A WEAK-TO-STRONG CONVERGENCE PRINCIPLE FOR FEJÉR-MONOTONE METHODS IN HILBERT SPACES

Heinz H. Bauschke† and Patrick L. Combettes‡

†Department of Mathematics and Statistics, Okanagan University College, Kelowna, BC V1V 1V7, CANADA. bauschke@cecm.sfu.ca

‡Department of Electrical Engineering, City College and Graduate School, City University of New York, New York, NY 10031, USA. plc@ee-mail.engr.ccny.cuny.edu

April 7, 1999

Abstract

We consider a wide class of iterative methods arising in numerical mathematics and optimization which are known to converge only weakly. Exploiting an idea originally proposed by Haugazeau, we present a simple modification of these methods which makes them strongly convergent without additional assumptions. Several applications are discussed.


Key words. Convex feasibility, Fejér-monotonicity, firmly nonexpansive mapping, fixed point, Haugazeau, maximal monotone operator, projection, proximal point algorithm, resolvent, subgradient algorithm.

1 Introduction

Let \( \mathcal{H} \) be a real Hilbert space with scalar product \( \langle \cdot, \cdot \rangle \), norm \( \| \cdot \| \), and distance \( d \). In 1965, Bregman [5] proposed a simple iterative method for finding a common point of \( m \) intersecting closed convex sets \( (S_i)_{1 \leq i \leq m} \) in \( \mathcal{H} \). He showed that, given an arbitrary starting point \( x_0 \in \mathcal{H} \), the sequence \( (x_n)_{n \geq 0} \) generated by the periodic projection algorithm

\[
(\forall n \in \mathbb{N}) \quad x_{n+1} = P_{n \mod m+1}x_n,
\]

where \( P_i \) denotes the projector onto \( S_i \) and where the mod \( m \) function takes values in \( \{0, \ldots, m-1\} \), converges weakly to a point in \( S = \bigcap_{i=1}^m S_i \). In [17] (see also [1], [2], [3], and [11]), certain regularity conditions on the sets were described that guaranteed strong convergence of the iterations. To this
day, however, it remains an open question whether the convergence of (1.1) can be strong without such conditions.

In his unpublished 1968 dissertation \[19\], Haugazeau proposed independently a strongly convergent variant of (1.1) requiring essentially the same kind of computations. To describe his method let us define for a given ordered triplet \((x, y, z) \in \mathcal{H}^3\)

\[
H(x, y) = \{u \in \mathcal{H} \mid \langle u - y, x - y \rangle \leq 0\}
\]

(1.2)

and let us denote by \(Q(x, y, z)\) the projection of \(x\) onto \(H(x, y) \cap H(y, z)\). Thus, \(H(x, x) = \mathcal{H}\) and, if \(x \neq y\), \(H(x, y)\) is a closed affine half-space onto which \(y\) is the projection of \(x\). Haugazeau showed that given an arbitrary starting point \(x_0 \in \mathcal{H}\), the sequence \((x_n)_{n \geq 0}\) generated by the algorithm

\[
(\forall n \in \mathbb{N}) \quad x_{n+1} = Q \left( x_0, x_n, P_{(n \mod m) + 1}x_n \right),
\]

(1.3)

converges strongly to the projection of \(x_0\) onto \(S\).

Algorithm (1.1) is Fejér-monotone with respect to the solution set \(S\) in the sense that every orbit \((x_n)_{n \geq 0}\) it generates satisfies

\[
(\forall x, S) (\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\| \leq \|x_n - x\|.
\]

(1.4)

Under this monotonicity condition, \((x_n)_{n \geq 0}\) converges weakly to a point in \(S\) if and only if all its weak cluster points lie in \(S\). This basic fact was used in [2] and [12] to unify and harmonize weak convergence results in numerous areas of numerical mathematics and optimization, including nonlinear fixed point theory, approximation theory, equilibrium theory for sums of monotone set-valued operators, variational inequalities, convex feasibility, and nonsmooth minimization.

A question that naturally arises in connection with Fejér-monotone algorithms is whether weak convergence can be improved to strong convergence without further assumptions. The answer is negative. In fact, [15] and [18] provide two counterexamples of Fejér-monotone methods for which weak convergence holds but strong convergence fails. Examples of restrictions on the constituents of the problem that yield strong convergence in specific applications can be found in [1], [2], [3], [7], [10], [11], [17], [20], [23] and [25]. Typically, these restrictions involve linearity, compactness, or Slater assumptions and they are therefore quite stringent.

The purpose of this paper is to present a generalization of Haugazeau’s method (1.3) for the general problem of finding a point in a possibly empty closed convex set \(S\) in \(\mathcal{H}\) and to analyze its convergence properties. Our main result is a weak-to-strong convergence principle which essentially states that a simple Haugazeau-like transformation of a weakly convergent Fejér-monotone method yields a strongly convergent method without any additional restrictions.

A general model for Fejér-monotone methods is proposed in Section 2 and an abstract Haugazeau method is introduced and analyzed in Section 3. The weak-to-strong convergence principle is derived in Section 4 and applied to various problems in Sections 5 and 6. Throughout, \(\text{Id}\) denotes the identity operator on \(\mathcal{H}\) and \(\text{Fix}T\) the set of fixed points of an operator \(T\). \(\text{PC}\) denotes the projector onto a nonempty closed and convex set \(C\) and \(N_C\) its normal cone. \(\partial f\) denotes the
subdifferential of a function \( f : \mathcal{H} \to \mathbb{R} \) and \( \text{lev}_{\leq \mu} f = \{ x \in \mathcal{H} \mid f(x) \leq \mu \} \) its lower level set at height \( \mu \in \mathbb{R} \). The expressions \( x_n \xrightarrow{n} x \) and \( x_n \xrightarrow{\text{STR}} x \) denote respectively the weak and strong convergence to \( x \) of a sequence \( (x_n)_{n \geq 0} \) and \( \{\text{W}(x_n)\}_{n \geq 0} \) its set of weak cluster points.

2 Fejér-monotone methods

Let us first recall an important result on the weak convergence of Fejér-monotone sequences.

**Proposition 2.1** [8, Lem. 6] Let \( F \) be a nonempty closed and convex subset of \( \mathcal{H} \). Suppose that \( (x_n)_{n \geq 0} \subset \mathcal{H} \) is Fejér-monotone with respect to \( F \). Then \( x_n \xrightarrow{n} x \in F \iff \{\text{W}(x_n)\}_{n \geq 0} \subset F \).

**Proof.** We provide a short proof for completeness (see also [2], [12]). Fix \( z \in F \). By monotonicity, \( (\|x_n - z\|^2)_{n \geq 0} \) converges and so does \( (\|x_n - 2(x_n - z)\|)_{n \geq 0} \). Now take \( z_1 \) and \( z_2 \) in \( F \cap \{\text{W}(x_n)\}_{n \geq 0} \). Then it follows that \( (\langle x_n - z_1, z_2 - z_1 \rangle)_{n \geq 0} \) converges. Hence \( \langle z_1, z_2 \rangle = \langle z_2, z_2 - z_1 \rangle \), i.e., \( z_1 = z_2 \). Thus, \( F \cap \{\text{W}(x_n)\}_{n \geq 0} \) contains at most one point. Since \( (x_n)_{n \geq 0} \) is bounded, the assertion is proved. \( \square \)

Our formalization of Fejér-monotonicity will revolve around the following class of operators.

**Definition 2.2** \( \mathcal{F} = \{T : \mathcal{H} \to \mathcal{H} \mid \text{dom} T = \mathcal{H} \text{ and } (\forall x \in \mathcal{H}) \text{ Fix } T \subset H(x, Tx)\} \).

The class \( \mathcal{F} \) contains several types of operators commonly found in various areas of applied mathematics, and in particular in approximation and optimization theory.

**Example 2.3** An operator \( T : \mathcal{H} \to \mathcal{H} \) lies in \( \mathcal{F} \) if one of the following conditions is satisfied.

(i) \( \text{dom} T = \mathcal{H} \) and \( T \) is firmly nonexpansive:

\[
(\forall (x, y) \in \mathcal{H}^2) \quad \langle (T - \text{Id}) x - (T - \text{Id}) y \mid Tx - Ty \rangle \leq 0.
\]

(ii) \( T \) is the resolvent of a maximal monotone operator \( A : \mathcal{H} \to 2^{\mathcal{H}} \), i.e., \( T = (\text{Id} + \gamma A)^{-1} \) where \( \gamma \in [0, +\infty[ \).

(iii) \( T \) is the projector onto a nonempty closed and convex set \( C \subset \mathcal{H} \).

(iv) Given a continuous convex function \( f : \mathcal{H} \to \mathbb{R} \) such that \( \text{lev}_{\leq 0} f \neq \emptyset \),

\[
T : x \mapsto \begin{cases} 
    x - \frac{f(x)}{\|t\|^2}t, & \text{where } t \in \partial f(x) \text{ if } f(x) > 0 \\
    x & \text{if } f(x) \leq 0.
\end{cases}
\]

\[\text{(2.2)}\]
Proof. (i) is straightforward. (ii) $\Leftrightarrow$ (i): see [9] (see also [25] for (ii) $\Rightarrow$ (i)). (iii) $\Rightarrow$ (ii): Take $A = A_c$ [6, Ex. 2.8.2]. Alternatively, (iii) $\Rightarrow$ (i) is shown in [16, Ch. 12]. (iv): First, observe that $T$ is well defined everywhere since $\operatorname{dom} \partial f = \mathcal{H}$ and $f(x) > 0 \Rightarrow f(x) > \inf_{y \in \mathcal{H}} f(y) \Rightarrow t \neq 0$. Moreover, $\operatorname{Fix} T = \{x \leq f\}$. Now take $x \in \mathcal{C} \operatorname{Fix} T$, $t \in \partial f(x)$, and $y \in \operatorname{Fix} T$. Then $\langle y-x \mid t \rangle + f(x) \leq f(y) \leq 0$. Consequently, $\langle y-Tx \mid x-Tx \rangle \leq 0$, i.e., $y \in H(x,Tx)$. We conclude $(\forall x \in \mathcal{H}) \operatorname{Fix} T \subset H(x,Tx)$. □

Proposition 2.4 Every $T$ in $\mathfrak{T}$ satisfies the following properties.

(i) $\operatorname{Fix} T = \bigcap_{x \in \mathcal{H}} H(x,Tx)$.

(ii) $\operatorname{Fix} T$ is closed and convex.

(iii) $(\forall \lambda \in [0,1])$ Id $+ \lambda(T - \text{Id}) \in \mathfrak{T}$.

Proof. (i): Definition 2.2 states that $\operatorname{Fix} T \subset \bigcap_{x \in \mathcal{H}} H(x,Tx)$ and we must therefore show that $\bigcap_{x \in \mathcal{H}} H(x,Tx) \subset \operatorname{Fix} T$. This follows from the implications $y \in \bigcap_{x \in \mathcal{H}} H(x,Tx) \Rightarrow y \in H(y,Ty) \Rightarrow \|y - Ty\|^2 \leq 0 \Rightarrow y \in \operatorname{Fix} T$. (i) $\Rightarrow$ (ii) since the sets $(H(x,Tx))_{x \in \mathcal{H}}$ are closed and convex. (iii): If $\lambda = 0$ the result is straightforward. Now take $x \in \mathcal{H}$, $y \in H(x,Tx)$, $\lambda \in [0,1]$, and let $T' = \text{Id} + \lambda(T - \text{Id})$. Then dom $T' = \text{dom} T = \mathcal{H}$ and

$$
\langle y-T'x \mid x-T'x \rangle = \lambda \langle y-Tx \mid x-Tx \rangle - \lambda(1-\lambda)\|x-Tx\|^2 \leq \lambda \langle y-Tx \mid x-Tx \rangle \leq 0. \tag{2.3}
$$

Thus, we get $\operatorname{Fix} T' = \operatorname{Fix} T \subset H(x,Tx) \subset H(x,T'x)$ and the proof is complete. □

We now describe a general scheme to construct Fejér-monotone sequences.

Proposition 2.5 Let $F$ be a nonempty closed and convex subset of $\mathcal{H}$. Then a sequence $(x_n)_{n \geq 0} \subset \mathcal{H}$ is Fejér-monotone with respect to $F$ if and only if

$$
(\forall n \in \mathbb{N}) \quad x_{n+1} = 2T_n x_n - x_n, \tag{2.4}
$$

where $(T_n)_{n \geq 0} \subset \mathfrak{T}$ and $F \subset \bigcap_{n \geq 0} \operatorname{Fix} T_n$.

Proof. Take a sequence $(x_n)_{n \geq 0}$ constructed as in (2.4), where $(T_n)_{n \geq 0} \subset \mathfrak{T}$ and $F \subset \bigcap_{n \geq 0} \operatorname{Fix} T_n$. Next, fix $z \in F$ and $n \in \mathbb{N}$. Then $z \in \operatorname{Fix} T_n \subset H(x_n,T_n x_n)$ and therefore

$$
\|x_{n+1} - z\|^2 = \|x_n - z\|^2 + 4\langle z - T_n x_n \mid x_n - T_n x_n \rangle \leq \|x_n - z\|^2, \tag{2.5}
$$

which shows that $(x_n)_{n \geq 0}$ is Fejér-monotone. Conversely, suppose that $(x_n)_{n \geq 0}$ is Fejér-monotone. For every $n \in \mathbb{N}$, let $T_n$ be the projector onto the nonempty closed convex set

$$
H_n = \{z \in \mathcal{H} \mid \|x_{n+1} - z\|^2 \leq \|x_n - z\|^2\} = \left\{z \in \mathcal{H} \mid \left\langle \frac{z - x_n + x_{n+1}}{2} \mid x_n - x_{n+1} \right\rangle \leq 0 \right\}. \tag{2.6}
$$
Then $T_n x_n = (x_n + x_{n+1})/2$ and the recursion (2.4) holds. Moreover, dom $T_n = \mathcal{H}$ and $F \subset H_n = \operatorname{Fix} T_n = H(x_n, T_n x_n)$. $T_n$ therefore satisfies the required conditions. \qed

Henceforth, we shall consider a slightly less general iterative model.

**Algorithm 2.6** Given $x_0 \in \mathcal{H}$ and $\varepsilon \in [0,1]$ a sequence $(x_n)_{n \geq 0}$ is constructed as follows. At iteration $n \in \mathbb{N}$, suppose that $x_n$ is given. Then select $T_n \in \mathcal{K}$ and set $x_{n+1} = x_n + (2-\varepsilon)(T_n x_n - x_n)$.

**Theorem 2.7** Let $(x_n)_{n \geq 0}$ be an arbitrary orbit of Algorithm 2.6 and suppose that $F = \bigcap_{n \geq 0} \operatorname{Fix} T_n \neq \emptyset$. Then:

(i) $x_n \xleftarrow{\mathcal{R}} x \in F \iff \mathcal{W}(x_n)_{n \geq 0} \subset F$.

(ii) $\sum_{n \geq 0} \|x_{n+1} - x_n\|^2 < +\infty$ and $\sum_{n \geq 0} \|x_n - T_n x_n\|^2 < +\infty$.

**Proof.** (i): For every $n \in \mathbb{N}$ and $z \in F$, $z \in \operatorname{Fix} T_n \subset H(x_n, T_n x_n)$ and therefore

$$
\|x_{n+1} - z\|^2 = \|x_n - z\|^2 + 2(2-\varepsilon)(z - T_n x_n \mid x_n - T_n x_n) - \varepsilon(2-\varepsilon) \|x_n - T_n x_n\|^2 \\
\leq \|x_n - z\|^2 - \varepsilon(2-\varepsilon) \|x_n - T_n x_n\|^2 \\
\leq \|x_n - z\|^2.
$$

Hence $(x_n)_{n \geq 0}$ is Fejér-monotone and the assertion follows from Proposition 2.1. (ii): By virtue of (2.7), we get

$$
\sum_{n \geq 0} \|x_{n+1} - x_n\|^2 = (2-\varepsilon)^2 \sum_{n \geq 0} \|x_n - T_n x_n\|^2 \leq (2-\varepsilon) \|x_0 - z\|^2 / \varepsilon.
$$

\qed

**Remark 2.8** The summability properties need not hold for the more general iteration (2.4): For instance, take $x_0 \neq 0$ and $\forall n \in \mathbb{N} \ T_n : x \mapsto 0$. Then $F = \{0\}$ and (2.4) produces the sequence $((-1)^n x_0)_{n \geq 0}$.

## 3 An abstract Haugazeau method

In this section, we investigate a generalization of (1.3) based on the same operator theoretic framework as in Algorithm 2.6.

**Algorithm 3.1** Given $x_0 \in \mathcal{H}$, a sequence $(x_n)_{n \geq 0}$ is constructed as follows. At iteration $n \in \mathbb{N}$, suppose that $x_n$ is given and select $T_n \in \mathcal{K}$. If $H(x_0, x_n) \cap H(x_n, T_n x_n) \neq \emptyset$, set $x_{n+1} = Q(x_0, x_n, T_n x_n)$; otherwise stop.
In [19], a necessary and sufficient condition was derived for $H(x, y) \cap H(y, z) = \emptyset$ as well as the expression of $Q(x, y, z)$ when $H(x, y) \cap H(y, z) \neq \emptyset$. With these results, the conceptual Algorithm 3.1 can be rewritten more explicitly.

**Algorithm 3.2** *(Explicit reformulation of Algorithm 3.1)* A sequence $(x_n)_{n \geq 0}$ is constructed as follows.

Step 0. *Set $n = 0$ and fix $x_0 \in \mathcal{H}$.*

Step 1. *Select $T_n \in \mathcal{X}$.*

Step 2. *Set $\pi_n = \langle x_0 - x_n \mid x_n - T_nx_n \rangle$, $\mu_n = \|x_0 - x_n\|^2$, $\nu_n = \|x_n - T_nx_n\|^2$, and $\rho_n = \mu_n\nu_n - \pi_n^2$.*

Step 3. *If $\rho_n = 0$ and $\pi_n < 0$ stop. Else set

$$
x_{n+1} = \begin{cases} 
T_nx_n & \text{if } \rho_n = 0 \text{ and } \pi_n \geq 0, \\
x_0 + (1 + \pi_n/\nu_n)(T_nx_n - x_n) & \text{if } \rho_n > 0 \text{ and } \pi_n\nu_n \geq \rho_n, \\
x_n + \frac{\nu_n}{\rho_n}(\pi_n(x_0 - x_n) + \mu_n(T_nx_n - x_n)) & \text{if } \rho_n > 0 \text{ and } \pi_n\nu_n < \rho_n,
\end{cases}
$$

then set $n = n + 1$ and go to Step 1.*

From a numerical standpoint, it is important to observe that, given $T_nx_n$, the update equation $x_{n+1} = Q(x_0, x_n, T_nx_n)$ requires only modest computations. Therefore the bulk of the execution cost resides in the determination of the sequence $(T_nx_n)_{n \geq 0}$, just as in Algorithm 2.6.

Some basic properties of Algorithm 3.2 are detailed below.

**Proposition 3.3** Let $(x_n)_{n \geq 0}$ be an arbitrary orbit of Algorithm 3.2. Then:

(i) *If at iteration $n$ the point $x_{n+1}$ is defined, then $\|x_0 - x_n\| \leq \|x_0 - x_{n+1}\|$.*

(ii) *$(\forall n > 0)$ $x_n = x_0 \iff x_n = x_{n-1} = \ldots = x_0 \iff x_0 \in \bigcap_{k=0}^{n-1} \text{Fix} T_k$.*

(iii) *Either the algorithm terminates or it generates an infinite sequence $(x_n)_{n \geq 0}$ such that $(\|x_0 - x_n\|)_{n \geq 0}$ is increasing.*

(iv) *The algorithm terminates at iteration $n > 0$ if and only if

$$
x_n \neq x_0 \text{ and } (\exists \alpha \in [0, +\infty)) \ T_nx_n = \alpha x_0 + (1 - \alpha)x_n.
$$

(v) *The algorithm generates an infinite sequence if $F = \bigcap_{n \geq 0} \text{Fix} T_n \neq \emptyset$.***
Proof. (i): Let us first recall that the projector onto a nonempty closed convex set \( C \subset \mathcal{H} \) is characterized by [16, Ch. 12]

\[
(\forall x \in \mathcal{H}) \quad P_C x \in C \quad \text{and} \quad \langle x - P_C x \mid C - P_C x \rangle \leq 0.
\]

Hence, \( x_n \) is the projection of \( x_0 \) onto \( H(x_0, x_n) \) and \( x_{n+1} = Q(x_0, x_n, T_n x_n) \Rightarrow x_{n+1} \in H(x_0, x_n) \Rightarrow \| x_0 - x_n \| \leq \| x_0 - x_{n+1} \| \). (ii): The first equivalence follows from (i) and the second one can be established by induction. Indeed, it holds for \( n = 1 \) since \( x_1 = Q(x_0, x_0, T_0 x_0) = T_0 x_0 \). Furthermore, if it holds for some \( n > 0 \), then

\[
x_{n+1} = x_n = \ldots = x_0 \Leftrightarrow \begin{cases} 
  x_0 \in \bigcap_{k=0}^{n-1} \text{Fix} T_k \\
  x_0 = x_{n+1} = Q(x_0, x_0, T_n x_0) = T_n x_0 \Leftrightarrow x_0 \in \bigcap_{k=0}^{n} \text{Fix} T_k.
\end{cases}
\]

(iii) follows from (i). (iv): By the Cauchy-Schwarz inequality \( \rho_n \geq 0 \) and the conditions \( \rho_n = 0 \) and \( \pi_n < 0 \) are equivalent to stating that the vectors \( x_0 - x_n \) and \( x_n - T_n x_n \) are linearly dependent, nonzero, and their scalar product is strictly negative, whence (3.2). (v): It is sufficient to show

\[
F \subset \bigcap_{n \geq 0} H(x_0, x_n) \Rightarrow F \subset H(x_0, x_n) \cap H(x_n, T_n x_n), \quad \text{i.e., since} \quad (T_n)_{n \geq 0} \subset \mathcal{F}, \quad F \subset \bigcap_{n \geq 0} H(x_0, x_n). \quad \text{For} \quad n = 0, \quad \text{it is clear that} \quad F \subset H(x_0, x_n) = \mathcal{H}. \quad \text{Furthermore, for every} \quad n \in \mathbb{N}, \quad \text{it results from (3.3) that}
\]

\[
F \subset H(x_0, x_n) \Rightarrow F \subset H(x_0, x_n) \cap H(x_n, T_n x_n) \Rightarrow F \subset H(x_0, Q(x_0, x_n, T_n x_n)) \Rightarrow F \subset H(x_0, x_{n+1}),
\]

which establishes the assertion by induction. \( \square \)

Next, we turn our attention to the convergence properties of Algorithm 3.2.

**Theorem 3.4** Let \( (x_n)_{n \geq 0} \) be an arbitrary orbit of Algorithm 3.2 and let \( F = \bigcap_{n \geq 0} \text{Fix} T_n \). Then

(i) *If the algorithm does not terminate, then* \( (x_n)_{n \geq 0} \) *is bounded* \( \Leftrightarrow \) \( (\| x_0 - x_n \|)_{n \geq 0} \) *converges.*

(ii) *If* \( \emptyset \neq \mathcal{O} \), *then* \( (x_n)_{n \geq 0} \) *is bounded and* \( (\forall n \in \mathbb{N}) \quad x_n \in F \Leftrightarrow x_n = P_F x_0. \)

(iii) *If* \( \emptyset \neq \mathcal{O} \), *then* \( (\| x_0 - x_n \|)_{n \geq 0} \) *converges and* \( \lim_n \| x_0 - x_n \| \leq \| x_0 - P_F x_0 \|. \)

(iv) *If* \( \emptyset \neq \mathcal{O} \), *then* \( x_n \xrightarrow{n} x_0 \Leftrightarrow P_F x_0 \subset \mathcal{M}(x_n)_{n \geq 0} \subset F. \)

(v) *If the algorithm does not terminate and* \( (x_n)_{n \geq 0} \) *is bounded, then* \( \sum_{n \geq 0} \| x_{n+1} - x_n \|^2 < +\infty \) *and* \( \sum_{n \geq 0} \| x_n - T_n x_n \|^2 < +\infty. \)

*Proof.* (i) follows from Proposition 3.3(i). (ii): As shown in the proof of Proposition 3.3(v), \( F \subset \bigcap_{n \geq 0} H(x_0, x_n) \). Hence,

\[
(\forall n \in \mathbb{N}) \quad P_F x_0 \in F \subset H(x_0, x_n) \Rightarrow \| x_0 - x_n \| \leq \| x_0 - P_F x_0 \|. \quad (3.6)
\]
(iii) follows from (ii), (i), and the previous inequality. (iv): The forward implication is trivial. For the reverse implication, fix \( x \in \mathcal{W}(x_n)_{n \geq 0} \), say \( x_n \xrightarrow{k} x \). Such a point does exist for \( (x_n)_{n \geq 0} \) is bounded by (ii). It follows from the weak lower semicontinuity of \( \| \cdot \| \) and (iii) that
\[
\| x_0 - x \| \leq \lim_k \| x_0 - x_n \| = \lim_n \| x_0 - x_n \| \leq \| x_0 - P_F x_0 \|. \tag{3.7}
\]
Consequently, since \( x \in F \), \( x = P_F x_0 \) and, in turn, \( \mathcal{W}(x_n)_{n \geq 0} = \{ P_F x_0 \} \). Next, since \( (x_n)_{n \geq 0} \) is bounded, we obtain \( x_n \xrightarrow{n} P_F x_0 \). The weak lower semicontinuity of \( \| \cdot \| \) and (iii) then yield
\[
\| x_0 - P_F x_0 \| \leq \lim_n \| x_0 - x_n \| \leq \| x_0 - P_F x_0 \|. \tag{3.8}
\]
Therefore \( \| x_0 - x_n \| \xrightarrow{n} \| x_0 - P_F x_0 \| \). However,
\[
\| x_n - P_F x_0 \|^2 = \| x_0 - x_n \|^2 - \| x_0 - P_F x_0 \|^2 + 2\langle x_n - P_F x_0 \mid x_0 - P_F x_0 \rangle. \tag{3.9}
\]
Hence, we conclude \( x_n \xrightarrow{n} P_F x_0 \). (v): For every \( n \in \mathbb{N} \), the inclusion \( x_{n+1} \in H(x_0, x_n) \) implies
\[
\| x_0 - x_{n+1} \|^2 - \| x_0 - x_n \|^2 = \| x_{n+1} - x_n \|^2 + 2\langle x_{n+1} - x_n \mid x_n - x_0 \rangle \geq \| x_{n+1} - x_n \|^2. \tag{3.10}
\]
Hence \( \sum_{n \geq 0} \| x_{n+1} - x_n \|^2 \leq \sup_{n \geq 0} \| x_0 - x_n \|^2 < +\infty \) for \( (x_n)_{n \geq 0} \) is bounded. In turn, since for every \( n \in \mathbb{N} \) the inclusion \( x_{n+1} \in H(x_n, T_n x_n) \) implies
\[
\| x_{n+1} - x_n \|^2 = \| x_{n+1} - T_n x_n \|^2 - 2\langle x_{n+1} - T_n x_n \mid x_n - T_n x_n \rangle + \| x_n - T_n x_n \|^2 \geq \| x_{n+1} - T_n x_n \|^2 + \| x_n - T_n x_n \|^2, \tag{3.11}
\]
we obtain \( \sum_{n \geq 0} \| x_n - T_n x_n \|^2 < +\infty \). \( \square \)

4 The weak-to-strong convergence principle

To achieve convergence in Algorithms 2.6 and 3.2, the sequence \( (T_n)_{n \geq 0} \) must be asymptotically well behaved, a notion that we formalize as follows.

**Definition 4.1** A sequence \( (T_n)_{n \geq 0} \subset \mathcal{X} \) is coherent if for every bounded sequence \( (y_n)_{n \geq 0} \subset \mathcal{H} \) there holds
\[
\begin{cases}
\sum_{n \geq 0} \| y_{n+1} - y_n \|^2 < +\infty \\
\sum_{n \geq 0} \| y_n - T_n y_n \|^2 < +\infty
\end{cases} \Rightarrow \mathcal{W}(y_n)_{n \geq 0} \subset \bigcap_{n \geq 0} \text{Fix } T_n. \tag{4.1}
\]

This property allows us to view the convergence of Algorithms 2.6 and 3.2 from a single perspective.

**Theorem 4.2** Suppose that \( (T_n)_{n \geq 0} \) is coherent and let \( F = \bigcap_{n \geq 0} \text{Fix } T_n \). Then:
(i) If $F \neq \emptyset$, then every orbit of Algorithm 2.6 converges weakly to a point in $F$.

(ii) (Trichotomy) For an arbitrary orbit $(x_n)_{n \geq 0}$ of Algorithm 3.2, exactly one of the following alternatives holds:

(a) $F \neq \emptyset$ and $x_n \xrightarrow{n} P_F x_0$.
(b) $F = \emptyset$ and $\|x_n\| \xrightarrow{n} +\infty$.
(c) $F = \emptyset$ and the algorithm terminates.

Proof. (i) follows from Theorem 2.7. (ii): Assertion (a) follows from items (ii) and (iv)-(v) in Theorem 3.4. It remains to prove (b), in which $F = \emptyset$ and $(x_n)_{n \geq 0}$ is fully defined. Suppose that $\|x_n\| \xrightarrow{n} +\infty$. Then it follows from Proposition 3.3(iii) that $(x_n)_{n \geq 0}$ is bounded, and then from Theorem 3.4(v) that $\sum_{n \geq 0} \|x_{n+1} - x_n\|^2 < +\infty$ and $\sum_{n \geq 0} \|x_n - T_n x_n\|^2 < +\infty$. Condition (4.1) then yields $\emptyset \neq \mathcal{W}(x_n)_{n \geq 0} \subset \bigcap_{n \geq 0} \text{Fix } T_n = F = \emptyset$, which is absurd. \(\Box\)

Remark 4.3 (The weak-to-strong convergence principle) The practical significance of Theorem 4.2 is that when the solution set $F$ is not empty and the sequence $(T_n)_{n \geq 0}$ is coherent, not only the generic Fejér-monotone method described by Algorithm 2.6 converges weakly to a solution but it can also easily be transformed into a strongly convergent method in the form of Algorithm 3.2: it suffices to replace the updating rule $x_{n+1} = x_n + (2 - \varepsilon)(T_n x_n - x_n)$ by $x_{n+1} = Q(x_0, x_n, T_n x_n)$. Furthermore, while the solution produced by Algorithm 2.6 is in general an undetermined point in $F$, that of Algorithm 3.2 is precisely the projection of the initial point $x_0$ onto $F$.

As the examples below show, all three cases may occur in the trichotomy described in Theorem 4.2(ii).

Example 4.4 In Algorithm 3.2, take $z \neq x_0 = 0$ and at every iteration $n$:

(a) $T_n = P_{H_n}$, where $H_n = \{x \in \mathcal{H} : \langle x - (1 - 2^{-n}) z, z \rangle \geq 0\} = \text{Fix } T_n$. Then $(\forall n \in \mathbb{N})$ $x_n = (1 - 2^{-n}) z$ and $F = \{x \in \mathcal{H} : \langle x - z, z \rangle \geq 0\} \neq \emptyset$. Hence $x_n \xrightarrow{n} P_F x_0 = z$.

(b) $T_n : x \mapsto x + z$. Then $F = \emptyset$ and $(\forall n \in \mathbb{N})$ $x_n = nz$.

(c) $T_n : x \mapsto x + (-1)^n z$. Then $F = \emptyset$ and we get successively $T_0 x_0 = z$, $x_1 = z$, and $T_1 x_1 = 0$. Hence $H(x_0, x_1) \cap H(x_1, T_1 x_1) = H(0, z) \cap H(z, 0) = \emptyset$ and the algorithm stops.

We conclude this section with a useful fact.

Proposition 4.5 Fix $\delta \in [0, 1]$, $(T_n)_{n \geq 0} \subset \mathcal{F}$, and define

$$\text{Fix } \delta \in [0, 1], (T_n)_{n \geq 0} \subset \mathcal{Y}, \text{ and define}$$

$$(\forall n \in \mathbb{N}) \quad T'_n = \text{Id} + \lambda_n (T_n - \text{Id}) \quad \text{where } \lambda_n \in [\delta, 1].$$

Then $(T'_n)_{n \geq 0}$ is coherent if and only if $(T_n)_{n \geq 0}$ is coherent.

Proof. $(T'_n)_{n \geq 0} \subset \mathcal{F}$ by Proposition 2.4(iii). Moreover, $\sum_{n \geq 0} \|y_n - T'_n y_n\|^2 < +\infty \Leftrightarrow \sum_{n \geq 0} \|y_n - T'_n y_n\|^2 < +\infty$, and $(\forall n \in \mathbb{N}) \text{ Fix } T'_n = \text{ Fix } T_n$. \(\Box\)
5 Constraint disintegration methods

A general computational strategy to solve complex problems is to disintegrate the solution set $S$ as an intersection of simpler sets and to devise an iterative method in which only one of these sets is activated at each iteration [21]. In this section, we present one such implementation of Algorithms 2.6 and 3.2 for solving the convex feasibility problem

$$\text{Find } x \in S = \bigcap_{i \in I} S_i,$$

where $(S_i)_{i \in I}$ is a countable family of possibly empty closed convex sets in $\mathcal{H}$. Alternative generalizations of (1.3) are studied in [13] from a distinct viewpoint.

**Assumption 5.1** $(\forall n \in \mathbb{N}) \; T_n \in \mathfrak{S} \; \text{and} \; \text{Fix} T_n = S_{\iota(n)}$, where

(i) The index control mapping $\iota : \mathbb{N} \to I$ satisfies

$$(\forall i \in I)(\exists M_i > 0)(\forall n \in \mathbb{N}) \; i \in \{\iota(n), \ldots, \iota(n + M_i - 1)\}. \tag{5.2}$$

(ii) For every $i \in I$, every bounded sequence $(y_n)_{n \geq 0} \subset \mathcal{H}$, and every strictly increasing sequence $(n_k)_{k \geq 0} \subset \mathbb{N}$ there holds

$$\begin{cases}
y_{n_k} \xrightarrow{k} y \\
(\forall k \in \mathbb{N}) \; \iota(n_k) = i \\
\sum_{n \geq 0} \|y_{n+1} - y_n\|^2 < +\infty \\
\sum_{n \geq 0} \|y_n - T_n y_n\|^2 < +\infty
\end{cases} \Rightarrow y \in S_i. \tag{5.3}$$

Condition (5.2) ensures that, for every index $i$, the set $S_i$ is activated at least once within any $M_i$ consecutive iterations.

**Example 5.2** Take $I = \mathbb{N}$. Then the mapping $\iota$ defined by $(\forall i \in I)(\forall k \in \mathbb{N}) \; \iota(2^k(2k + 1) - 1) = i$ satisfies (5.2) with $(\forall i \in I) \; M_i = 2^{i+1}$.

We now state and prove the convergence to a solution of (5.1) of two constraint disintegration schemes based on Algorithms 2.6 and 3.2.

**Theorem 5.3** Fix $x_0 \in \mathcal{H}$ and $\varepsilon \in [0, 1]$. Then, under Assumption 5.1:

(i) Every orbit of the algorithm

$$(\forall n \in \mathbb{N}) \; x_{n+1} = x_n + \lambda_n (T_n x_n - x_n) \quad \text{where} \quad \lambda_n \in [\varepsilon, 2 - \varepsilon] \tag{5.4}$$

converges weakly to a point in $S$ if $S \neq \emptyset$. 

10
(ii) Every orbit of the algorithm

\[(\forall n \in \mathbb{N}) \quad x_{n+1} = Q(x_0, x_n, x_n + \lambda_n(T_n x_n - x_n)) \quad \text{where} \quad \lambda_n \in [\varepsilon, 1] \quad (5.5)\]

converges strongly to \( P_S x_0 \) if \( S \neq \emptyset \); if \( S = \emptyset \), either (5.5) terminates or \( \|x_n\| \to +\infty \).

\[ \text{Proof.} \quad \text{Let us first establish that } (T_n)_{n \geq 0} \text{ is coherent, i.e., that (4.1) holds. To this end, take a bounded sequence } (y_n)_{n \geq 0} \subset \mathcal{H} \text{ such that } \sum_{n \geq 0} \|y_{n+1} - y_n\|^2 < +\infty \text{ and } \sum_{n \geq 0} \|y_n - T_n y_n\|^2 < +\infty. \]

Next, fix \( i \in I \) and \( y \in \mathfrak{M}(y_n)_{n \geq 0} \), say \( y_n \xrightarrow{k} y \). Then, since by Assumption 5.1(i)

\[ \bigcap_{n \geq 0} \text{Fix } T_n = \bigcap_{n \geq 0} S_{\iota(n)} = \bigcap_{i \in I} S_i = S, \quad (5.6) \]

it suffices to show \( y \in S_i \). Condition (5.2) guarantees the existence of a strictly increasing sequence \( (p_k)_{k \geq 0} \subset \mathbb{N} \) such that

\[ (\forall k \in \mathbb{N}) \quad n_k \leq p_k \leq n_k + M_i - 1 \quad \text{and} \quad \iota(p_k) = i. \quad (5.7) \]

Hence, upon invoking the Cauchy-Schwarz inequality, we get

\[ (\forall k \in \mathbb{N}) \quad \|y_{p_k} - y_{n_k}\| \leq \sum_{l = n_k}^{n_k + M_i - 1} \|y_{l+1} - y_l\| \leq \sqrt{M_i} \sum_{l \geq n_k} \|y_{l+1} - y_l\|^2. \quad (5.8) \]

Consequently

\[ \sum_{n \geq 0} \|y_{n+1} - y_n\|^2 < +\infty \Rightarrow y_{p_k} - y_{n_k} \xrightarrow{k} 0 \Rightarrow y_{p_k} \xrightarrow{k} y. \quad (5.9) \]

In view of (5.7) and Assumption 5.1(ii), \( y \in S_i \) and consequently \( (T_n)_{n \geq 0} \) is coherent. However, for any relaxation sequence \( (\lambda_n)_{n \geq 0} \subset [\varepsilon/(2 - \varepsilon), 1] \), \( (\text{Id} + \lambda'_n(T_n - \text{Id}))_{n \geq 0} \) is also coherent by virtue of Proposition 4.5. It therefore follows from Theorem 4.2(i) that every orbit of the algorithm

\[ (\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + (2 - \varepsilon) \lambda'_n(T_n x_n - x_n) \quad \text{where} \quad \lambda'_n \in [\varepsilon/(2 - \varepsilon), 1] \quad (5.10) \]

or, equivalently, of (5.4) converges weakly to a point in \( \bigcap_{n \geq 0} \text{Fix } T_n \). In light of (5.6), assertion (i) is proved. Likewise, as Proposition 4.5 asserts that \( (\text{Id} + \lambda_n(T_n - \text{Id}))_{n \geq 0} \) is coherent, assertion (ii) follows from Theorem 4.2(ii). \( \Box \)

6 Applications

In this section the weak-to-strong convergence principle is applied, in the form of Theorem 5.3, to specific situations. Further examples can be constructed by considering the Fejér-monotone methods described in [2], [12], and [20].
6.1 Common zeros of monotone operators

Let \((A_i)_{i \in I}\) be a countable family of maximal monotone operators from \(\mathcal{H}\) into \(2^\mathcal{H}\). Our first application concerns the problem of constructing a common zero of the operators \((A_i)_{i \in I}\), i.e.,

\[
\text{Find } x \in \mathcal{H} \text{ such that } (\forall i \in I) \ 0 \in A_i x.
\]  

(6.1)

Alternatively, since the set of zeros of a maximal monotone operator is closed and convex, this problem can be formulated in the format (5.1) by letting \(S\) be the set of common zeros and \((\forall i \in I) S_i = A_i^{-1}0\).

Henceforth, \(A_i, \gamma = (\text{Id} - (\text{Id} + \gamma A_i)^{-1})/\gamma\) denotes the Yosida approximation of \(A_i\) of index \(\gamma \in ]0, +\infty[\) and \(\text{gr}A_i = \{(v, w) \in \mathcal{H}^2 \mid w \in A_i v\}\) its graph. We shall exploit the fact that \(\text{gr}A_i\) is weakly-strongly closed [6, Prop. 2.5]:

\[
(\forall (v_n, w_n))_{n \geq 0} \subset \text{gr}A_i \quad \begin{cases} v_n \overset{n}{\rightarrow} v \\ w_n \overset{n}{\rightarrow} w \end{cases}
\]  

\(\Rightarrow (v, w) \in \text{gr}A_i.
\]  

(6.2)

**Corollary 6.1** Fix \(x_0 \in \mathcal{H}, \ \varepsilon \in ]0, 1[, \) and \((\gamma_n)_{n \geq 0} \subset ]0, +\infty[\). Suppose that \(i: \mathbb{N} \to I\) satisfies condition (5.2) and that, for every \(i \in I\) and every strictly increasing sequence \((n_k)_{k \geq 0} \subset \mathbb{N}\) such that \((\forall k \in \mathbb{N}) i(n_k) = i\), there holds \(\inf_{k \geq 0} \gamma_{n_k} > 0\). Then:

(i) **Every orbit of the algorithm**

\[
(\forall n \in \mathbb{N}) \ x_{n+1} = x_n - \gamma_n \lambda_n A_{i(n), \gamma_n} x_n \text{ where } \lambda_n \in [\varepsilon, 2 - \varepsilon]
\]  

converges weakly to a point in \(S\) if \(S \neq \emptyset\).

(ii) **Every orbit of the algorithm**

\[
(\forall n \in \mathbb{N}) \ x_{n+1} = Q(x_0, x_n, x_n - \gamma_n \lambda_n A_{i(n), \gamma_n} x_n) \text{ where } \lambda_n \in [\varepsilon, 1]
\]  

converges strongly to \(P_S x_0\) onto \(S\) if \(S \neq \emptyset\); if \(S = \emptyset\) either (6.4) terminates or \(\|x_n\| \overset{n \rightarrow +\infty}{\rightarrow} +\infty\).

**Proof.** Let \((T_n)_{n \geq 0} = (\text{Id} - \gamma_{n} A_{i(n), \gamma_n})_{n \geq 0}\). Then, for every \(n \in \mathbb{N}\), Fix \(T_n = A_{i(n)}^{-1} 0 = S_{i(n)}\) and \(T_n \in \mathcal{T}\) by Example 2.3(ii). Next, we observe that (6.3) conforms to (5.4) and (6.4) to (5.5). Therefore, the announced result will follow from Theorem 5.3 if we show that Assumption 5.1(ii) is satisfied. To this end, fix \(i \in I\) and take a bounded sequence \((y_n)_{n \geq 0}\) from which a subsequence \((y_{n_k})_{k \geq 0}\) can be extracted such that \(y_{n_k} \overset{k}{\rightarrow} y\) and \((\forall k \in \mathbb{N}) i(n_k) = i\). Next define \((\forall k \in \mathbb{N}) v_k = y_{n_k} - \gamma_{n_k} A_{i, \gamma_{n_k}} y_{n_k}\) and note that \((\forall k \in \mathbb{N}) A_{i, \gamma_{n_k}} y_{n_k} \in A_i v_k\). Now suppose \(\sum_{n \geq 0} \gamma_n^2 \|A_{i(n), \gamma_n} y_n\|^2 < +\infty\). Then \(v_k - y_{n_k} \overset{k}{\rightarrow} 0\) and therefore \(v_k \overset{k}{\rightarrow} y\). Moreover, it follows from the assumption \(\inf_{k \geq 0} \gamma_{n_k} > 0\) that \(A_{i, \gamma_{n_k}} y_{n_k} \overset{k}{\rightarrow} 0\). To sum up,

\[
(\forall k \geq 0) ((v_k, A_{i, \gamma_{n_k}} y_{n_k})) \subset \text{gr}A_i
\]  

\[
v_k \overset{k}{\rightarrow} y
\]  

(6.5)

\[
A_{i, \gamma_{n_k}} y_{n_k} \overset{k}{\rightarrow} 0.
\]

12
Therefore (6.2) implies $0 \in A_{1}y$ and (5.3) ensues. \(\square\)

As discussed in [25], a special case of (6.1) of great interest is the problem of finding a zero of a single maximal monotone operator $A : \mathcal{H} \to 2^\mathcal{H}$. In this case, trichotomy reduces to dichotomy and Corollary 6.1 with $\lambda_n = 1$ for every $n \in \mathbb{N}$ specializes to

**Corollary 6.2** Fix $x_0 \in \mathcal{H}$ and $(\gamma_n)_{n \geq 0} \subset ]0, +\infty]$ such that $\inf_{n \geq 0} \gamma_n > 0$. Then:

(i) *Every orbit of the algorithm*

\[
(\forall n \in \mathbb{N}) \quad x_{n+1} = (\text{Id} + \gamma_n A)^{-1} x_n \quad \text{(6.6)}
\]

converges weakly to a zero of $A$ if $A^{-1} 0 \neq \emptyset$.

(ii) *Every orbit of the algorithm*

\[
(\forall n \in \mathbb{N}) \quad x_{n+1} = Q \left( x_0, x_n, (\text{Id} + \gamma_n A)^{-1} x_n \right) \quad \text{(6.7)}
\]

converges strongly to the projection of $x_0$ onto $A^{-1} 0$ if $A^{-1} 0 \neq \emptyset$; if $A$ has no zero then $\|x_n\| \xrightarrow{n \to +\infty} +\infty$.

**Proof.** In light of Corollary 6.1, it is enough to assume $A^{-1} 0 = \emptyset$ and to show that algorithm (6.7) does not terminate. Our proof is patterned after that of [26, Th. 9] and is based on a truncation argument found in [25]. Suppose that the iterates $(x_k)_{0 \leq k \leq n}$ are well defined for some $n > 0$ (this is certainly true for $n = 1$). Let

\[
(\forall k \in \{0, \ldots, n\}) \quad v_k = (\text{Id} + \gamma_k A)^{-1} x_k \quad \text{and} \quad A' = A + N_C,
\]

where

\[
C = \left\{ v \in \mathcal{H} \mid \|v\| \leq 1 + \max_{0 \leq k \leq n} \|v_k\| \right\}. \quad \text{(6.8)}
\]

Recall that $N_C$ is maximal monotone with $\text{dom} N_C = C$ and that $(\forall v \in \text{int}(C)) \ N_C v = \{0\}$ [6, Ex. 2.8.2]. Thus, since by construction $(v_k)_{0 \leq k \leq n} \subset \text{dom} A \cap \text{int}(C)$, we deduce on the one hand that

\[
(\forall k \in \{0, \ldots, n\}) \quad x_k \in v_k + \gamma_k A v_k = v_k + \gamma_k A' v_k, \quad \text{i.e.,} \quad v_k \in (\text{Id} + \gamma_k A')^{-1} x_k \quad \text{(6.9)}
\]

and, on the other hand, that $A'$ is maximal monotone by virtue of Rockafellar's sum theorem [6, Cor. 2.7]. Consequently, its resolvents are single-valued [6, Prop. 2.2] and we derive from (6.9) that $(\forall k \in \{0, \ldots, n\}) \ v_k = (\text{Id} + \gamma_k A')^{-1} x_k$. Therefore, up to iteration $n$, replacing $A$ by $A'$ does not affect the behavior of algorithm (6.7). However, since $\text{dom} A' \subset C$ is bounded, $A'$ is surjective [6, Cor. 2.2] and it therefore has zeros. Hence, arguing as in the proof of Proposition 3.3(v), we obtain $\emptyset \neq (A')^{-1} 0 \subset \bigcap_{k=0}^{n} H(x_0, x_k) \cap H(x_k, v_k)$. We conclude that $x_{n+1}$ is well defined. \(\square\)

**Remark 6.3** Corollary 6.2(i) gives the weak convergence to a zero of $A$ of the classical proximal point algorithm, i.e., of the composition product $\prod_{n \geq 0} (\text{Id} + \gamma_n A)^{-1} x_0$. Such results go back to the prox-regularization method of [22] (see also [7] and [25] for extensions to inexact iterations and strong convergence conditions). An instance of Corollary 6.2(i) in which strong convergence fails is constructed in [18, Cor. 5.1].

13
Remark 6.4 A result closely related to Corollary 6.2(ii) was obtained via a different analysis in [26]. In the method considered there, the update $x_{n+1}$ is the projection of $x_0$ onto $H(x_0, x_n) \cap H(x_n + u_n, y_n)$, where $y_n = (\text{Id} + \gamma_n A)^{-1}(x_n + u_n)$ and $\|u_n\| \leq \sigma \max\{\|x_n - y_n\|, \|x_n + u_n - y_n\|\}$ for some $\sigma \in [0, 1]$.

6.2 Common fixed points of nonexpansive operators

Given a countable family of nonexpansive operators $(R_i)_{i \in I}$ defined everywhere from $\mathcal{H}$ into $\mathcal{H}$, we consider the common fixed point problem

$$\text{Find } x \in \mathcal{H} \text{ such that } \forall i \in I \text{ } R_i x = x.$$  \hspace{1cm} (6.10)

Recall that $R_i$ is nonexpansive if $(\forall (x, y) \in \mathcal{H}^2) \|R_i x - R_i y\| \leq \|x - y\|$, which implies that $\text{Id} - R_i$ is maximal monotone and therefore demiclosed [8, Lem. 4]:

$$(\forall (y_n)_{n \geq 0} \subset \mathcal{H}) \begin{cases} y_n \rightarrow y \\ y_n - R_i y_n \rightarrow w \Rightarrow w = y - R_i y. \end{cases} \hspace{1cm} (6.11)$$

Since the fixed point set of a nonexpansive operator is closed and convex [16, Lem. 3.4], (6.10) can be cast in the form of (5.1) by letting $S$ be the set of common fixed points and $(\forall i \in I) \; S_i = \text{Fix } R_i$.

Corollary 6.5 Fix $x_0 \in \mathcal{H}$ and $\varepsilon \in [0, 1]$ and suppose that $\iota: \mathbb{N} \rightarrow I$ satisfies condition (5.2). Then:

(i) Every orbit of the algorithm

$$\forall n \in \mathbb{N} \quad x_{n+1} = x_n + \lambda_n (R_{\iota(n)} x_n - x_n) \text{ where } \lambda_n \in [\varepsilon, 1 - \varepsilon] \hspace{1cm} (6.12)$$

converges weakly to a point in $S$ if $S \neq \emptyset$ [8, Th. 5].

(ii) Every orbit of the algorithm

$$\forall n \in \mathbb{N} \quad x_{n+1} = Q (x_0, x_n, x_n + \lambda_n (R_{\iota(n)} x_n - x_n)) \text{ where } \lambda_n \in [\varepsilon, 1/2] \hspace{1cm} (6.13)$$

converges strongly to $P_S x_0$ if $S \neq \emptyset$; if $S = \emptyset$, either (6.13) terminates or $\|x_n\| \nrightarrow +\infty$.

Proof Let $(T_n)_{n \geq 0} = ((R_{\iota(n)} + \text{Id})/2)_{n \geq 0}$. Then, for every $n \in \mathbb{N}$, $\text{Fix } T_n = \text{Fix } R_{\iota(n)} = S_{\iota(n)}$ and $T_n$ is firmly nonexpansive [16, Th. 12.1]. Hence, $T_n \in \mathcal{F}$ by Example 2.3(i) and it follows from (6.11) that condition (5.3) is satisfied. Since (6.12) conforms to (5.4) and (6.13) to (5.5), the assertions follow from Theorem 5.3. \[\square\]

In the case of a single nonexpansive operator $R: \mathcal{H} \rightarrow \mathcal{H}$ with domain $\mathcal{H}$, dichotomy rather than trichotomy occurs in (ii) and we obtain

Corollary 6.6 Fix $x_0 \in \mathcal{H}$ and $\varepsilon \in [0, 1]$. Then:


(i) Every orbit of the algorithm

\[(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Rx_n - x_n) \text{ where } \lambda_n \in [\varepsilon, 1 - \varepsilon] \quad (6.14)\]

converges weakly to a point in \(\text{Fix} \, R\) if \(\text{Fix} \, R \neq \emptyset\) \([14, \text{ Th. 8}]\).

(ii) Every orbit of the algorithm

\[(\forall n \in \mathbb{N}) \quad x_{n+1} = Q(x_0, x_n, x_n + \lambda_n(Rx_n - x_n)) \text{ where } \lambda_n \in [\varepsilon, 1/2] \quad (6.15)\]

converges strongly to the projection of \(x_0\) onto \(\text{Fix} \, R\) if \(\text{Fix} \, R \neq \emptyset\); if \(R\) has no fixed point, then \(\|x_n\| \nrightarrow +\infty\).

Proof. This is an application of Corollary 6.5 and it must simply be shown that algorithm (6.15) does not terminate. Suppose that \(\text{Fix} \, R = \emptyset\) and, as is true for \(n = 1\), that the iterates \((x_k)_{0 \leq k \leq n}\) are well defined for some \(n > 0\). Let

\[R' = P_C \circ R, \quad \text{where } C = \left\{ x \in \mathcal{H} \mid \|x\| \leq \max_{0 \leq k \leq n} \|Rx_k\| \right\}. \quad (6.16)\]

Then \((\forall k \in \{0, \ldots, n\}) \quad R'x_k = Rx_k\) and it follows that, up to iteration \(n\), replacing \(R\) by \(R'\) does not affect the behavior of algorithm (6.15). However, \(R'\) maps the nonempty, closed, bounded, and convex set \(C\) into itself and it follows from the Browder-Göhde-Kirk theorem \([16, \text{ Th. 4.1}]\) that \(\text{Fix} \, R' \neq \emptyset\). Hence, we can argue as in the proof of Proposition 3.3(v) and establish that \(Q(x_0, x_n, x_n + \lambda_n(R'x_n - x_n))\) exists. Since \(R'x_n = Rx_n, x_{n+1}\) is therefore well defined. \(\square\)

Remark 6.7 An instance of Corollary 6.6(i) in which strong convergence fails is constructed in [15].

Remark 6.8 In some applications, a nonexpansive operator \(R\) may be defined only on a closed convex subset \(C\) of \(\mathcal{H}\). By setting \(R' = P_C \circ R \circ P_C\), we obtain a nonexpansive operator \(R'\) that is defined everywhere with \(\text{Fix} \, R' = \text{Fix} \, R\). Thus, we can apply Corollary 6.6 to \(R'\) rather than \(R\) to find a fixed points of \(R\). A similar remark can be made for Corollary 6.5.

6.3 Subgradient methods

Let \((f_i)_{i \in I}\) be a countable family of continuous convex functions from \(\mathcal{H}\) into \(\mathbb{R}\) such that the sets \((\text{lev}_{\leq 0} f_i)_{i \in I}\) are nonempty. Under consideration is the problem

\[\text{Find } x \in \mathcal{H} \text{ such that } \sup_{i \in I} f_i(x) \leq 0. \quad (6.17)\]

Upon calling \(S\) its set of solutions and setting \((\forall i \in I) \quad S_i = \text{lev}_{\leq 0} f_i, (6.17)\) is seen to fit into (5.1).
Subsequently, for every \( i \in I \) the operator \( G_i \) is defined to be

\[
G_i : x \mapsto \begin{cases} 
  x - \frac{f_i(x)}{\|t\|^2} t, & \text{if } f_i(x) > 0 \\
  x & \text{if } f_i(x) \leq 0
\end{cases}
\]

(6.18)

and we shall say that the subdifferential \( \partial f_i \) is bounded if it maps bounded sets into bounded sets (see [4] for a discussion of this property).

**Corollary 6.9** Fix \( x_0 \in \mathcal{H} \) and \( \varepsilon \in [0,1] \). Suppose that \( i : \mathbb{N} \to I \) satisfies condition (5.2) and that the subdifferentials \( (\partial f_i)_{i \in I} \) are bounded. Then:

(i) Every orbit of the algorithm

\[
(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n (G_{i(n)} x_n - x_n)
\]

converges weakly to a point in \( S \) if \( S \neq \emptyset \).

(ii) Every orbit of the algorithm

\[
(\forall n \in \mathbb{N}) \quad x_{n+1} = Q (x_0, x_n, x_n + \lambda_n (G_{i(n)} x_n - x_n))
\]

where \( \lambda_n \in [\varepsilon, 2-\varepsilon] \) and \( \lambda_n \in [\varepsilon, 1] \)

(6.20)

converges strongly to \( P_S x_0 \) if \( S \neq \emptyset \); if \( S = \emptyset \), either (6.20) terminates or \( \| x_n \| \xrightarrow{n} +\infty \).

**Proof.** Let \( (T_n)_{n \geq 0} = (G_{i(n)})_{n \geq 0} \). Then, for every \( n \in \mathbb{N} \), \( \text{Fix} T_n = \text{lev}_{\leq 0} f_{i(n)} = S_{i(n)} \) and \( T_n \in \mathcal{S} \) by Example 2.3(iv). Hence, in order to apply Theorem 5.3, it suffices to verify that Assumption 5.1(ii) holds. Fix \( i \in I \) and take a bounded sequence \( (y_n)_{n \geq 0} \) such that \( \sum_{n \geq 0} \| y_n - G_{i(n)} y_n \|^2 < +\infty \) and containing a subsequence \( (y_{n_k})_{k \geq 0} \) such that \( y_{n_k} \xrightarrow{k} y \) and \( (\forall k \in \mathbb{N}) \quad \iota(n_k) = i \). Then we must show \( f_i(y) \leq 0 \). Since \( f_i \) is weak lower semicontinuous, \( f_i(y) \leq \liminf_k f_i(y_{n_k}) \). Passing to a subsequence if necessary, we can assume \( (y_{n_k})_{k \geq 0} \subseteq \mathcal{S}_i \) (otherwise the conclusion is immediate). However, since \( \partial f_i \) is bounded, \( y_{n_k} - G_{i} y_{n_k} \xrightarrow{k} 0 \Rightarrow f_i(y_{n_k})/\| y_{n_k} \| \xrightarrow{k} 0 \Rightarrow f_i(y_{n_k}) \xrightarrow{k} 0 \). Therefore \( f_i(y) \leq 0 \). \( \square \)

The above results are applicable to the approximate minimization of a continuous convex function \( f : \mathcal{H} \to \mathbb{R} \) over a nonempty closed convex set \( C \subseteq \mathcal{H} \). Let us fix \( \mu \in \mathbb{R} \) such that \( \text{lev}_{\leq \mu} f \neq \emptyset \) and define the approximate solution set as \( S = C \cap \text{lev}_{\leq \mu} f \). Then the problem is a special instance of (6.17) with \( I = \{1, 2\} \), \( f_1 = f - \mu \), and \( f_2 = d(\cdot, C) \). Furthermore, since \( (\forall x \in \mathbb{C}) \partial d(x, C) = \{(x - P_C x)/d(x, C)\} \), (6.18) gives \( G_2 = P_C \) and Corollary 6.9 with (6.21) yields

\[
(\forall n \in \mathbb{N}) \quad \lambda_{2n} = \alpha_n, \lambda_{2n+1} = 1, \iota(2n) = 1, \text{ and } \iota(2n+1) = 2
\]

(6.21)

**Corollary 6.10** Fix \( x_0 \in \mathcal{H} \) and \( \varepsilon \in [0,1] \). Suppose that \( \Pi \triangleq \inf_{x \in C} f(x) > -\infty \) and that \( \partial f \) is bounded. Then:
(i) Every orbit of the algorithm
\[(\forall n \in \mathbb{N}) \; x_{n+1} = P_C(x_n + \alpha_n(G_1x_n - x_n)) \quad \text{where} \quad \alpha_n \in [\varepsilon, 2 - \varepsilon] \quad (6.22)\]
converges weakly to a point in \(S\) if \(S \neq \emptyset\), i.e., if \(\mu > \mu^*\) or if \(f\) has a minimizer on \(C\) and \(\mu = \mu^*\).

(ii) Every orbit of the algorithm
\[(\forall n \in \mathbb{N}) \; x_{n+1} = Q(x_0, z_n, P_Cz_n), \quad \text{where} \quad z_n = Q(x_0, x_n, x_n + \alpha_n(G_1x_n - x_n)) \quad \text{and} \quad \alpha_n \in [\varepsilon, 1] \quad (6.23)\]
converges strongly to \(P_Sx_0\) if \(S \neq \emptyset\); if \(S = \emptyset\), either (6.23) terminates or \(\|x_n\| \nrightarrow +\infty\).

Now, suppose that \(f\) admits a nonglobal minimizer over \(C\) and that \(\partial f\) maps the bounded subsets of \(C\) into bounded sets. Then we deduce from Corollary 6.10(i) that every orbit of the algorithm
\[
\begin{cases}
x_0 \in C \\
\varepsilon \in [0, 1]
\end{cases} \quad \text{and} \quad \text{(6.24)} \quad (\forall n \in \mathbb{N}) \; x_{n+1} = P_C\left(x_n + \alpha_n \frac{\mu - f(x_n)}{\|t_n\|^2}t_n\right) \quad \text{where} \quad \left\{ \begin{array}{l}
t_n \in \partial f(x_n) \\
\alpha_n \in [\varepsilon, 2 - \varepsilon]
\end{array} \right.
\]
converges weakly to a minimizer of \(f\) over \(C\). This classical result is due to Polyak [24, Th. 1].

### 6.4 Bregman’s and Haugazeau’s methods

We conclude the paper by revisiting the two algorithms which motivated the present work. More specifically, we recover the convergence results of Bregman’s and Haugazeau’s algorithms mentioned in the Introduction and describe the behavior of the latter in the inconsistent case. It is noteworthy that these results are consequences of any of Corollaries 6.1, 6.5, or 6.9.

**Corollary 6.11** Let \((S_i)_{1 \leq i \leq m}\) be nonempty closed convex subsets of \(H\) with intersection \(S\) and fix \(x_0 \in H\). Then:

(i) The orbit of algorithm (1.1) converges weakly to a point in \(S\) if \(S \neq \emptyset\) [5, Th. 1].

(ii) The orbit of algorithm (1.3) converges strongly to \(P_Sx_0\) if \(S \neq \emptyset\) [19, Th. 3-2]; if \(S = \emptyset\), either (1.3) terminates or \(\|x_n\| \nrightarrow +\infty\).

**Proof.** Take \(I = \{1, \ldots, m\}\) and \(\iota: n \mapsto n \mod m + 1\). Then (5.2) is satisfied and the assertions follow from any of the following results:

1. Corollary 6.1 with \((\forall i \in I) \; A_i = N_{S_i}\) and \((\forall n \in \mathbb{N}) \; \gamma_n = \lambda_n = 1\).
(2) Corollary 6.5 with \((\forall i \in I) R_i = P_i\) and \((\forall n \in \mathbb{N}) \lambda_n = 1\).

(3) Corollary 6.9 with \((\forall i \in I) f_i = d(\cdot, S_i)\) and \((\forall n \in \mathbb{N}) \lambda_n = 1\).

\[\square\]

**Acknowledgement.** The authors wish to thank Mikhail Solodov for sending them [26].

**References**


