# Projection and proximal point methods: convergence results and counterexamples

Heinz H. Bauschke, Eva Matoušková, and Simeon Reich

June 14, 2003 — Version 1.15

**Abstract**

Recently, Hundal has constructed a hyperplane $H$, a cone $K$, and a starting point $y_0$ in $\ell_2$ such that the sequence of alternating projections $\big((P_K P_H)^n y_0\big)_{n\in\mathbb{N}}$ 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, Elfving, and Herman. Extensions to Banach spaces and the situation for the Hilbert ball are discussed as well.

**2000 Mathematics Subject Classification:** 47H09, 47J25, 90C25.

**Keywords:** Alternating projections, averaged projections, Hilbert space, nonexpansive, proximal point algorithm, weak convergence.

## 1 Introduction

Throughout (most of) this paper, we assume that

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

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

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

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

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

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 [50] 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] and [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

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

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]; 1938) and Merzlyakov's method of extrapolated averaged projectors for solving linear inequalities ([39]; 1963). 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] and [49]) and for the string-averaging projection method recently introduced by Censor, Elfving, and Herman [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, $S^\ominus = \{x^* \in X : (\forall s \in S)\ \langle x^*, s\rangle \leq 0\}$, $\overline{\operatorname{span}}\, S$, $\overline{\operatorname{cone}}\, S$, $\operatorname{conv} S$, $\operatorname{int} S$, $\iota_S$) 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 $\partial f$ (respectively, $\nabla f$, $f^*$), 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*

$$(5) \qquad (\forall x \in C)(\forall y \in C) \quad \Phi_{x,y}\colon [0,1] \to [0, +\infty[\ : \lambda \mapsto \big\|\big((1-\lambda)x + \lambda Tx\big) - \big((1-\lambda)y + \lambda Ty\big)\big\|.$$

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

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

$$(6) \qquad (\forall x \in C)(\forall y \in C) \quad \|Tx - Ty\| \leq \|x - y\|.$$

**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) $(\forall x \in C)(\forall y \in C)$ $\|Tx - Ty\|^2 \leq \langle Tx - Ty, x - y\rangle$.

(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*

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

*The corresponding map $P_C \colon X \to C$ 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 1.11.2], 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 $P_{U(C)} = UP_CU^*$.*

*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) $(P_H P_C)\big|_H = \frac{1}{2}(P_C + P_D)\big|_H$.

*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)&(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 non-orthogonal 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 $F = \bigcap_{i=1}^{m} \operatorname{Fix} T_i \neq \varnothing$, then $\operatorname{Fix} T = F$.*

*Proof.* [18, Proposition 1.1] and [18, Proposition 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 $\mu \geq 0$). 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* $\lambda \in [0,1[$.

(iii) $T : X \to X$ *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 $F = \varnothing$, then $\lim_n \|T^n x\| = +\infty$.*

(ii) *If $F \neq \varnothing$, 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 $\ell_p$ 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*

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

*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

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

In particular,

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

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

$$(\forall n \in \mathbb{N}) \quad (\forall k \in \mathbb{N})(\forall l \in \mathbb{N}) \qquad 1 \le k = l - n \quad \Rightarrow \quad \|x_k - x_l\|^2 \le \|x_k\|^2 - \|x_l\|^2. \tag{11}$$

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

*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

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

*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

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

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 $\|x_{n+1} - c\| \le \|x_n - c\|$. After squaring and simplifying, this inequality turns into

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

Since $n$ and $t$ were chosen arbitrarily, we conclude

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

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

**Remark 3.2** See [50], [51], and [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 on p. 105]. The part determining the limit is borrowed from [13, Fact 2.2.(v)]; see also [4, Theorem 6.2.2] and [12, 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 $A \cap B \neq \varnothing$. Define the sequence of alternating projections by*

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

*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), $P_B P_A$ is strongly nonexpansive with $\mathrm{Fix}(P_B P_A) = A \cap B$. 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 $\|x_{2n} - c\| \geq \|x_{2n+1} - c\| \geq \|x_{2(n+1)} - c\|$, for all $n \in \mathbb{N}$; consequently, $\|x_{2n} - c\| - \|P_A x_{2n} - c\| \to 0$. Since $P_A$ is strongly nonexpansive, it follows that $x_{2n} - x_{2n+1} = x_{2n} - P_A x_{2n} \to 0$. 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 $A \cap B = \varnothing$), 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 $F = \mathrm{Fix}(T_1) \cap \mathrm{Fix}(T_2) \neq \varnothing$. Then for every $x_0 \in C$, the sequence $(x_n)_{n\in\mathbb{N}}$ generated by

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

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 $\ell_2 = \overline{\mathrm{span}}\,\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \}$, where the $n^{\text{th}}$ 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* $X = \ell_2 = \overline{\operatorname{span}}\big\{\mathbf{e}_1, \mathbf{e}_2, \ldots,\big\}$ *and define* $\mathbf{v}$ *by*

$$\begin{aligned}\mathbf{v}\colon\ &[0, +\infty[\ \to \ell_2\\ &r \mapsto \exp\big(-100r^3\big)\mathbf{e}_1 + \cos\Big(\big(r - \lfloor r\rfloor\big)\pi/2\Big)\mathbf{e}_{\lfloor r\rfloor+2} + \sin\Big(\big(r - \lfloor r\rfloor\big)\pi/2\Big)\mathbf{e}_{\lfloor r\rfloor+3}.\end{aligned} \tag{20}$$

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

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

*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* $\{\mathbf{e}_1\}^{\ominus} = \{x \in \ell_2 : \langle \mathbf{e}_1, x\rangle \leq 0\}$. This is fully equivalent to how we stated his example because $K \cap \operatorname{int}\big(\{\mathbf{e}_1\}^{\ominus}\big) = \varnothing$. However, for our purposes, it is more convenient to work with the hyperplane $H$ instead.

(iii) For future use, we point out now that

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

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 $\|y_n\| = \|y_n - 0\| \geq \|P_H y_n - P_H 0\| = \|P_H y_n - 0\| \geq \|P_K P_H y_n - P_K 0\| = \|y_{n+1}\|$, 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, equation (14)]. For $\xi \in \mathbb{R}$, set $x(\xi) = p(\eta) + \xi \mathbf{e}_1$. For all $\xi$ sufficiently negative, we have $\langle \mathbf{e}_1, x(\xi)\rangle \leq 0$. Then $\langle \mathbf{e}_1, P_H x(\xi)\rangle = 0$ for all $\xi$ sufficiently negative and hence $P_H x(\xi) = P_1 p(\eta)$. By [34, Theorem 3.14], there exists $\widehat{\eta} \in \left]\eta, \lfloor \eta \rfloor + 1\right[$ and $\widehat{\alpha} \in \left]0,1\right[$ such that $P_K P_H x(\xi) = P_K P_1 p(\eta) = \widehat{\alpha} p(\widehat{\eta})$. Note that the last term is *independent* of $\xi$ and that $\beta = \langle \mathbf{e}_1, \widehat{\alpha} p(\widehat{\eta})\rangle = \widehat{\alpha} + h(\eta) > 0$. Thus

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

The last term, $(h(\eta)+\xi)\beta$, can be made arbitrarily negative by choosing $\xi$ sufficiently negative. In particular, for all $\xi$ 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 $\|P_H P_K x\|^2 \leq \langle P_H P_K x, x\rangle$, 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, $\langle P_H P_K x, P_{K^\ominus} x\rangle \geq 0$ and so

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

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

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

Further, $P_H(K) \subset P_H(X^+) \subset X^+$ and hence $-P_H(K) \subset -X^+$. Now (27) yields $-\|P_H(K)\|^2 \geq 0$ 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*

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

*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 $y_{n+1} = P_H P_K y_n = P_H P_K z_n = \big(\frac{1}{2}P_K + \frac{1}{2}P_L\big)z_n = z_{n+1}$. 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 $C = \bigcap_{i=1}^m C_i \neq \varnothing$. 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 $f : X \to \left]-\infty, +\infty\right]$ by $\operatorname{Prox}(f)$:

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

For example, if $C$ is a nonempty closed convex set in $X$, then $\operatorname{Prox}(\iota_C) = P_C$. Rockafellar proved that $\partial f$ 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 $\left]-\infty, +\infty\right]$ that are convex, lower semicontinuous, and proper. Suppose $\lambda_1$ and $\lambda_2$ belong to $\left]0, 1\right[$, $\lambda_1 + \lambda_2 = 1$, and set*

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

*Then $f$ is a proper lower semicontinuous convex function such that*

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

*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*

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

*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

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

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

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

where • in the 1[st] 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 2[nd] 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 3[rd] equality; the 4[th] 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, equation (9.1)] or prove it directly via Fenchel duality) implies the 5[th] equality; • the two remaining equalities follow from the definitions. Since $x^*$ was chosen arbitrarily, we have proved the announced identity for $f^*$, namely

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

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

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

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

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

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

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

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

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

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 $f(x) = \tfrac{1}{2}d^2\big(0, (K-x)\cap(x-L)\big)$, for every $x \in X$.*

*Then $f$ is convex, lower semicontinuous, proper, and its proximal map is* $\operatorname{Prox}(f) = \frac{1}{2}P_K + \frac{1}{2}P_L$. *Moreover, the sequence* $\big((\operatorname{Prox}(f))^n z_0\big)_{n\in\mathbb{N}}$ *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

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

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 $C = C_1 \cap C_2 \cap C_3 \neq \varnothing$ 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, Elfving, and Herman [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, Example 2.14 and Example 2.20] for another example of a string-averaging method.) For convergence results in finite-dimensional spaces, see [20, 21, 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

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

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

$$\text{(46)}\quad (\forall x \in X)(\forall y \in X) \quad \sigma(x, y) = \frac{\big(1 - \|x\|^2\big)\big(1 - \|y\|^2\big)}{\big|1 - \langle x, y\rangle\big|^2} \quad \text{and} \quad \rho(x, y) = \operatorname{arctanh} \sqrt{1 - \sigma(x, y)}.$$

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

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

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

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

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

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

**Fact 9.1 (projectors in the Hilbert ball)** *Suppose $C$ is a $\rho$-closed and $\rho$-convex set in $X$. Then for every $x \in X$, there exists a unique point in $C$, denoted $\widetilde{P}_C x$, such that $\rho(x, \widetilde{P}_C x) = \inf_{y \in C} \rho(x, y)$. The induced map $\widetilde{P}_C : X \to C$ is firmly $\rho$-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 *$\rho$-ball* of radius $s$ centered at $y$ (see also [31, Section 15] for characterizations). A set $S$ in $X$ is *$\rho$-bounded*, if it is contained in some open $\rho$-ball. A sequence $(x_n)_{n \in \mathbb{N}}$ in $X$ is $\rho$-bounded, if its orbit $\{x_n : n \in \mathbb{N}\}$ is.

A map $T : X \to X$ is called *para-strongly $\rho$-nonexpansive* (see [47], where this was strongly nonexpansive for brevity), if $T$ is $\rho$-nonexpansive, $\operatorname{Fix} T \neq \varnothing$, and for every $\rho$-bounded sequence $(x_n)_{n \in \mathbb{N}}$ and every $y \in \operatorname{Fix} T$, 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 *$\rho$-retraction* onto $C$, if $T$ is continuous in $(X, \rho)$ and $\operatorname{Fix} T = C$.

The following basic results on para-strongly $\rho$-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 $\rho$-nonexpansive of the first kind and $F \neq \varnothing$, then $T$ is para-strongly $\rho$-nonexpansive.*

(ii) *If $T$ is para-strongly $\rho$-nonexpansive, then the map*

$$X \to F : x \mapsto \operatorname{weak\,lim}_n T^n x \tag{50}$$

*is a well-defined $\rho$-nonexpansive $\rho$-retraction onto $F$.*

(iii) *If each $T_i$ is para-strongly $\rho$-nonexpansive and $F_1 \cap F_2 \neq \varnothing$, then $T_2 T_1$ is para-strongly $\rho$-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 $\rho$-closed $\rho$-convex sets in $X$ such that $A \cap B \neq \varnothing$. Then for every $x \in X$, the sequence $\big((\widetilde{P}_B \widetilde{P}_A)^n x\big)_{n \in \mathbb{N}}$ converges weakly to some point in $A \cap B$.*

*Proof.* (See also [47].) The projectors $\widetilde{P}_A$ and $\widetilde{P}_B$ are firmly $\rho$-nonexpansive of the first kind (Fact 9.1), hence para-strongly $\rho$-nonexpansive (Fact 9.2(i)). By Fact 9.2(iii), $T = \widetilde{P}_B \widetilde{P}_A$ is para-strongly $\rho$-nonexpansive and $\operatorname{Fix} T = A \cap B$. 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 $\rho$-nonexpansive maps from $X$ to $X$ such that* $\operatorname{Fix}(T_1) \cap \operatorname{Fix}(T_2) \neq \varnothing$. *Suppose $\lambda_1, \lambda_2$ 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 $\rho$-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, $\mathrm{Fix}(T_1)\cap\mathrm{Fix}(T_2)\subset\mathrm{Fix}\,T$. Conversely, fix $y\in\mathrm{Fix}(T_1)\cap\mathrm{Fix}(T_2)$ and pick $x\in\mathrm{Fix}\,T$. By [37, Lemma 3.3(i)], $\rho(x,y)=\rho(Tx,Ty)=\rho(Tx,y)\leq\max\{\rho(T_1x,y),\rho(T_2x,y)\}$. Without loss of generality, assume that $\max\{\rho(T_1x,y),\rho(T_2x,y)\}=\rho(T_1x,y)$. Since $T_1$ is para-strongly $\rho$-nonexpansive, we obtain $\rho(x,y)\leq\rho(T_1x,y)\leq\rho(x,y)$ and hence $T_1x=x$. Since $x=Tx$, this implies $T_2x=x$. We conclude that $\mathrm{Fix}\,T\subset\mathrm{Fix}(T_1)\cap\mathrm{Fix}(T_2)$. Next, pick a $\rho$-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 [37, Lemma 3.3(i)] again, and after repeatedly passing to subsequences and relabelling if necessary, we obtain $\epsilon>0$ and $M\geq 0$ such that

$$(\forall n\in\mathbb{N})\quad \epsilon\leq\rho(x_n,Tx_n),\;\; \max\big\{\rho(Tx_n,y),\rho(T_2x_n,y)\big\}\leq\rho(T_1x_n,y),\;\; \text{and}\;\; \rho(x_n,y)\to M. \tag{52}$$

Hence $0\leftarrow\rho(Tx_n,y)-\rho(x_n,y)\leq\rho(T_1x_n,y)-\rho(x_n,y)\leq 0$ and thus

$$\rho(T_1x_n,y)-\rho(x_n,y)\to 0. \tag{53}$$

Since $T_1$ is para-strongly $\rho$-nonexpansive, it follows that

$$\rho(T_1x_n,x_n)\to 0. \tag{54}$$

Moreover, (51), (52) and (53) imply that

$$\rho(T_1x_n,y)\to M,\;\; \limsup\rho(T_2x_n,y)\leq M,\;\; \text{and}\;\; \rho(Tx_n,y)\to M. \tag{55}$$

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:

$$(\forall n\in\mathbb{N})\quad \begin{cases} u_n\in\mathrm{conv}\{T_1x_n,Tx_n\},\;\; v_n\in\mathrm{conv}\{Tx_n,T_2x_n\},\;\; \text{and}\\ \|u_n-Tx_n\|=\|v_n-Tx_n\|=\min\big\{\|T_1x_n-Tx_n\|,\|T_2x_n-Tx_n\|\big\}.\end{cases} \tag{56}$$

Then, by construction,

$$(\forall n\in\mathbb{N})\quad \tfrac{1}{2}u_n+\tfrac{1}{2}v_n=Tx_n\;\; \text{and}\;\; \|T_1x_n-T_2x_n\|=\frac{\|u_n-v_n\|}{2\min\{\lambda_1,\lambda_2\}}. \tag{57}$$

Using [37, Lemma 3.3(i)] and (52), we obtain $\rho(u_n,y)\leq\max\{\rho(T_1x_n,y),\rho(Tx_n,y)\}=\rho(T_1x_n,y)$ and $\rho(v_n,y)\leq\max\{\rho(Tx_n,y),\rho(T_2x_n,y)\}\leq\rho(T_1x_n,y)$, for every $n\in\mathbb{N}$. Combining these inequalities with (55) and (57), we see that

$$\limsup\rho(u_n,y)\leq M,\;\; \limsup\rho(v_n,y)\leq M,\;\; \text{and}\;\; \lim\rho\big(\tfrac{1}{2}u_n+\tfrac{1}{2}v_n,y\big)=M. \tag{58}$$

Now (58) and [46, Lemma 4] result in $u_n - v_n \to 0$; equivalently (see (57)), $T_1x_n - T_2x_n \to 0$. By (55), the sequences $(T_1x_n)_{n\in\mathbb{N}}$ and $(T_2x_n)_{n\in\mathbb{N}}$ are both $\rho$-bounded; therefore, using [37, Theorem 3.4], we conclude that $\rho(T_1x_n, T_2x_n) \to 0$. The triangle inequality and (54) yield

$$\rho(T_2x_n, x_n) \to 0. \tag{59}$$

Finally, (52), (54), (59), and [37, Lemma 3.3(i)] imply the contradiction $\epsilon \leq \rho(Tx_n, x_n) \leq \max\{\rho(T_1x_n, x_n), \rho(T_2x_n, x_n)\} \to 0$. ■

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

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

*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 *$\rho$-averaged* (respectively, *$\rho$-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 $\rho$-nonexpansive and $\lambda \in [0,1[$; see [46] — if furthermore $\lambda = \frac{1}{2}$, we say that the map is *$\rho$-midpoint averaged*. It is known that if $T$ is $\rho$-averaged of either kind with $\operatorname{Fix} T \neq \varnothing$ 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 [36] (see also [37, Example 10.6 on p. 475]) constructed $\rho$-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 $\frac{1}{2}\widetilde{P}_A \oplus \frac{1}{2}\widetilde{P}_B$ or $\frac{1}{2}\widetilde{P}_A + \frac{1}{2}\widetilde{P}_B$, for some $\rho$-closed $\rho$-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].

## Acknowledgments

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 and H. Brézis, Duality for the sum of convex functions in general Banach spaces, in *Aspects of Mathematics and its Applications*, North-Holland Math. Library, vol. 34, North-Holland, Amsterdam, The Netherlands, pp. 125–133, 1986.

[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 and S. Reich, On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces, *Houston Journal of Mathematics*, vol. 4, pp. 1–9, 1978.

[4] H. H. Bauschke, *Projection Algorithms and Monotone Operators*, PhD 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 *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications (Haifa 2000)*, D. Butnariu, Y. Censor, and S. Reich (editors), Elsevier, Amsterdam, The Netherlands, pp. 11–22, 2001.

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

[7] H. H. Bauschke and J. M. Borwein, Dykstra's alternating projection algorithm for two sets, *Journal of Approximation Theory,* vol. 79, pp. 418–443, 1994.

[8] H. H. Bauschke and J. M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review*, vol. 38, pp. 367–426, 1996.

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

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

[11] H. H. Bauschke and P. L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, *Mathematics of Operations Research*, vol. 26, pp. 248–264, 2001.

[12] H. H. Bauschke, F. Deutsch, H. Hundal, and S.-H. Park, Accelerating the convergence of the method of alternating projections, to appear in *Transactions of the American Mathematical Society*, 2003.

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

[14] Y. Benyamini and 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, *Soviet Mathematics Doklady*, vol. 6, pp. 688–692, 1965.

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

[17] H. Brézis and P.-L. Lions, Produits infinis de résolvantes, *Israel Journal of Mathematics*, vol. 29, pp. 329–345, 1978.

[18] R. E. Bruck and S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston Journal of Mathematics*, vol. 3, pp. 459–470, 1977.

[19] D. Butnariu, Y. Censor, and S. Reich (editors), *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 *Visualization and Optimization Techniques*, Y. Censor and M. Ding (editors), Proceedings of SPIE, vol. 4553, SPIE — The International Society for Optical Engineering, Bellingham, WA, pp. 1-9, 2001.

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

[22] Y. Censor and E. Tom, Convergence of string-averaging projection schemes for inconsistent convex feasibility problems, preprint, 2003.

[23] Y. Censor and 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, pp. 326–333, 1938.

[25] P. L. Combettes, Hilbertian convex feasibility problem: convergence of projection methods, *Applied Mathematics and Optimization*, vol. 35, pp. 311–330, 1997.

[26] G. Crombez, Finding common fixed points of strict paracontractions by averaging strings of sequential iterations, *Journal of Nonlinear and Convex Analysis*, vol. 3, pp. 345–351, 2002.

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

[28] J. M. Dye and S. Reich, On the unrestricted iteration of projections in Hilbert space, *Journal of Mathematical Analysis and Applications*, vol. 156, pp. 101–119, 1991.

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

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

[31] K. Goebel and 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 Journal on Control and Optimization*, vol. 29, pp. 403–419, 1991.

[33] J.-B. Hiriart-Urruty, A general formula on the conjugate of the difference of functions, *Canadian Mathematical Bulletin*, vol. 29, pp. 482–485, 1986.

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

[35] M. D. Kirszbraun, Über die zusammenziehenden und Lipschitzschen Transformationen, *Fundamenta Mathematicae*, vol. 22, pp. 77–108, 1934.

[36] T. Kuzcumow and A. Stachura, Extensions of nonexpansive mappings in the Hilbert ball with the hyperbolic metric II, *Commentationes Mathematicae Universitatis Carolinae*, vol. 29, pp. 403–410, 1988.

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

[38] B. Martinet, Régularisation d'inéquations variationnelles par approximations successives, *Revue Française d'Informatique et de Recherche Opérationnelle,* vol. 4, pp. 154–158, 1970.

[39] Y. I. Merzlyakov, On a relaxation method of solving systems of linear inequalities, *U.S.S.R. Computational Mathematics and Mathematical Physics*, vol. 2, pp. 504–510, 1963.

[40] G. J. Minty, Monotone (nonlinear) operators in Hilbert space, *Duke Mathematical Journal*, vol. 29, pp. 341–346, 1962.

[41] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, *Bulletin de la Société Mathématique de France*, vol. 93, pp. 273–299, 1965.

[42] S. Reich, Product formulas, nonlinear semigroups, and accretive operators, *Journal of Functional Analysis*, vol. 36, pp. 147–168, 1980.

[43] S. Reich, On the asymptotic behavior of nonlinear semigroups and the range of accretive operators, *Journal of Mathematical Analysis and Applications*, vol. 79, pp. 113–126, 1981.

[44] S. Reich, On the asymptotic behavior of nonlinear semigroups and the range of accretive operators II, *Journal of Mathematical Analysis and Applications*, vol. 87, pp. 134–146, 1982.

[45] S. Reich, A limit theorem for projections, *Linear and Multilinear Algebra*, vol. 13, pp. 281-290, 1983.

[46] S. Reich, Averaged mappings in the Hilbert ball, *Journal of Mathematical Analysis and Applications*, vol. 109, pp. 199–206, 1985.

[47] S. Reich, The alternating algorithm of von Neumann in the Hilbert ball, *Dynamic Systems and Applications*, vol. 2, pp. 21–25, 1993.

[48] R. T. Rockafellar, On the maximal monotonicity of subdifferential mappings, *Pacific Journal of Mathematics*, vol. 33, pp. 209–216, 1970.

[49] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM Journal on Control and Optimization*, vol. 14, pp. 877-898, 1976.

[50] 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.)

[51] J. von Neumann, On rings of operators. Reduction theory, *Annals of Mathematics*, vol. 50, pp. 401–485, 1949.

[52] N. Wiener, On the factorization of matrices, *Commentarii Mathematici Helvetici*, vol. 29, pp. 97–111, 1955.

[53] I. Yamada, The hybrid steepest descent method for the variational inequality problem over the intersection of fixed points sets of nonexpansive mappings, in *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications (Haifa 2000)*, D. Butnariu, Y. Censor, and S. Reich (editors), Elsevier, Amsterdam, The Netherlands, pp. 473–504, 2001.

[54] C. Zălinescu, *Convex Analysis in General Vector Spaces,* World Scientific Publishing, River Edge, NJ, 2002.

[55] E. H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory, in *Contributions to Nonlinear Functional Analysis,* E. H. Zarantonello (editor), Academic Press, New York, NY, pp. 237–424, 1971.

Heinz H. Bauschke
Department of Mathematics and Statistics
University of Guelph
Guelph, Ontario N1G 2W1
Canada
hbauschk@uoguelph.ca

Eva Matoušková
Mathematical Institute
Academy of Sciences of the Czech Republic
Žitná 25
CZ-11567 Prague
Czech Republic
matouse@math.cas.cz

Simeon Reich
Department of Mathematics
The Technion — Israel Institute of Technology
32000 Haifa
Israel
sreich@tx.technion.ac.il