilbert (and even Banach) space statement is true. It turns out that the corresponding proofs of [47, Lemmata 3 and 4] carry over if we utilize [31, Lemma 2.17.1] at the appropriate places.

(ii). Clearly, $\operatorname{Fix}(T_1) \cap \operatorname{Fix}(T_2) \subset \operatorname{Fix} T$. Conversely, fix $y \in \operatorname{Fix}(T_1) \cap \operatorname{Fix}(T_2)$ and pick $x \in \operatorname{Fix} T$. By [36, Lemma 3.3(i)], $\rho(x, y) = \rho(Tx, Ty) = \rho(Tx, y) \leqslant \max\{\rho(T_1 x, y), \rho(T_2 x, y)\}$. Without loss of generality, assume that $\max\{\rho(T_1 x, y), \rho(T_2 x, y)\} = \rho(T_1 x, y)$. Since T_1 is para-strongly ρ-nonexpansive, we obtain $\rho(x, y) \leqslant \rho(T_1 x, y) \leqslant \rho(x, y)$ and hence $T_1 x = x$. Since $x = Tx$, this implies $T_2 x = x$. We conclude that $\operatorname{Fix} T \subset \operatorname{Fix}(T_1) \cap \operatorname{Fix}(T_2)$. Next, pick a ρ-bounded sequence $(x_n)_{n \in \mathbb{N}}$ such that

$$\rho(x_n, y) - \rho(Tx_n, y) \to 0. \tag{51}$$

We have to show that $\rho(x_n, Tx_n) \to 0$. Assume to the contrary this were false. Using [36, Lemma 3.3(i)] again, and after repeatedly passing to subsequences and relabelling if necessary, we obtain $\varepsilon > 0$ and $M \geqslant 0$ such that

$$(\forall n \in \mathbb{N}) \ \ \varepsilon \leqslant \rho(x_n, Tx_n), \ \ \max\{\rho(Tx_n, y), \rho(T_2 x_n, y)\} \leqslant \rho(T_1 x_n, y), \\ \text{and} \ \ \rho(x_n, y) \to M. \tag{52}$$

Hence $0 \leftarrow \rho(Tx_n, y) - \rho(x_n, y) \leqslant \rho(T_1 x_n, y) - \rho(x_n, y) \leqslant 0$ and thus

$$\rho(T_1 x_n, y) - \rho(x_n, y) \to 0. \tag{53}$$

Since T_1 is para-strongly ρ-nonexpansive, it follows that

$$\rho(T_1 x_n, x_n) \to 0. \tag{54}$$

Moreover, (51)–(53) imply that

$$\rho(T_1 x_n, y) \to M, \quad \limsup \rho(T_2 x_n, y) \leqslant M \quad \text{and} \quad \rho(Tx_n, y) \to M. \tag{55}$$

Now let $(u_n)_{n \in \mathbb{N}}$ and $(v_n)_{n \in \mathbb{N}}$ be the two sequences in X that are uniquely defined by the following properties:

$$(\forall n \in \mathbb{N}) \quad \begin{cases} u_n \in \operatorname{conv}\{T_1 x_n, Tx_n\}, \quad v_n \in \operatorname{conv}\{Tx_n, T_2 x_n\} \quad \text{and} \\ \|u_n - Tx_n\| = \|v_n - Tx_n\| = \min\{\|T_1 x_n - Tx_n\|, \quad \|T_2 x_n - Tx_n\|\}. \end{cases} \tag{56}$$

Then, by construction,

$$(\forall n \in \mathbb{N}) \quad \tfrac{1}{2} u_n + \tfrac{1}{2} v_n = Tx_n \quad \text{and} \quad \|T_1 x_n - T_2 x_n\| = \frac{\|u_n - v_n\|}{2 \min\{\lambda_1, \lambda_2\}}. \tag{57}$$

Using [36, Lemma 3.3(i)] and (52), we obtain $\rho(u_n, y) \leqslant \max\{\rho(T_1 x_n, y), \rho(T x_n, y)\} = \rho(T_1 x_n, y)$ and $\rho(v_n, y) \leqslant \max\{\rho(T x_n, y), \rho(T_2 x_n, y)\} \leqslant \rho(T_1 x_n, y)$, for every $n \in \mathbb{N}$. Combining these inequalities with (55) and (57), we see that

$$\limsup \rho(u_n, y) \leqslant M, \quad \limsup \rho(v_n, y) \leqslant M \quad \text{and}$$

$$\lim \rho\left(\tfrac{1}{2} u_n + \tfrac{1}{2} v_n, y\right) = M. \tag{58}$$

Now (58) and [46, Lemma 4] result in $u_n - v_n \to 0$; equivalently (see (57)), $T_1 x_n - T_2 x_n \to 0$. By (55), the sequences $(T_1 x_n)_{n\in\mathbb{N}}$ and $(T_2 x_n)_{n\in\mathbb{N}}$ are both ρ-bounded; therefore, using [36, Theorem 3.4], we conclude that $\rho(T_1 x_n, T_2 x_n) \to 0$. The triangle inequality and (54) yield

$$\rho(T_2 x_n, x_n) \to 0. \tag{59}$$

Finally, (52), (54), (59), and [36, Lemma 3.3(i)] imply the contradiction $\varepsilon \leqslant \rho(T x_n, x_n) \leqslant \max\{\rho(T_1 x_n, x_n), \rho(T_2 x_n, x_n)\} \to 0$. □

Corollary 9.6. *Suppose A and B are ρ-closed ρ-convex sets in X such that $A \cap B \neq \emptyset$. Then for every $x \in X$, the sequences*

$$\left(\left(\tfrac{1}{2}\tilde{P}_A \oplus \tfrac{1}{2}\tilde{P}_B\right)^n x\right)_{n\in\mathbb{N}} \quad \textit{and} \quad \left(\left(\tfrac{1}{2}\tilde{P}_A + \tfrac{1}{2}\tilde{P}_B\right)^n x\right)_{n\in\mathbb{N}} \tag{60}$$

both converge weakly to some point in $A \cap B$.

Proof. Combine Theorem 9.5 with Fact 9.2(ii). □

Remark 9.7. Whether the convergence in Corollary 9.6 is actually strong is an interesting open problem. Theorem 5.1 illustrates the failure of norm convergence in the corresponding Hilbert space setting.

Remark 9.8. A map $T : X \to X$ is *ρ-averaged* (respectively, *ρ-averaged of the second kind*), if it is of the form $T = (1-\lambda)I \oplus \lambda T'$ (respectively, $T = (1-\lambda)I + T'$), where T' is ρ-nonexpansive and $\lambda \in [0, 1[$; see [46]—if furthermore $\lambda = \frac{1}{2}$, we say that the map is *ρ-midpoint averaged*. It is known that if T is ρ-averaged of either kind with $\operatorname{Fix} T \neq \emptyset$ and $x \in X$, then $(T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $\operatorname{Fix} T$ (see [46, Theorems 3 and 5]).

Remark 9.9. Kuczumow and Stachura [37] (see also [36, Example 10.6 p. 475]) constructed ρ-midpoint averaged maps of both types such that some sequence of iterates fails to converge in norm—these counterexamples are similar to (and their construction is based upon) the corresponding counterexample by Genel and Lindenstrauss in Hilbert space [29]. We do not know whether the Kuczumow–Stachura maps can be expressed as either $\frac{1}{2}\tilde{P}_A \oplus \frac{1}{2}\tilde{P}_B$ or $\frac{1}{2}\tilde{P}_A + \frac{1}{2}\tilde{P}_B$, for some ρ-closed ρ-convex sets A, B in X.

We conclude with the following striking difference between Hilbert space and the Hilbert ball. If C is a closed convex nonempty set in a Hilbert space, then (Fact 2.4) R_C is nonexpansive and thus $P_C = \frac{1}{2}I + \frac{1}{2}R_C$ is (even midpoint) averaged. In the Hilbert ball, the corresponding statement is known to be false; see [31, Example 2.22.1].

Acknowledgements

We thank Patrick Combettes for providing several helpful comments and references. H.H. Bauschke's work was partially supported by the Natural Sciences and Engineering Research Council of Canada. E. Matoušková's work was partially supported by the Czech Academy of Sciences Grant GAAV-A1019103 and by Kepler University in Linz, Austria. S. Reich's work was partially supported by the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities (Grant 592/00), by the Fund for the Promotion of Research at the Technion (Grant 100-104), and by the Technion VPR Fund—Argentinian Research Fund (Grant 100-217).

References

[1] H. Attouch, H. Brézis, Duality for the sum of convex functions in general Banach spaces, in: J.A. Barroso (Ed.), Aspects of Mathematics and its Applications, North-Holland Math. Library, Vol. 34, North-Holland, Amsterdam, The Netherlands, 1986, pp. 125–133.

[2] A. Auslender, Méthodes Numériques pour la Résolution des Problèmes d'Optimisation avec Constraintes, Thèse, Faculté des Sciences, Grenoble, 1969.

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

[4] H.H. Bauschke, Projection Algorithms and Monotone Operators, Ph.D. Thesis, Simon Fraser University, 1996. Available as Preprint 96:080 at www.cecm.sfu.ca/preprints/1996pp.html.

[5] H.H. Bauschke, Projection algorithms: results and open problems, in: D. Butnariu, Y. Censor, S. Reich (Eds.), Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, Haifa, 2000, Elsevier, Amsterdam, The Netherlands, 2001, pp. 11–22.

[6] H.H. Bauschke, J.M. Borwein, On the convergence of von Neumann's alternating projection algorithm for two sets, Set-Valued Analysis, 1 (1993) 185–212.

[7] H.H. Bauschke, J.M. Borwein, Dykstra's alternating projection algorithm for two sets, J. Approx. Theory 79 (1994) 418–443.

[8] H.H. Bauschke, J.M. Borwein, On projection algorithms for solving convex feasibility problems, SIAM Rev. 38 (1996) 367–426.

[9] H.H. Bauschke, J.M. Borwein, A.S. Lewis, The method of cyclic projections for closed convex sets in Hilbert space, in: Y. Censor, S. Reich (Eds.), Recent Developments in Optimization Theory and Nonlinear Analysis, Jerusalem, 1995, Contemporary Mathematics, Vol. 204, American Mathematical Society, Providence, RI, 1997, pp. 1–38.

[10] H.H. Bauschke, J.V. Burke, F.R. Deutsch, H.S. Hundal, J.D. Vanderwerff, A new proximal point iteration that converges weakly but not in norm, Proc. Amer. Math. Soc., available at http://www.cecm.sfu.ca/preprints/2002pp.html as Preprint 02:191.

[11] H.H. Bauschke, P.L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, Math. Oper. Res. 26 (2001) 248–264.

[12] H.H. Bauschke, F. Deutsch, H. Hundal, S.-H. Park, Fejér monotonicity and weak convergence of an accelerated method of projections, in: M. Théra (Ed.), Constructive, Experimental, and Nonlinear Analysis, Limoges, 1999, American Mathematical Society, Providence, RI, 2000, pp. 1–6.

[13] H.H. Bauschke, F. Deutsch, H. Hundal, S.-H. Park, Accelerating the convergence of the method of alternating projections, Trans. Am. Math. Soc. 355 (2003) 3433–3461.

[14] Y. Benyamini, J. Lindenstrauss, Geometric Nonlinear Functional Analysis, Vol. 1, American Mathematical Society, Providence, RI, 2000.

[15] L.M. Bregman, The method of successive projection for finding a common point of convex sets, Sov. Math. Dokl. 6 (1965) 688–692.

[16] H. Brézis, Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert, North-Holland, Amsterdam, 1973.

[17] H. Brézis, P.-L. Lions, Produits infinis de résolvantes, Isr. J. Math. 29 (1978) 329–345.

[18] R.E. Bruck, S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, Houston J. Math. 3 (1977) 459–470.

[19] D. Butnariu, Y. Censor, S. Reich (Eds.), Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, Elsevier, Amsterdam, The Netherlands, 2001.

[20] Y. Censor, On sequential and parallel projection algorithms for feasibility and optimization, in: Y. Censor, M. Ding (Eds.), Visualization and Optimization Techniques, Proceedings of SPIE, Vol. 4553, SPIE—The International Society for Optical Engineering, Bellingham, WA, 2001, pp. 1–9.

[21] Y. Censor, T. Elfving, G.T. Herman, Averaging strings of sequential iterations for convex feasibility problems, in: D. Butnariu, Y. Censor, S. Reich (Eds.), Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, Haifa, 2000, Elsevier, Amsterdam, The Netherlands, 2001, pp. 101–114.

[22] Y. Censor, E. Tom, Convergence of string-averaging projection schemes for inconsistent convex feasibility problems, Optim. Methods Software 18 (2003) 543–554.

[23] Y. Censor, S.A. Zenios, Parallel Optimization, Oxford University Press, New York, NY, 1997.

[24] G. Cimmino, Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari, La Ricerca Scientifica ed il Progresso Tecnico nell'Economia Nazionale, Roma, Vol. 1, 1938, pp. 326–333.

[25] P.L. Combettes, Hilbertian convex feasibility problem: convergence of projection methods, Appl. Math. Optim. 35 (1997) 311–330.

[26] G. Crombez, Finding common fixed points of strict paracontractions by averaging strings of sequential iterations, J. Nonlinear and Convex Anal. 3 (2002) 345–351.

[27] F. Deutsch, Best Approximation in Inner Product Spaces, Springer, New York, NY, 2001.

[28] J.M. Dye, S. Reich, On the unrestricted iteration of projections in Hilbert space, J. Math. Anal. Appl. 156 (1991) 101–119.

[29] A. Genel, J. Lindenstrauss, An example concerning fixed points, Isr. J. Math. 22 (1975) 81–86.

[30] K. Goebel, W.A. Kirk, Topics in Metric Fixed Point Theory, Cambridge University Press, Cambridge, UK, 1990.

[31] K. Goebel, S. Reich, Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings, Marcel Dekker, New York, NY, 1984.

[32] O. Güler, On the convergence of the proximal point algorithm for convex minimization, SIAM J. Control Optim. 29 (1991) 403–419.

[33] J.-B. Hiriart-Urruty, A general formula on the conjugate of the difference of functions, Can. Math. Bull. 29 (1986) 482–485.

[34] H. Hundal, An alternating projection that does not converge in norm, Nonlinear Anal., available at http://www.cecm.sfu.ca/preprints/2002pp.html as Preprint 02:189.

[35] M.D. Kirszbraun, Über die zusammenziehenden und Lipschitzschen Transformationen, Fund Math. 22 (1934) 77–108.

[36] T. Kuczumow, S. Reich, D. Shoikhet, Fixed points of holomorphic mappings: a metric approach, in: W.A. Kirk, B. Sims (Eds.), Handbook of Metric Fixed Point Theory, Kluwer, Dordrecht, The Netherlands, 2001, pp. 437–515.

[37] T. Kuzcumow, A. Stachura, Extensions of nonexpansive mappings in the Hilbert ball with the hyperbolic metric II, Commentat. Math. Univ. Carolin. 29 (1988) 403–410.

[38] B. Martinet, Régularisation d'inéquations variationnelles par approximations successives, Rev. Fr. d'Informatique Recherche Opér. 4 (1970) 154–158.

[39] Y.I. Merzlyakov, On a relaxation method of solving systems of linear inequalities, USSR Comput. Math. Math. Phys. 2 (1963) 504–510.

[40] G.J. Minty, Monotone (nonlinear) operators in Hilbert space, Duke Math. J. 29 (1962) 341–346.

[41] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. France 93 (1965) 273–299.

[42] S. Reich, Product formulas, nonlinear semigroups, and accretive operators, J. Funct. Anal. 36 (1980) 147–168.

[43] S. Reich, On the asymptotic behavior of nonlinear semigroups and the range of accretive operators, J. Math. Anal. Appl. 79 (1981) 113–126.

[44] S. Reich, On the asymptotic behavior of nonlinear semigroups and the range of accretive operators II, J. Math. Anal. Appl. 87 (1982) 134–146.
[45] S. Reich, A limit theorem for projections, Linear and Multilinear Algebra 13 (1983) 281–290.
[46] S. Reich, Averaged mappings in the Hilbert ball, J. Math. Anal. Appl. 109 (1985) 199–206.
[47] S. Reich, The alternating algorithm of von Neumann in the Hilbert ball, Dynamic Systems Appl. 2 (1993) 21–25.
[48] R.T. Rockafellar, On the maximal monotonicity of subdifferential mappings, Pacific J. Math. 33 (1970) 209–216.
[49] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976) 877–898.
[50] J. von Neumann, On rings of operators. Reduction theory, Ann. Math. 50 (1949) 401–485.
[51] J. von Neumann, Functional Operators II: The Geometry of Orthogonal Spaces, Princeton University Press, Princeton, NJ, 1950 (This is a reprint of mimeographed lecture notes first distributed in 1933.).
[52] N. Wiener, On the factorization of matrices, Comment. Math. Helv. 29 (1955) 97–111.
[53] I. Yamada, The hybrid steepest descent method for the variational inequality problem over the intersection of fixed points sets of nonexpansive mappings, in: D. Butnariu, Y. Censor, S. Reich (Eds.), Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, Haifa, 2000, Elsevier, Amsterdam, The Netherlands, 2001, pp. 473–504.
[54] C. Zălinescu, Convex Analysis in General Vector Spaces, World Scientific, River Edge, NJ, 2002.
[55] E.H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory, in: E.H. Zarantonello (Ed.), Contributions to Nonlinear Functional Analysis, Academic Press, New York, NY, 1971, pp. 237–424.