Abstract

We consider the problem of finding a best approximation pair, i.e., two points which achieve the minimum distance between two closed convex sets in a Hilbert space. When the sets intersect, the method under consideration, termed ASR for averaged successive reflections, is a special instance of an algorithm due to Lions and Mercier for finding a zero of the sum of two maximal monotone operators. We investigate systematically the asymptotic behavior of ASR in the general case when the sets do not necessarily intersect and show that the method produces best approximation pairs provided they exist. Finitely many sets are handled in a product space, in which case the ASR method is shown to coincide with a special case of Spingarn’s method of partial inverses.

Keywords: Best approximation pair, convex set, firmly nonexpansive map, Hilbert space, hybrid projection-reflection method, method of partial inverses, normal cone, projection, reflection, weak convergence.

1 Introduction

Throughout this paper,

\[ X \text{ is a real Hilbert space with inner product } \langle \cdot , \cdot \rangle \text{ and induced norm } \| \cdot \| , \]

\[ ^{\dagger} \text{Published version: Journal of Approximation Theory 127 (2004) 178-192} \]
\[ ^{\dagger} \text{Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario N1G 2W1, Canada. E-mail: hbauschk@uoguelph.ca.} \]
\[ ^{\ddagger} \text{Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie – Paris 6, 75005 Paris, France. E-mail: plc@math.jussieu.fr.} \]
\[ ^{§} \text{The Pacific Institute for the Mathematical Sciences, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada. E-mail: rluke@cecm.sfu.ca.} \]
and

(2) \( A \) and \( B \) are two nonempty closed convex (possibly non-intersecting) sets in \( X \).

Let \( I \) denote the identity operator on \( X \), and let \( P_A \) and \( P_B \) be the projectors (best approximation operators) onto \( A \) and \( B \), respectively. Given a point \( a \in X \), the standard best approximation problem relative to \( B \) is to [16]

(3) \[
\text{find } b \in B \text{ such that } \|a - b\| = \inf \|a - B\|.
\]

A natural extension of this problem is to find a best approximation pair relative to \((A, B)\), i.e., to

(4) \[
\text{find } (a, b) \in A \times B \text{ such that } \|a - b\| = \inf \|A - B\|.
\]

If \( A = \{a\} \), (4) reduces to (3) and its solution is \( P_B a \). On the other hand, when the problem is consistent, i.e., \( A \cap B \neq \emptyset \), then (4) reduces to the well-known convex feasibility problem for two sets [4, 13] and its solution set is \( \{(x, x) \in X \times X : x \in A \cap B\} \). The formulation (4) captures a wide range of problems in applied mathematics and engineering [11, 23, 26, 29, 34].

The method of alternating projections applied to the sets \( A \) and \( B \) is perhaps the most straightforward algorithm to obtain a best approximation pair. It is described by the algorithm

(5) \[
\text{Take } x_0 \in X \text{ and set } (\forall n \in \mathbb{N}) \ x_n = (P_A P_B)^n x_0.
\]

It was shown in [10, Theorem 2] that if \( A \) or \( B \) is compact, then the sequence \((x_n, P_B x_n)_{n \in \mathbb{N}}\) converges in norm to a best approximation pair. Best approximation pairs may not exist in general; however, if they do, then the sequence generated by (5) solves (4) in the sense that \((x_n, P_B x_n)_{n \in \mathbb{N}}\) converges weakly to some best approximation pair. This happens in particular when one of the sets is bounded [3, 11, 23].

While simple and elegant, the method of alternating projections can suffer from slow convergence, as theoretical [5, 20] and numerical [12] investigations have shown. We analyze an alternative strategy based on reflections rather than projections. Denote the reflectors with respect to \( A \) and \( B \) by \( R_A = 2P_A - I \) and \( R_B = 2P_B - I \) respectively, and consider the successive approximation method

(6) \[
\text{Take } x_0 \in X \text{ and set } (\forall n \in \mathbb{N}) \ x_n = T^n x_0,
\]

where

(7) \[
T = \frac{1}{2}(R_A R_B + I).
\]

The algorithm described by (6) iterates the operator \( T \) which is the average between the successive reflectors \( R_A R_B \) and \( I \). We thus refer to (6)–(7) as the Averaged Successive Reflections (ASR) method. This algorithm does not appear to be well known in approximation theory; consequently, let us now provide some motivation for it.
• If $A \cap B \neq \emptyset$, then the ASR method is a special case of a nonlinear variant of the Douglas-Rachford algorithm [17] proposed by Lions and Mercier in [25] to find a zero of the sum of two maximal monotone operators (in our setting, the normal cone maps of $A$ and $B$).

• Our work in the field of imaging [6, 7] motivated us to study the ASR method. In [7], we used a relaxed version of (6)–(7), which we called the Hybrid Projection-Reflection (HPR) method, to solve the (nonconvex) phase retrieval problem. This algorithm was inspired by our attempt to use reliable convex optimization techniques as a basis to analyze current state-of-the-art techniques in phase retrieval [6]. In fact, more than thirty years of numerical experience with the phase retrieval problem have shown that HPR-type methods converge to an acceptable neighborhood of the solution in fewer iterations than alternating projections.

• If $B$ is the Cartesian product of finitely many halfspaces and $A$ is the diagonal in the corresponding product space, then the ASR method coincides with Spingarn’s method of partial inverses for solving linear inequalities; see Section 4 for further details. Spingarn reports on page 61 in [32] that his algorithm “does better on certain classes of poorly conditioned problems,” although it “is outperformed by cyclic projections with overrelaxation, at least on well-conditioned problems.” An interesting open problem is to obtain some more precise guidelines on when one should prefer the ASR method to cyclic projections and vice versa.

• The following example in the Euclidean plane illustrates a simple setting where the ASR method is superior to the method of alternating projections. Let $A = \{(r, s) \in \mathbb{R}^2 : s \leq 0\}$ and $B = \{(r, s) \in \mathbb{R}^2 : r \leq s\}$. Fix $x_0 = (8, 4)$ as a starting point for the sequence $(x_n)_{n \in \mathbb{N}}$ generated the ASR method (6)–(7). Then $x_1 = (6, -2)$, $x_2 = (2, -4)$, $x_3 = (-1, -3)$, and $x_4 = x_n = (-2, -2)$, for every $n \geq 4$. Thus the ASR method finds the point $(-2, -2) \in A \cap B$ in just four iterations. In contrast, the sequence generated by the method of alternating projections (5) with the same starting point $(8, 4)$ converges to $(0, 0) \in A \cap B$, but not in finitely many steps.

We now recall the known convergence results for the ASR method.

**Fact 1.1** Suppose that $A \cap B \neq \emptyset$ and let $(x_n)_{n \in \mathbb{N}}$ be an arbitrary sequence generated by (6)–(7). Then the following hold.

(i) $(x_n)_{n \in \mathbb{N}}$ converges weakly to some fixed point $x$ of $T$ and $P_B x \in A \cap B$.

(ii) The “shadow” sequence $(P_B x_n)_{n \in \mathbb{N}}$ is bounded and each of its weak cluster points belongs to $A \cap B$.

**Proof.** See [25, Theorem 1] (specialized to the normal cone maps of $A$ and $B$), or the more direct proof of [6, Fact 5.9].

The aim of this paper is to analyze completely the asymptotic behavior of the ASR method (6)–(7), covering in particular the case when $A \cap B = \emptyset$. In addition, we shall briefly explore extensions of our main results to the setting of finitely many sets.
The paper is organized as follows. We provide basic facts on the geometry of two closed convex sets in Section 2. In Section 3, we show that, for any sequence \((x_n)_{n \in \mathbb{N}}\) generated by (6)–(7), either \(\|P_B x_n\| \to +\infty\) and (4) has no solution, or \(((P_A R_B x_n, P_B x_n))_{n \in \mathbb{N}}\) is bounded and its weak cluster points are solutions of (4). Additional results are presented for the case when \(A\) is a linear subspace. This is of relevance in Section 4, where finitely many sets are handled in a product space. We conclude by establishing a connection with Spingarn’s method of partial inverses [31, 32, 33].

Notation. The closure of a set \(C \subset X\) is denoted by \(\overline{C}\) and its interior by \(\text{int} C\); its recession cone is \(\text{rec}(C) = \{x \in X : x + C \subset C\}\) (note that \(\text{rec } \emptyset = X\)) and its normal cone map is given by \(N_C : x \mapsto \{\{u \in X : (\forall c \in C) \langle c - x, u \rangle \leq 0\}\}\), if \(x \in C\); \(\emptyset\), otherwise.

If \(C\) is a convex cone, its polar cone is \(C^\ominus = \{x \in X : (\forall c \in C) \langle c, x \rangle \leq 0\}\) and \(C^\oplus = -C^\ominus\). The range of an operator \(T\) is denoted by \(\text{ran } T\) (with closure \(\overline{\text{ran } T}\)) and its fixed point set by \(\text{Fix } T\). Finally, \(\rightharpoonup\) denotes weak convergence and \(\mathbb{N}\) is the set of nonnegative integers.

2 The geometry of two closed convex sets

Recall (see [21, Theorem 12.1]) that an operator \(\tilde{T}\) from \(X\) to \(X\) is firmly nonexpansive, i.e.,
\[
(\forall x \in X)(\forall y \in X) \quad \|\tilde{T} x - \tilde{T} y\|^2 + \|(I - \tilde{T}) x - (I - \tilde{T}) y\|^2 \leq \|x - y\|^2,
\]
if and only if \(\tilde{R} = 2\tilde{T} - I\) is nonexpansive, i.e.,
\[
(\forall x \in X)(\forall y \in X) \quad \|\tilde{R} x - \tilde{R} y\| \leq \|x - y\|.
\]

**Fact 2.1** 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 \(\|x - P_C x\| = \inf \|x - C\|\). The point \(P_C x\) is characterized by
\[
P_C x \in C \quad \text{and} \quad (\forall c \in C) \quad \langle c - P_C x, x - P_C x \rangle \leq 0.
\]
The operator \(P_C : X \to C : x \mapsto P_C x\) is called the projector onto \(C\); it is firmly nonexpansive and consequently, the reflector \(R_C = 2P_C - I\) is nonexpansive.

**Proof.** See [16, Theorems 4.1 and 5.5], [21, Chapter 12], [22, Propositions 3.5 and 11.2], or [35, Lemma 1.1].\(\Box\)

**Fact 2.2** Suppose that \(C\) is a nonempty closed convex set in \(X\). Then \(\text{ran } (I - P_C) = (\text{rec}(C))^\ominus\).

**Proof.** See [35, Theorem 3.1].\(\Box\)
In order to study the geometry of the given two closed convex sets $A$ and $B$, it is convenient to introduce the following objects, which we use throughout the paper:

\begin{equation}
D = B - A, \quad v = P_D(0), \quad E = A \cap (B - v), \quad \text{and} \quad F = (A + v) \cap B.
\end{equation}

It follows at once from (10) that

\begin{equation}
-v \in N_D(v).
\end{equation}

Note also that if $A \cap B \neq \emptyset$, then $E = F = A \cap B$. However, even when $A \cap B = \emptyset$, the sets $E$ and $F$ may be nonempty and they serve as substitutes for the intersection. Indeed, $\|v\|$ measures the “gap” between the sets $A$ and $B$.

Fact 2.3

(i) $\|v\| = \inf \|A - B\|$, and the infimum is attained if and only if $v \in B - A$.

(ii) $E = \text{Fix}(P_A P_B)$ and $F = \text{Fix}(P_B P_A)$.

(iii) $E + v = F$.

(iv) If $e \in E$ and $f \in F$, then $P_B e = P_F e = e + v$ and $P_A f = P_E f = f - v$.

(v) $E$ and $F$ are nonempty provided one of the following conditions holds:

(a) $A \cap B \neq \emptyset$.

(b) $B - A$ is closed.

(c) $A$ or $B$ is bounded.

(d) $A$ and $B$ are polyhedral sets (intersections of finitely many halfspaces).

(e) $\text{rec}(A) \cap \text{rec}(B)$ is a linear subspace, and $A$ or $B$ is locally compact.

Proof. See [2, Section 5] and [3, Section 2]. \(\Box\)

Proposition 2.4 Suppose that $f \in F$ and $y \in N_D(v)$, and set $e = f - v \in E$. Then the following hold.

(i) $N_D(v) = N_B(f) \cap (-N_A(e))$.

(ii) $P_B(f + y) = f$.

(iii) $P_A(e - y) = e$.

Proof. (i) follows from (11). (ii) and (ii) follow from (i) and (10). \(\Box\)
Proposition 2.5 Suppose that \((a_n)_{n \in \mathbb{N}}\) and \((b_n)_{n \in \mathbb{N}}\) are sequences in \(A\) and \(B\), respectively. Then

\begin{equation}
(13) \quad b_n - a_n \to v \iff \|b_n - a_n\| \to \|v\|.
\end{equation}

Now assume that \(b_n - a_n \to v\). Then the following hold.

(i) \(b_n - P_A b_n \to v\) and \(P_B a_n - a_n \to v\).

(ii) The weak cluster points of \((a_n)_{n \in \mathbb{N}}\) and \((P_A b_n)_{n \in \mathbb{N}}\) (resp. \((b_n)_{n \in \mathbb{N}}\) and \((P_B a_n)_{n \in \mathbb{N}}\)) belong to \(E\) (resp. \(F\)). Consequently, the weak cluster points of the sequences

\[ ((a_n, b_n))_{n \in \mathbb{N}}, ((a_n, P_B a_n))_{n \in \mathbb{N}}, ((P_A b_n, b_n))_{n \in \mathbb{N}} \]

are best approximation pairs relative to \((A, B)\).

(iii) If \(E = \emptyset\) (or, equivalently, \(F = \emptyset\)), then \(\min \{\|a_n\|, \|P_A b_n\|, \|b_n\|, \|P_B a_n\|\} \to +\infty\).

Proof. The implication \(\Rightarrow\) is clear. Conversely, let \((\forall n \in \mathbb{N}) \ d_n = b_n - a_n \in B - A \subset \overline{B - A} = D\). It follows from \((10)\) that \((\forall n \in \mathbb{N}) \ <d_n - v, v> \geq 0\). Hence

\begin{equation}
(14) \quad (\forall n \in \mathbb{N}) \quad \|d_n\|^2 - \|v\|^2 = \|d_n - v\|^2 + 2\langle d_n - v, v \rangle \geq \|d_n - v\|^2,
\end{equation}

which proves \((13)\). Assume for the remainder of the proof that \(b_n - a_n \to v\) or, equivalently, \(\|b_n - a_n\| \to \|v\|\). Since

\begin{equation}
(\forall n \in \mathbb{N}) \quad \|b_n - a_n\| \geq \max \{\|b_n - P_A b_n\|, \|P_B a_n - a_n\|\}
\geq \min \{\|b_n - P_A b_n\|, \|P_B a_n - a_n\|\}
\geq \|v\|,
\end{equation}

we conclude that \((\|b_n - P_A b_n\|)_{n \in \mathbb{N}}\) and \((\|P_B a_n - a_n\|)_{n \in \mathbb{N}}\) both converge to \(\|v\|\). As just proved, this now yields \(b_n - P_A b_n \to v\) and \(P_B a_n - a_n \to v\). Hence (i) holds. Let \(a \in A\) be a weak cluster point of \((a_n)_{n \in \mathbb{N}}\), say \(a_n \to a\). Then \(b_{k_n} \to v + a \in B \cap (v + A) = F\). Hence \(a \in A \cap (B - v) = E\). The remaining three sequences are treated similarly and thus (ii) is verified. Finally, (iii) is a direct consequence of (ii). \(\square\)

Remark 2.6 Sequences conforming to the assumptions described in Proposition 2.5 can be generated by \((5)\), upon rewriting it as

\begin{equation}
(15) \quad \text{Take } b_{-1} \in B \text{ and set } (\forall n \in \mathbb{N}) \ a_n = P_A b_{n-1} \text{ and } b_n = P_B a_n.
\end{equation}

Indeed, [3, Theorem 4.8] implies that \(b_n - a_n \to v\) (see also [11]). This happens also for the iterates generated by Dykstra’s algorithm [3, Theorem 3.8]. In Theorem 3.13 below, we shall see that the ASR method also gives rise to sequences with this behavior.

Corollary 2.7 \(v \in (P_B - I)(A) \cap (I - P_A)(B) \subset (\text{rec } B)^\circ \cap (\text{rec } A)^\circ\).
Proof. In view of Proposition 2.5(i) and Remark 2.6,
\[
(16) \quad v \in (P_B - I)(A) \cap (I - P_A)(B) \subset \text{ran} (P_B - I) \cap \text{ran} (I - P_A).
\]
Now apply Fact 2.2. □

Remark 2.8 Corollary 2.7 can be refined in certain cases.

(i) First assume that \( A = a + K \) and \( B = b + L \), where \( K \) and \( L \) are closed convex cones. Then 
\[ \text{rec}(A) = K \] and \( \text{rec}(B) = L \). Hence, by Corollary 2.7, \( v \in L^\oplus \cap K^\ominus \). In fact, [2, Ex. 2.2] shows that 
\[ v = P_{L^\oplus \cap K^\ominus} (b - a). \]

(ii) Now assume that \( A \) is a closed affine subspace, say \( A = a + K \), where \( K \) is a closed linear subspace. Then \( K = A - A \) and hence 
\[ v \in (A - A)^\perp. \]

3 The Averaged Successive Reflections (ASR) method

Let us start with a key observation concerning the operator \( T = (R_AR_B + I)/2 \).

Proposition 3.1 \( T \) is firmly nonexpansive and defined on \( X \).

Proof. By Fact 2.1, the projectors \( P_A \) and \( P_B \) are firmly nonexpansive. As pointed out in the beginning of Section 2, the corresponding reflectors \( R_A \) and \( R_B \) are nonexpansive. It follows that \( R_AR_B \) is nonexpansive as well and, hence, that \( T \) is firmly nonexpansive. □

Several fundamental results on firmly nonexpansive maps have been discovered over the past four decades. Specializing these to \( T \), we obtain the following.

Fact 3.2 Let \( x_0 \in X \). Then:

(i) \( (T^n x_0 - T^{n+1} x_0)_{n \in \mathbb{N}} \) converges in norm to the unique element of minimum norm in \( \text{ran} (I - T) \);

(ii) \( \text{Fix} T \neq \emptyset \iff (T^n x_0)_{n \in \mathbb{N}} \) converges weakly to some point in \( \text{Fix} T \);

(iii) \( \text{Fix} T = \emptyset \iff \|T^n x_0\| \to +\infty \).

Proof. (See also [9].) (i): [1, Corollary 2.3] and [28, Corollary 2] (ii): [27, Theorem 3]. (iii): [1, Corollary 2.2]. □

The following identities will be useful later.
Proposition 3.3 Let \( x \in X \). Then:

(i) \( x - Tx = P_Bx - P_AR_Bx \):

(ii) \( \|x - Tx\|^2 = \|x - P_Bx\|^2 + \langle x - P_AR_Bx, R_Bx - P_AR_Bx \rangle \).

Proof. Indeed,

\[
x - Tx = x - \frac{1}{2}(R_AR_Bx + x) = \frac{1}{2}(x - 2P_AR_Bx + R_Bx) = \frac{1}{2}(x - 2P_AR_Bx + 2Bx - x)
\]

\[
= P_Bx - P_AR_Bx.
\]

Hence (i) holds, and we obtain further

\[
\|x - Tx\|^2 = \|P_Bx - P_AR_Bx\|^2
\]

\[
= \|P_Bx - x\|^2 + \|x - P_AR_Bx\|^2 + 2\langle x - P_AR_Bx, P_Bx - x \rangle
\]

\[
= \|P_Bx - x\|^2 + \|x - P_AR_Bx\|^2 + \langle x - P_AR_Bx, R_Bx - x \rangle
\]

\[
= \|P_Bx - x\|^2 + \|x - P_AR_Bx\|^2 + \langle x - P_AR_Bx, (R_Bx - P_AR_Bx) - (x - P_AR_Bx) \rangle
\]

\[
= \|P_Bx - x\|^2 + \langle x - P_AR_Bx, R_Bx - P_AR_Bx \rangle,
\]

as announced in (ii). \( \Box \)

Theorem 3.4 The unique element of minimum norm in \( \text{ran}(I - T) \) is \( v \).

Proof. It follows from Fact 3.2(i) that \( \text{ran}(I - T) \) possesses a unique element of minimum norm, say \( w \). We shall show that \( w = v \). On the one hand, by Proposition 3.3(i), we have \( \text{ran}(I - T) \subseteq B - A \) and hence \( w \in B - A = D \). On the other hand, it follows from Proposition 3.3(ii) and (10) that, for every \( a \in A \),

\[
\|w\|^2 \leq \|a - Ta\|^2 = \|P_Ba - a\|^2 + \langle a - P_AR_Ba, R_Ba - P_AR_Ba \rangle \leq \|P_Ba - a\|^2 = \inf \|B - a\|^2.
\]

Hence \( \|w\| \leq \inf \|B - A\| \) and, therefore, \( w = P_D0 = v \). \( \Box \)

Theorem 3.5 The set \( \text{Fix}(T + v) \) is closed and convex. Moreover,

\[
(17) \quad F + N_D(v) \subseteq \text{Fix}(T + v) \subseteq v + F + N_D(v).
\]

Proof. Since \( T \) is firmly nonexpansive, so is \( T + v \). Hence \( \text{Fix}(T + v) \) is closed and convex (see, for instance, [21, Lemma 3.4] or [22, Proposition 5.3]). Now pick \( f \in F, y \in N_D(v) \), and set \( x = f + y \). By Proposition 2.4(ii), we have \( P_Bx = f \). Hence \( R_Bx = 2P_Bx - x = 2f - (f + y) = f - y \). Now, let \( e = f - v \). It follows from (12) that \( y - v \in N_D(v) \). Therefore, using Proposition 2.4(iii), we obtain \( P_AR_Bx = P_A(f - y) = P_A(e - (y - v)) = e = f - v \). Hence \( P_Bx - P_AR_Bx = f - (f - v) = v \).
By Proposition 3.3(i), \( x - Tx = P_B x - P_A R_B x = v \) and, in turn, \( x = Tx + v \), i.e., \( x \in \text{Fix}(T + v) \). Thus,

\[
(18) \quad F + N_D(v) \subset \text{Fix}(T + v).
\]

To establish the remaining inclusion, pick \( x \in \text{Fix}(T + v) \). Then \( x - Tx = v \) or, equivalently (see Proposition 3.3), \( P_B x - P_A R_B x = v \). Let \( f = P_B x = v + P_A R_B x \) and \( y = x - v - f \). Then \( f \in B \cap (A + v) = F \) and \( x = v + f + y \). It now suffices to show that \( y \in N_D(v) \). To see this, pick \( a \in A \) and \( b \in B \). On the one hand, since \( f = P_B x \), Fact 2.1 results in \( \langle b - f, x - f \rangle \leq 0 \). Using the definition of \( y \), we write the last inequality equivalently as

\[
(19) \quad \langle b - f, y + v \rangle \leq 0.
\]

On the other hand, \( P_A(2f - x) = P_A(2P_B x - x) = P_A R_B x = f - v \). Again using Fact 2.1, we deduce \( \langle a - f + v, -y \rangle = \langle a - (f - v), (2f - x) - (f - v) \rangle \leq 0 \). Hence

\[
(20) \quad \langle f - a - v, y \rangle \leq 0.
\]

Adding (19) and (20), we obtain \( \langle b - a - v, y \rangle + \langle b - f, v \rangle \leq 0 \). This inequality, (12), Proposition 2.4(ii), and Fact 2.1 now yield \( \langle b - a - v, y \rangle \leq \langle b - f, -v \rangle = \langle b - f, (f - v) \rangle \leq 0 \). We conclude that \( y \in N_D(v) \).

Remark 3.6 A little care with (17) shows that \( \text{rec}(F) + N_D(v) \subset \text{rec}(\text{Fix}(T + v)) \). In particular, if \( F \neq \emptyset \), then \( -v \in \text{rec}(\text{Fix}(T + v)) \) (use (12)).

The next two examples illustrate that the bracketing given for \( \text{Fix}(T + v) \) in Theorem 3.5 is tight.

Example 3.7 Let \( X = \mathbb{R} \), \( A = \{0\} \), and \( B = [1, +\infty[ \). Then \( D = B \), \( v = 1 \), \( F = \{1\} \), and \( \text{Fix}(T + v) = F + N_D(v) \).

Example 3.8 Let \( X = \mathbb{R} \), \( A = [1, +\infty[ \), and \( B = \{0\} \). Then \( D = ]-\infty, -1[ \), \( v = -1 \), \( F = B \), and \( \text{Fix}(T + v) = v + F + N_D(v) \).

The following result, which improves upon [6, Fact A1], gives a complete description of \( \text{Fix} T \) in the consistent case.

Corollary 3.9 Suppose that \( A \cap B \neq \emptyset \). Then \( \text{Fix} T = (A \cap B) + N_D(0) \) and \( P_B(\text{Fix} T) = A \cap B \).

Proof. Since \( A \cap B \neq \emptyset \), we have \( v = 0 \) and \( F = A \cap B \). The formula for \( \text{Fix} T \) (resp. \( P_B(\text{Fix} T) \)) follows from Theorem 3.5 (resp. Proposition 2.4(ii)).

Remark 3.10 We show that if the sets \( A \) and \( B \) do not “overlap sufficiently”, then \( \text{Fix} T \) may be strictly larger than \( A \cap B \). Indeed, let \( X = \mathbb{R} \), \( A = \{0\} \), and \( B = [0, +\infty[ \). Then \( D = B \), \( v = 0 \), \( F = \{0\} = A \cap B \), yet \( \text{Fix} T = ]-\infty, 0[ \). This simple example shows that iterating \( T \) alone may not yield a point in \( A \cap B \). Hence it is important to monitor the “shadow sequence” \( (P_B T^n x_0)_{n \in \mathbb{N}} \); see Fact 1.1 and Theorem 3.13 below.
Remark 3.11 If $0 \in \text{int}(B-A)$ (a fortiori if the Slater-type condition $(A \cap \text{int}(B)) \cup (B \cap \text{int}(A)) \neq \emptyset$ holds), then $N_D(0) = \{0\}$ and consequently (Corollary 3.9) $\text{Fix} T = A \cap B$.

Lemma 3.12 Suppose that $F \neq \emptyset$, let $y_0 \in \text{Fix}(T + v)$ and set $y_n = T^n y_0$, for all $n \in \mathbb{N}$. Then $(y_n)_{n \in \mathbb{N}} = (y_0 - nv)_{n \in \mathbb{N}}$ lies in $\text{Fix}(T + v)$. Moreover,

$$(21) \quad (\forall n \in \mathbb{N}) \quad \|x_{n+1} - y_0 + (n+1)v\|^2 + \|x_n - x_{n+1} - v\|^2 \leq \|x_n - y_0 + nv\|^2.$$ 

Proof. The proof proceeds by induction on $n$. Clearly, $y_0 - 0v = y_0 \in \text{Fix}(T + v)$. Now assume that $y_n = y_0 - nv \in \text{Fix}(T + v)$, for some $n \in \mathbb{N}$. Then $y_{n+1} - nv = y_n = (T + v)(y_n) = Ty_n + v = y_{n+1} + v$ and hence $y_{n+1} = y_0 - (n+1)v$. Moreover, (17) is precisely what is needed to show that $y_{n+1} \in \text{Fix}(T + v)$. Hence the claims regarding $(y_n)_{n \in \mathbb{N}}$ are proven. Next, (21) follows from the firm nonexpansiveness of $T$ (Proposition 3.1) applied to $x_n$ and $y_n = y_0 - nv$. \(\square\)

Theorem 3.13 (Averaged Successive Reflections (ASR) method) Let $x_0 \in X$ and set $x_n = T^n x_0$, for all $n \in \mathbb{N}$. Then the following hold.

(i) $x_n - x_{n+1} = P_B x_n - P_A R_B x_n \to v$ and $P_B x_n - P_A P_B x_n \to v$.

(ii) If $A \cap B \neq \emptyset$, then $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point in $\text{Fix}(T) = (A \cap B) + N_D(0)$; otherwise, $\|x_n\| \to +\infty$.

(iii) Exactly one of the following two alternatives holds.

(a) $F = \emptyset$, $\|P_B x_n\| \to +\infty$, and $\|P_A P_B x_n\| \to +\infty$.

(b) $F \neq \emptyset$, the sequences $(P_B x_n)_{n \in \mathbb{N}}$ and $(P_A P_B x_n)_{n \in \mathbb{N}}$ are bounded, and their weak cluster points belong to $F$ and $E$, respectively; in fact, the weak cluster points of

$$(22) \quad ((P_A R_B x_n, P_B x_n))_{n \in \mathbb{N}} \quad \text{and} \quad ((P_A P_B x_n, P_B x_n))_{n \in \mathbb{N}}$$

are best approximation pairs relative to $(A, B)$.

Proof. (i): On the one hand, Proposition 3.3(i) yields

$$(23) \quad (\forall n \in \mathbb{N}) \quad x_n - x_{n+1} = x_n - Tx_n = P_B x_n - P_A R_B x_n.$$ 

On the other hand, Fact 3.2(i) and Theorem 3.4 imply

$$(24) \quad x_n - x_{n+1} = T^n x_0 - T^{n+1} x_0 \to v.$$ 

Altogether, we obtain the first claim and, by Proposition 2.5(i), $P_B x_n - P_A P_B x_n \to v$. (ii): This follows immediately from Fact 3.2(ii)&(iii) and Corollary 3.9. (iii): If $F = \emptyset$, then (i) and Proposition 2.5(iii) yield $\|P_B x_n\| \to +\infty$ and $\|P_A P_B x_n\| \to +\infty$. Now assume that $F \neq \emptyset$. We claim that $(P_B x_n)_{n \in \mathbb{N}}$ is bounded. Indeed, fix $f \in F \subset \text{Fix}(T + v)$ (see Theorem 3.5). Repeated
Remark 3.16
Pick this and further properties.

Theorem 3.17 (when $A$ is an affine subspace) Suppose that $A$ is a closed affine subspace and $x_0 \in X$. Let $x_n = T^n x_0$, for all $n \in \mathbb{N}$. Then

\begin{equation}
0 \leq \|x_n - (f - nv)\| \leq \|x_0 - f\|, \quad \text{for all } n \in \mathbb{N}.
\end{equation}

Also, since $P_B(f - nv) = f$ (Proposition 2.4(ii)) and $P_B$ is nonexpansive (Fact 2.1), we have

\[
(\forall n \in \mathbb{N}) \quad \|P_B x_n - f\| = \|P_B x_n - P_B(f - nv)\| \leq \|x_n - (f - nv)\| \leq \|x_0 - f\|.
\]

Hence $(P_B x_n)_{n \in \mathbb{N}}$ is bounded. The remaining statements regarding the weak cluster points now follow from (i) and Proposition 2.5(ii).

Remark 3.14 The conclusions of Theorem 3.13 can be strengthened provided $A$ or $B$ has additional properties.

(i) Best approximation pairs exist and can be found as described in Theorem 3.13(iii)(b) whenever (at least) one of the conditions listed in Fact 2.3(v) is satisfied.

(ii) Suppose that best approximation pairs relative to $(A,B)$ exist, i.e., $F \neq \emptyset$. If $P_B$ is weakly continuous (as is the case when $X$ is finite-dimensional or $B$ is a closed affine subspace), then $((P_BR x_n, P_B x_n))_{n \in \mathbb{N}}$ and $((P_A P_B x_n, P_B x_n))_{n \in \mathbb{N}}$ both converge weakly to such a pair.

We shall discuss the important case when $A$ is an affine or linear subspace in Theorem 3.17 and Proposition 3.19 below.

Remark 3.15 If $x_0 \in X$ and $y_0 \in \text{Fix}(T + v)$, then (21) implies that $(\|T^n x_0 + nv - y_0\|)_{n \in \mathbb{N}}$ is decreasing. Consequently, $(T^n x_0 + nv)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\text{Fix}(T + v)$. In certain settings, Fejér monotonicity sheds further light on the behavior of the sequence $(T^n x_0 + nv)_{n \in \mathbb{N}}$.

For instance, if int $\text{Fix}(T + v) \neq \emptyset$, then $(T^n x_0 + nv)_{n \in \mathbb{N}}$ must converge in norm. See [4, 14] for this and further properties.

Remark 3.16 Pick $x_0 \in X$ and set $x_n = T^n x_0$, for every $n \in \mathbb{N}$.

(i) Theorem 3.13(i) states that $P_B x_n - P_A P_B x_n \to v$. Hence, using Fact 2.3(i), $(\delta_n)_{n \in \mathbb{N}} = (\|P_B x_n - P_A P_B x_n\|^2)_{n \in \mathbb{N}}$ converges to $\|v\|^2 = \inf \|A - B\|^2$, the (squared) gap between $A$ and $B$. In [7, Section 4], a normalized version of $\delta_n$ was employed as a stopping criterion and error measure in an application of the ASR method to image processing.

(ii) By [28, Corollary 2], $x_n/n \to -v$. Hence, one can monitor the value of $\|x_n/n\|$ during the execution of the ASR method as an approximation of the gap $\|v\|$.

Theorem 3.17 (when $A$ is an affine subspace) Suppose that $A$ is a closed affine subspace and $x_0 \in X$. Let $x_n = T^n x_0$, for all $n \in \mathbb{N}$. Then

\begin{equation}
P_B x_n - P_A x_n \to v.
\end{equation}

If $F \neq \emptyset$, then $(P_A x_n)_{n \in \mathbb{N}}$ is bounded and its weak cluster points belong to $E$. If furthermore $A \cap B \neq \emptyset$, then $(x_n)_{n \in \mathbb{N}}$ converges weakly to some point $x \in (A \cap B) + N_D(0)$. Moreover, $(P_A x_n)_{n \in \mathbb{N}}$ and $(P_B x_n)_{n \in \mathbb{N}}$ converge weakly to $P_A x \in A \cap B$.  

11
Proof. Since $A$ is an affine subspace, $P_A$ is an affine operator. It follows that $P_AR_B = P_A(P_B + P_B - I) = P_AP_B + P_AP_B - P_A = 2P_AP_B - P_A$ and hence $P_B - P_AR_B = P_B + P_A - 2P_AP_B = 2(P_B - P_AP_B) + P_A - P_B$. In turn, this implies

$$P_B - P_A = 2(P_B - P_AP_B) - (P_B - P_AR_B).$$  \tag{26}

Now apply (26) to $(x_n)_{n \in \mathbb{N}}$, invoke Theorem 3.13(i), and deduce that $P_Bx_n - P_Ax_n \to v$. If $F \neq \emptyset$, then $(P_Bx_n)_{n \in \mathbb{N}}$ is bounded (Theorem 3.13(iii)(b)). Consequently, (25) implies that $(P_Ax_n)_{n \in \mathbb{N}}$ is bounded and that every weak cluster point of $(P_Ax_n)_{n \in \mathbb{N}}$ belongs to $E$ (Proposition 2.5(ii)). Now assume that $A \cap B \neq \emptyset$, whence $v = 0$ and $E = F = A \cap B$. It follows from Theorem 3.13(ii) that $x_n \to x \in (A \cap B) + N_D(0)$. Since $P_A$ is weakly continuous, we have $P_Ax_n \to P_Ax$. By (25) and the weak closedness of $B$, we conclude that $P_Bx_n \to P_Ax \in A \cap B$. \qed

Remark 3.18 The convergence statement (25) need not hold if $A$ is not an affine subspace: indeed, if $x_0 = 0$ in Example 3.8, then $P_Bx_n - P_Ax_n = -\max\{1,n\} \to -\infty$.

When $A$ is a linear subspace, an additional property complements the results of Theorem 3.17.

Proposition 3.19 (when $A$ is a linear subspace) Suppose that $A$ is a closed linear subspace. Then $P_A(\text{Fix}(T + v)) = E$. If $A \cap B \neq \emptyset$, then $P_A(\text{Fix}(T)) = A \cap B$.

Proof. In view of Theorem 3.5 and Fact 2.3(iii), we may assume that $E \neq \emptyset$. Pick $e \in E$. Adding $A^\perp$ to (17) yields

$$F + N_D(v) + A^\perp \subset \text{Fix}(T + v) + A^\perp \subset v + F + N_D(v) + A^\perp. \tag{27}$$

On the other hand, $v \in A^\perp$ (Remark 2.8(ii)) and $N_D(v) \subset -N_A(e) = A^\perp$ (Proposition 2.4(i)). Hence (27) implies $F + A^\perp = \text{Fix}(T + v)$. Together with Fact 2.3(iv), this yields $P_A(\text{Fix}(T + v)) = E$. Now suppose $A \cap B \neq \emptyset$. By (11), $v = 0$ and $E = A \cap B$. Therefore, $P_A(\text{Fix}(T + v)) = E$ becomes $P_A(\text{Fix}(T)) = A \cap B$. \qed

We conclude this section with another special case.

Remark 3.20 Suppose that $A$ is an obtuse cone, i.e., $A^\oplus \subset A$. Pick $x_0 \in A$ and set $x_n = T^n x_0$, for all $n \in \mathbb{N}$. Since $\text{ran} R_A = A$ [8], the entire sequence $(x_n)_{n \in \mathbb{N}}$ lies in $A$.

4 Finitely many sets

In this final section, we show how one can adapt the two-set results of Section 3 to problems with finitely many sets. We assume that

$$C_1, \ldots, C_J \text{ are finitely many nonempty closed convex sets in } X. \tag{28}$$
The following product space technique was first introduced in [30]. Pick \( (\lambda_j)_{1 \leq j \leq J} \) in \([0, 1]\) such that \( \sum_{j=1}^{J} \lambda_j = 1 \) and denote by \( X \) the Hilbert space obtained by equipping the Cartesian product \( X^J \) with the inner product \((x_j)_{1 \leq j \leq J}, (y_j)_{1 \leq j \leq J} \mapsto \sum_{j=1}^{J} \lambda_j \langle x_j, y_j \rangle \). Let
\[
A = \{ (x, \ldots, x) \in X : x \in X \} \quad \text{and} \quad B = C_1 \times \cdots \times C_J.
\]
Then the set \( \bigcap_{j=1}^{J} C_j \) in \( X \) corresponds to the set \( A \cap B \) in \( X \). Moreover, the projections of \( x = (x_j)_{1 \leq j \leq J} \in X \) onto \( A \) and \( B \) are given by
\[
P_A x = (\sum_{j=1}^{J} \lambda_j x_j, \ldots, \sum_{j=1}^{J} \lambda_j x_j) \quad \text{and} \quad P_B x = (P_{C_1} x_1, \ldots, P_{C_J} x_J),
\]
respectively. By analogy with (11), we now set
\[
D = \overline{B} - \overline{A}, \quad v = P_D(0), \quad E = A \cap (B - v), \quad \text{and} \quad F = (A + v) \cap B.
\]
Then a point \( (e, \ldots, e) \in X \) belongs to \( E \) if and only if \( e \) minimizes the proximity function
\[
x \mapsto \sum_{j=1}^{J} \lambda_j \| x - P_{C_j} x \|^2
\]
or, equivalently, if \( e \in \text{Fix} \sum_{j=1}^{J} \lambda_j P_{C_j} \) (see [3, 11, 15] for details). Further, let
\[
T = \frac{1}{2}(R_A R_B + I),
\]
fix \( x_0 \in A \), and set \( x_n = T^n x_0 \), for all \( n \in \mathbb{N} \). Then we obtain the ASR method in \( X \) for the two sets \( A \) and \( B \) and, as seen in Remark 3.10, the pertinent sequence to monitor is the “shadow sequence” \( (P_B x_n)_{n \in \mathbb{N}} \). The results of Section 3 can be applied to this product space setting which, in turn, yield new convergence results for algorithms operating in the original space \( X \) via (30). Rather than detailing these counterparts, we shall bring to light a particularly interesting connection with Spingarn’s method of partial inverses [31] (see also [18, 19, 24]).

**Remark 4.1 (Spingarn’s method of partial inverses)** Since \( A \) is a closed linear subspace, Theorem 3.17 is applicable and one can thus monitor the sequence \( (P_B x_n)_{n \in \mathbb{N}} \) or the sequence \( (P_A x_n)_{n \in \mathbb{N}} \). The latter corresponds precisely to Spingarn’s *method of partial inverses* for finding a zero of \( \sum_{j=1}^{J} \lambda_j N_{C_j} = \sum_{j=1}^{J} N_{C_j} \), i.e., for finding a point in \( \bigcap_{j=1}^{J} C_j \); see [31, Section 6]. It is noteworthy that the main convergence result of Spingarn [31, Corollary 5.1] in this setting can also be deduced from Theorem 3.13 and Proposition 3.19.

Spingarn analyzed further the case when \( X \) is a Euclidean space and each set \( C_j \) is a halfspace in [32] and [33]. Specifically, he proved that \( F \neq \emptyset \) (this can also be deduced from Fact 2.3(v)(d)), that \( (P_A x_n)_{n \in \mathbb{N}} \) converges linearly to some point in \( E \) [33, Theorems 1 and 2], and that convergence occurs in finitely many steps provided that \( \text{int} \bigcap_{j=1}^{J} C_j \neq \emptyset \) [32, Theorem 2].

**Acknowledgments**

H. H. Bauschke’s work was supported in part by the Natural Sciences and Engineering Research Council of Canada. D. R. Luke’s work was supported in part by a Postdoctoral Fellowship with the Pacific Institute for the Mathematical Sciences, and in part by the Institute for Numerical and Applied Mathematics at the University of Göttingen.
References

[1] J. B. Baillon, R. E. Bruck, and S. Reich, On the asymptotic behavior of nonexpansive mappings and


in Hilbert space, in Recent Developments in Optimization Theory and Nonlinear Analysis, pp. 1–38,


448–450, 1959.


[14] P. L. Combettes, Quasi-Fejérian analysis of some optimization algorithms, in Inherently Parallel Algo-
rithms for Feasibility and Optimization (D. Butnariu, Y. Censor, and S. Reich, Eds.), pp. 115–152. New


[18] J. Eckstein, Splitting Methods for Monotone Operators with Applications to Parallel Optimization, PhD
thesis, Department of Civil Engineering, Massachusetts Institute of Technology, 1989.


