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

[‡]Laboratoire d'Analyse Numérique, Université Pierre et Marie Curie – Paris 6
75005 Paris, France
plc@ann.jussieu.fr

April 7, 1999; revised December 27, 2000

Abstract

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

AMS 1991 subject classification. Primary: 65J15, 47N10; secondary 41A29, 47H05, 47H09, 65K10, 90C25.

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 $\langle \cdot \mid \cdot \rangle$, norm $\|\cdot\|$, and distance d. In 1965, Bregman [8] 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

$$(\forall n \in \mathbb{N}) \;\; x_{n+1} = P_{n \,(\mathrm{mod}\, m)+1} x_n, \tag{1.1}$$

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 [23] (see also [3], [4], [5], and [15]), 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In his unpublished 1968 dissertation [25], Haugazeau proposed independently a strongly convergent variant of (1.1), requiring essentially the same kind of computations.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 $(x, y, z) \in \mathcal{H}^3$

$$H(x,y) = \big\{u \in \mathcal{H} \mid \langle u - y \mid x - y \rangle \leq 0\big\} \tag{1.2}$$

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

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Q\left(x_0, x_n, P_{n \,(\mathrm{mod}\, m)+1} x_n\right) \tag{1.3}$$

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

$$(\forall x \in S)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\| \leq \|x_n - x\|. \tag{1.4}$$

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 [4] and [17] 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.

In many disciplines, including economics [26], image recovery [14], mechanics [18], electromagnetics [18], quantum physics [18], and control theory [20], problems arise in infinite dimensional spaces. In such problems, norm convergence is often much more desirable than weak convergence for it translates the physically tangible property that the energy $\|x_n - x\|^2$ of the error between the iterate x_n and a solution x eventually becomes arbitrarily small. The importance of strong convergence is also underlined in [24], where a convex function f is minimized via the proximal point algorithm: It is shown that the rate of convergence of the value sequence $(f(x_n))_{n\geq 0}$ is better when $(x_n)_{n\geq 0}$ converges strongly than when it converges weakly. Such properties have a direct impact when the algorithm is executed directly in the underlying infinite-dimensional space, as is the case, for instance, in optical signal processing [38]. They are also relevant when the algorithm is implemented in a finite dimensional setting through discretization, since the behavior of certain iterative methods is closely related to that of their discretized counterpart and the number of iterations required by the two methods to converge to within a given tolerance is essentially the same [1], [2].

A question that naturally arises in connection with Fejér-monotone algorithms in infinite dimensional spaces is whether weak convergence can be improved to strong convergence without further assumptions. The following simple example shows that the answer is negative.

Example 1.1 Let $(x_n)_{n\geq 0}$ to be an orthonormal sequence in $\mathcal{H}$. Then $(\forall n \in \mathbb{N})\ \|x_n\| = 1$, and, by Bessel's inequality, $(\forall x \in \mathcal{H})\ \sum_{n\geq 0} |\langle x \mid x_n\rangle|^2 \leq \|x\|^2$. Hence, $(x_n)_{n\geq 0}$ is Fejér-monotone with respect to $S = \{0\}$, $x_n \overset{n}{\not\to} 0$, and $x_n \overset{n}{\rightharpoonup} 0$.

More elaborate constructions of Fejér-monotone methods for which weak convergence holds but strong convergence fails are provided in [21] and [24]. Of course, in specific applications, it is usually possible to achieve strong convergence at the expense of additional restrictions on the constituents of the problem [3], [4], [5], [10], [13], [15], [23], [27], [31], [32], [34], and [36]. 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, Id denotes the identity operator on $\mathcal{H}$ and $\operatorname{Fix} T$ the set of fixed points of an operator T. P_C denotes the projector onto a nonempty closed and convex set C, d_C the distance function to C, N_C its normal cone, $\complement C$ its complement, and 1_C its characteristic function, which takes value 1 on C and 0 on $\complement C$. ∂f denotes the subdifferential of a function $f\colon \mathcal{H} \to \mathbb{R}$ and $\operatorname{lev}_{\leq \mu} f = \{x \in \mathcal{H} \mid f(x) \leq \mu\}$ its lower level set at height $\mu \in \mathbb{R}$. The expressions $x_n \overset{n}{\rightharpoonup} x$ and $x_n \overset{n}{\to} x$ denote, respectively, the weak and strong convergence to x of a sequence $(x_n)_{n\geq 0}$, and $\mathfrak{W}(x_n)_{n\geq 0}$ 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 *[11, 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 $x_n \overset{n}{\rightharpoonup} x \in F \Leftrightarrow \mathfrak{W}(x_n)_{n\geq 0} \subset F$.*

Proof. We provide a short proof for completeness (see also [4], [17]). Fix $z \in F$. By monotonicity, $(\|x_n - z\|^2)_{n\geq 0}$ converges and so does $(\|x_n\|^2 - 2\langle x_n \mid z\rangle)_{n\geq 0}$. Now take z_1 and z_2 in $F \cap \mathfrak{W}(x_n)_{n\geq 0}$. Then it follows that $(\langle x_n \mid z_1 - z_2\rangle)_{n\geq 0}$ converges. Hence $\langle z_1 \mid z_1 - z_2\rangle = \langle z_2 \mid z_1 - z_2\rangle$, i.e., $z_1 = z_2$. Thus, $F \cap \mathfrak{W}(x_n)_{n\geq 0}$ contains at most one point. Since $(x_n)_{n\geq 0}$ is bounded, the assertion is proved. □

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

Definition 2.2 $\mathfrak{T} = \big\{T \colon \mathcal{H} \to \mathcal{H} \mid \operatorname{dom} T = \mathcal{H}$ *and* $(\forall x \in \mathcal{H}) \ \ \operatorname{Fix} T \subset H(x, Tx)\big\}$.

The class $\mathfrak{T}$ is fundamental because (i) it contains several types of operators commonly found in various areas of applied mathematics and in particular in approximation and optimization theory (Proposition 2.3); and (ii) it allows us to completely characterize Fejér-monotone sequences (Proposition 2.7). Before we proceed with specific examples, we need to recall that an operator $T \colon \mathcal{H} \to \mathcal{H}$ with $\operatorname{dom} T = \mathcal{H}$ is *firmly nonexpansive* if

$$(\forall (x,y) \in \mathcal{H}^2) \ \ \langle (T - \operatorname{Id})x - (T - \operatorname{Id})y \mid Tx - Ty\rangle \leq 0 \tag{2.1}$$

or, equivalently, if

$$(\forall (x,y) \in \mathcal{H}^2) \ \ \|Tx - Ty\|^2 \leq \|x - y\|^2 - \|(T - \operatorname{Id})x - (T - \operatorname{Id})y\|^2; \tag{2.2}$$

nonexpansive if

$$(\forall (x,y) \in \mathcal{H}^2) \ \ \|Tx - Ty\| \leq \|x - y\|; \tag{2.3}$$

and *quasi-nonexpansive* if

$$(\forall (x,y) \in \mathcal{H} \times \operatorname{Fix} T) \ \ \|Tx - y\| \leq \|x - y\|. \tag{2.4}$$

Clearly, (2.2) $\Rightarrow$ (2.3) $\Rightarrow$ (2.4).

Proposition 2.3 *Consider the following properties:*

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

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

(iii) $\operatorname{dom} T = \mathcal{H}$ *and T is firmly nonexpansive.*

(iv) *T is a subgradient projector relative to a continuous convex function $f \colon \mathcal{H} \to \mathbb{R}$ such that $\operatorname{lev}_{\leq 0} f \neq \varnothing$, i.e.,*

$$T : x \mapsto \begin{cases} x - \dfrac{f(x)}{\|g(x)\|^2} g(x) & \text{if } f(x) > 0 \\ x & \text{if } f(x) \leq 0, \end{cases} \tag{2.5}$$

where g is a selection of ∂f.

(v) $\operatorname{dom} T = \mathcal{H}$ *and* $2T - \operatorname{Id}$ *is quasi-nonexpansive.*

(vi) $T \in \mathfrak{T}$.

Then:

$$\begin{array}{ccccc} \text{(i)} & \Rightarrow & \text{(ii)} & \Leftrightarrow & \text{(iii)} \\ \Downarrow & & & & \Downarrow \\ \text{(iv)} & \Rightarrow & \text{(v)} & \Leftrightarrow & \text{(vi).} \end{array}$$

Proof. (i) $\Rightarrow$ (ii): Take $A = N_C$ [9, Ex. 2.8.2]. Alternatively, (i) $\Rightarrow$ (iii) is shown in [22, Ch. 12]. (i) $\Rightarrow$ (iv): Take $f = d_C$ [4, Remark 7.6]. (ii) $\Leftrightarrow$ (iii): see [12] (see also [36] for (ii) $\Rightarrow$ (iii)). (iii) $\Rightarrow$ (vi) follows immediately from (2.1). (iv) $\Rightarrow$ (vi): First, observe that T is well defined everywhere, since $\operatorname{dom} \partial f = \mathcal{H}$ and $f(x) > 0 \Rightarrow f(x) > \inf_{y \in \mathcal{H}} f(y) \Rightarrow g(x) \neq 0$. Moreover, $\operatorname{Fix} T = \operatorname{lev}_{\leq 0} f$. Now take $x \in \complement \operatorname{Fix} T$ and $y \in \operatorname{Fix} T$. Then $\langle y - x \mid g(x) \rangle + f(x) \leq f(y) \leq 0$. Consequently, $\langle y - Tx \mid x - Tx \rangle \leq 0$, i.e., $y \in H(x, Tx)$. We conclude $(\forall x \in \mathcal{H}) \;\; \operatorname{Fix} T \subset H(x, Tx)$. (v) $\Leftrightarrow$ (vi): Put $R = 2T - \operatorname{Id}$. Then

$$(\forall (x,y) \in \mathcal{H}^2) \;\; 4\langle y - Tx \mid x - Tx \rangle = \|2(Tx - x) + (x - y)\|^2 - \|x - y\|^2 = \|Rx - y\|^2 - \|x - y\|^2. \tag{2.6}$$

It therefore follows that

$$(\forall (x,y) \in \mathcal{H}^2) \;\; y \in H(x, Tx) \;\; \Leftrightarrow \;\; \|Rx - y\| \leq \|x - y\|. \tag{2.7}$$

Since $\operatorname{Fix} T = \operatorname{Fix} R$, we are done. ☐

The above proposition calls for some remarks.

Remark 2.4 The equivalence (v) $\Leftrightarrow$ (vi) parallels the well-known fact [22, Thm. 12.1]

$$2T - \operatorname{Id} \;\text{ is nonexpansive } \Leftrightarrow \; T \text{ is firmly nonexpansive.} \tag{2.8}$$

However, firmly nonexpansive operators form a proper subset of $\mathfrak{T}$. Thus,

- (iv) $\not\Rightarrow$ (iii): take $\mathcal{H} = \mathbb{R}$, $f \colon x \mapsto \max\{x+1, 2x+1\}$, and $g = 1_{]-\infty,0]} + 2 \cdot 1_{]0,+\infty[}$. Then (2.5) yields
$$(\forall x \in \mathbb{R}) \;\; Tx = \begin{cases} x & \text{if } x \leq -1 \\ -1 & \text{if } -1 < x \leq 0 \\ -1/2 & \text{if } x > 0. \end{cases} \tag{2.9}$$
Consequently the inequality in (2.1) fails for $x = -y = 1/8$.
- (v) $\not\Rightarrow$ (iii): Take $\mathcal{H} = \mathbb{R}$ and $T = (3/4)1_{\mathbb{Q}} \operatorname{Id}$. Then certainly $\operatorname{Fix} T = \{0\}$ and $2T - \operatorname{Id} = (1_{\mathbb{Q}}/2 - 1_{\complement\mathbb{Q}}) \operatorname{Id}$ is quasi-nonexpansive. However, the inequality in (2.1) fails for $x = 1$ and $y = \pi$.

Remark 2.5 Take $\lambda \in \,]0, +\infty[$ and $R \colon \mathcal{H} \to \mathcal{H}$ with $\operatorname{dom} R = \mathcal{H}$. In [30], Măruşter studied the property

$$(\forall (x,y) \in \mathcal{H} \times \operatorname{Fix} R) \;\; \langle y - x \mid Rx - x \rangle \geq \|Rx - x\|^2/\lambda. \tag{2.10}$$

Now suppose $T \colon \mathcal{H} \to \mathcal{H}$ and $\operatorname{dom} T = \mathcal{H}$. Then: $T \in \mathfrak{T} \Leftrightarrow R = \operatorname{Id} + \lambda(T - \operatorname{Id})$ satisfies (2.10). Indeed, $\operatorname{Fix} T = \operatorname{Fix} R$ and, for every $(x,y) \in \mathcal{H} \times \operatorname{Fix} T$, $y \in H(x, Tx) \Leftrightarrow \langle y - Tx \mid Tx - x \rangle \geq 0 \Leftrightarrow \langle y - x \mid Tx - x \rangle \geq \|Tx - x\|^2 \Leftrightarrow \langle y - x \mid Rx - x \rangle \geq \|Rx - x\|^2/\lambda$.

Proposition 2.6 *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) $(\forall \lambda \in [0,1])\ \ \mathrm{Id} + \lambda(T - \mathrm{Id}) \in \mathfrak{T}$.

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 $T' = \mathrm{Id} + \lambda(T - \mathrm{Id})$. Then $\operatorname{dom} T' = \operatorname{dom} T = \mathcal{H}$ and

$$\langle y - T'x \mid x - T'x \rangle = \lambda \langle y - Tx \mid x - Tx \rangle - \lambda(1 - \lambda)\|x - Tx\|^2 \leq \lambda \langle y - Tx \mid x - Tx \rangle \ \leq\ 0. \quad (2.11)$$

Thus, $y \in H(x, T'x)$ and hence $\operatorname{Fix} T' = \operatorname{Fix} T \subset H(x, Tx) \subset H(x, T'x)$, which completes the proof. □

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

Proposition 2.7 *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*

$$(\forall n \in \mathbb{N})\ \ x_{n+1} = 2T_n x_n - x_n, \quad (2.12)$$

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

Proof. Take a sequence $(x_n)_{n \geq 0}$ constructed as in (2.12), where $(T_n)_{n \geq 0}$ lies in $\mathfrak{T}$ 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, by Proposition 2.3, $2T_n - \mathrm{Id}$ is quasi-nonexpansive with fixed point set $\operatorname{Fix} T_n$. Therefore

$$\|x_{n+1} - z\| = \|(2T_n - \mathrm{Id})x_n - z\| \leq \|x_n - z\|, \quad (2.13)$$

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

$$\begin{aligned} H_n &= \left\{ z \in \mathcal{H} \mid \|x_{n+1} - z\|^2 \leq \|x_n - z\|^2 \right\} \\ &= \left\{ z \in \mathcal{H} \;\middle|\; \left\langle z - \frac{x_n + x_{n+1}}{2} \;\middle|\; x_n - x_{n+1} \right\rangle \leq 0 \right\}. \end{aligned} \quad (2.14)$$

Then $T_n x_n = (x_n + x_{n+1})/2$ and the recursion (2.12) holds. Moreover, $\operatorname{dom} T_n = \mathcal{H}$ and $F \subset H_n = \operatorname{Fix} T_n = H(x_n, T_n x_n)$. T_n therefore satisfies the required conditions. □

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

Algorithm 2.8 *Given $\varepsilon \in \,]0,1]$, 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 $T_n \in \mathfrak{T}$ and set $x_{n+1} = x_n + (2 - \varepsilon)(T_n x_n - x_n)$.*

Theorem 2.9 *Let $(x_n)_{n\geq 0}$ be an arbitrary orbit of Algorithm 2.8 and suppose that $F = \bigcap_{n\geq 0} \operatorname{Fix} T_n \neq \emptyset$. Then $(x_n)_{n\geq 0}$ is bounded. Moreover:*

(i) $x_n \overset{n}{\rightharpoonup} x \in F \Leftrightarrow \mathfrak{W}(x_n)_{n\geq 0} \subset F$.

(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. 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

$$\begin{aligned}\|x_{n+1} - z\|^2 &= \|x_n - z\|^2 + 2(2-\varepsilon)\langle z - T_n x_n \mid x_n - T_n x_n\rangle - \varepsilon(2-\varepsilon)\|x_n - T_n x_n\|^2 \\ &\leq \|x_n - z\|^2 - \varepsilon(2-\varepsilon)\|x_n - T_n x_n\|^2 \qquad (2.15)\\ &\leq \|x_n - z\|^2.\end{aligned}$$

Hence $(x_n)_{n\geq 0}$ is Fejér-monotone and, thereby, bounded. (i): Apply Proposition 2.1. (ii): By virtue of (2.15), we get

$$\sum_{n\geq 0} \|x_{n+1} - x_n\|^2 = (2-\varepsilon)^2 \sum_{n\geq 0} \|x_n - T_n x_n\|^2 \leq (2-\varepsilon)\|x_0 - z\|^2/\varepsilon. \qquad (2.16)$$

□

Remark 2.10 The summability properties need not hold for the more general iteration (2.12): For instance, take $x_0 \neq 0$ and $(\forall n \in \mathbb{N})$ $T_n\colon x \mapsto 0$. Then $F = \{0\}$ and (2.12) 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 2.8.

Algorithm 3.1 *At iteration $n \in \mathbb{N}$, suppose that x_n is given and select $T_n \in \mathfrak{T}$. 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 [25], 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)*

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

Step 1. *Select $T_n \in \mathfrak{T}$.*

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

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

$$x_{n+1} = \begin{cases} T_n x_n & \text{if } \rho_n = 0 \text{ and } \pi_n \geq 0, \\ x_0 + (1 + \pi_n/\nu_n)(T_n x_n - x_n) & \text{if } \rho_n > 0 \text{ and } \pi_n \nu_n \geq \rho_n, \\ x_n + \dfrac{\nu_n}{\rho_n}(\pi_n(x_0 - x_n) + \mu_n(T_n x_n - x_n)) & \text{if } \rho_n > 0 \text{ and } \pi_n \nu_n < \rho_n, \end{cases} \tag{3.1}$$

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

From a numerical standpoint, it is important to observe that, given $T_n x_n$, the update equation $x_{n+1} = Q(x_0, x_n, T_n x_n)$ requires only modest computations. Therefore, the bulk of the execution cost of iteration n resides in the determination of $T_n x_n$, just as in Algorithm 2.8.

Some basic properties of Algorithm 3.2 are detailed below.

Definition 3.3 *An* orbit *is a finite or infinite sequence* (x_n) *generated by an algorithm. An infinite orbit is denoted by* $(x_n)_{n \geq 0}$.

Proposition 3.4 *Let* (x_n) *be an arbitrary orbit of Algorithm 3.2. Then:*

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

(ii) *If* x_n *is defined then* $x_n = x_0 \Leftrightarrow x_n = x_{n-1} = \ldots = x_0 \Leftrightarrow x_0 \in \bigcap_{k=0}^{n-1} \operatorname{Fix} T_k$.

(iii) *If* $(x_n)_{n \geq 0}$ *is defined then* $(\|x_0 - x_n\|)_{n \geq 0}$ *is increasing.*

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

$$x_n \neq x_0 \text{ and } (\exists\, \alpha \in]0, +\infty[)\ T_n x_n = \alpha x_0 + (1 - \alpha) x_n. \tag{3.2}$$

(v) $(x_n)_{n \geq 0}$ *is defined if* $F = \bigcap_{n \geq 0} \operatorname{Fix} T_n \neq \varnothing$.

Proof. (i): Let us first recall that the projector onto a nonempty closed convex set $C \subset \mathcal{H}$ is characterized by [22, Ch. 12]

$$(\forall x \in \mathcal{H})\ P_C x \in C \text{ and } C \subset H(x, P_C x). \tag{3.3}$$

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

$$x_{n+1} = x_n = \ldots = x_0 \Leftrightarrow \begin{cases} x_0 \in \bigcap_{k=0}^{n-1} \operatorname{Fix} T_k \\ x_0 = x_{n+1} = Q(x_0, x_0, T_n x_0) = T_n x_0 \end{cases} \Leftrightarrow x_0 \in \bigcap_{k=0}^{n} \operatorname{Fix} T_k. \tag{3.4}$$

(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 $F \subset \bigcap_{n\geq 0} H(x_0, x_n) \cap H(x_n, T_n x_n)$, i.e., since $(T_n)_{n\geq 0} \subset \mathfrak{T}$, $F \subset \bigcap_{n\geq 0} H(x_0, x_n)$. For $n = 0$, it is clear that $F \subset H(x_0, x_n) = \mathcal{H}$. Furthermore, for every $n \in \mathbb{N}$, it results from (3.3) and Proposition 2.6(i) that

$$\begin{aligned} F \subset H(x_0, x_n) &\Rightarrow F \subset H(x_0, x_n) \cap H(x_n, T_n x_n) \\ &\Rightarrow F \subset H(x_0, Q(x_0, x_n, T_n x_n)) \\ &\Rightarrow F \subset H(x_0, x_{n+1}), \end{aligned} \tag{3.5}$$

which establishes the assertion by induction. □

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

Theorem 3.5 *Let (x_n) be an arbitrary orbit of Algorithm 3.2 and let $F = \bigcap_{n\geq 0} \operatorname{Fix} T_n$. Then:*

(i) *If $(x_n)_{n\geq 0}$ is defined then: $(x_n)_{n\geq 0}$ is bounded $\Leftrightarrow$ $(\|x_0 - x_n\|)_{n\geq 0}$ converges.*

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

(iii) *If $F \neq \emptyset$, 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 $F \neq \emptyset$, then: $x_n \overset{n}{\to} P_F x_0 \Leftrightarrow \mathfrak{W}(x_n)_{n\geq 0} \subset F$.*

(v) *If $(x_n)_{n\geq 0}$ is defined and 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 3.4(i). (ii): By definition, for every $n \in \mathbb{N}$, x_n is the projection of x_0 onto $H(x_0, x_n)$. On the other hand, as shown in the proof of Proposition 3.4(v), $F \subset \bigcap_{n\geq 0} H(x_0, x_n)$. Hence,

$$(\forall n \in \mathbb{N})\ \ \|x_0 - x_n\| \leq \|x_0 - P_F x_0\| \tag{3.6}$$

and the claim follows at once. (iii) follows from (ii), (i), and the previous inequality. (iv): The forward implication is trivial. For the reverse implication, assume $\mathfrak{W}(x_n)_{n\geq 0} \subset F$ and fix $x \in \mathfrak{W}(x_n)_{n\geq 0}$, say $x_{n_k} \overset{k}{\rightharpoonup} x$. 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

$$\|x_0 - x\| \leq \underline{\lim}_k \|x_0 - x_{n_k}\| = \lim_n \|x_0 - x_n\| \leq \|x_0 - P_F x_0\|. \tag{3.7}$$

Consequently, since $x \in F$, $x = P_F x_0$ and, in turn, $\mathfrak{W}(x_n)_{n\geq 0} = \{P_F x_0\}$. Next, since $(x_n)_{n\geq 0}$ is bounded, we obtain $x_n \overset{n}{\rightharpoonup} P_F x_0$. 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 $\|x_0 - x_n\| \overset{n}{\to} \|x_0 - P_F x_0\|$. 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 $x_n \overset{n}{\to} P_F x_0$. (v): For every $n \in \mathbb{N}$, the inclusion $x_{n+1} \in H(x_0, x_n)$ implies

$$\begin{aligned} \|x_0 - x_{n+1}\|^2 - \|x_0 - x_n\|^2 &= \|x_{n+1} - x_n\|^2 + 2\langle x_{n+1} - x_n \mid x_n - x_0\rangle \\ &\geq \|x_{n+1} - x_n\|^2. \end{aligned} \tag{3.10}$$

Hence $\sum_{n\geq 0} \|x_{n+1} - x_n\|^2 \leq \sup_{n\geq 0} \|x_0 - x_n\|^2 < +\infty$ since $(x_n)_{n\geq 0}$ is bounded by assumption. In turn, since for every $n \in \mathbb{N}$ the inclusion $x_{n+1} \in H(x_n, T_n x_n)$ implies

$$\begin{aligned} \|x_{n+1} - x_n\|^2 &= \|x_{n+1} - T_n x_n\|^2 - 2\langle x_{n+1} - T_n x_n \mid x_n - T_n x_n\rangle + \|x_n - T_n x_n\|^2 \\ &\geq \|x_{n+1} - T_n x_n\|^2 + \|x_n - T_n x_n\|^2, \end{aligned} \tag{3.11}$$

we obtain $\sum_{n\geq 0} \|x_n - T_n x_n\|^2 < +\infty$. ☐

4 The weak-to-strong convergence principle

To achieve convergence in Algorithms 2.8 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 $(T_n)_{n\geq 0} \subset \mathfrak{T}$ is* coherent *if for every bounded sequence $(y_n)_{n\geq 0}$ in $\mathcal{H}$ there holds*

$$\begin{cases} \sum_{n\geq 0} \|y_{n+1} - y_n\|^2 < +\infty \\ \sum_{n\geq 0} \|y_n - T_n y_n\|^2 < +\infty \end{cases} \quad \Rightarrow \quad \mathfrak{W}(y_n)_{n\geq 0} \subset \bigcap_{n\geq 0} \operatorname{Fix} T_n. \tag{4.1}$$

This property allows us to view the convergence of Algorithms 2.8 and 3.2 from a single perspective.

Theorem 4.2 *Suppose that $(T_n)_{n\geq 0}$ is coherent and let $F = \bigcap_{n\geq 0} \operatorname{Fix} T_n$. Then:*

(i) *If $F \neq \varnothing$, then every orbit of Algorithm 2.8 converges weakly to a point in F.*

(ii) *(Trichotomy) For an arbitrary orbit (x_n) of Algorithm 3.2, exactly one of the following alternatives holds:*

  (a) *$F \neq \varnothing$ and $x_n \overset{n}{\to} P_F x_0$.*

  (b) *$F = \varnothing$ and $\|x_n\| \overset{n}{\to} +\infty$.*

  (c) *$F = \varnothing$ and the algorithm terminates.*

Proof. (i) follows from Theorem 2.9. (ii): If $F \neq \varnothing$, then it follows from items (ii) and (iv)-(v) in Theorem 3.5 that $x_n \overset{n}{\rightarrow} P_F x_0$, whence (a). Thus, we now assume $F = \varnothing$. It remains to prove (b), in which $F = \varnothing$ and $(x_n)_{n\geq 0}$ is defined. Suppose that $\|x_n\| \overset{n}{\not\rightarrow} +\infty$. Then it follows from Proposition 3.4(iii) that $(x_n)_{n\geq 0}$ is bounded, and then from Theorem 3.5(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 $\varnothing \neq \mathfrak{W}(x_n)_{n\geq 0} \subset \bigcap_{n\geq 0} \operatorname{Fix} T_n = F = \varnothing$, which is absurd. □

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 2.8 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 2.8 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 elementary 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:

(i) $T_n = P_{H_n}$, where $H_n = \{x \in \mathcal{H} \mid \langle x - (1-2^{-n-1})z \mid z\rangle \geq 0\} = \operatorname{Fix} T_n$. Then $(\forall n \in \mathbb{N})\ \ x_n = (1-2^{-n})z$ and $F = \{x \in \mathcal{H} \mid \langle x - z \mid z\rangle \geq 0\} \neq \varnothing$. Hence $x_n \overset{n}{\rightarrow} P_F x_0 = z$.

(ii) $T_n \colon x \mapsto x + z$. Then $F = \varnothing$ and $(\forall n \in \mathbb{N})\ \ x_n = nz$.

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

Moreover in each case $(T_n)_{n\geq 0}$ is coherent.

We conclude this section with a useful fact.

Proposition 4.5 *Fix $\delta \in \,]0,1]$, $(T_n)_{n\geq 0}$ in $\mathfrak{T}$, and define*

$$(\forall n \in \mathbb{N})\ \ T_n' = \operatorname{Id} + \lambda_n (T_n - \operatorname{Id}) \ \ \textit{where}\ \ \lambda_n \in [\delta, 1]. \tag{4.2}$$

Then $(T_n')_{n\geq 0}$ is coherent if and only if $(T_n)_{n\geq 0}$ is coherent.

Proof. $(T_n')_{n\geq 0} \subset \mathfrak{T}$ by Proposition 2.6(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 $(\forall n \in \mathbb{N})\ \ \operatorname{Fix} T_n' = \operatorname{Fix} T_n$. □

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 acted upon at each iteration [28]. In this section, we present one such implementation of Algorithms 2.8 and 3.2 for solving the convex feasibility problem

$$\text{Find } x \in S = \bigcap_{i \in I} S_i, \tag{5.1}$$

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 [16] from a distinct viewpoint.

Assumption 5.1 $(\forall n \in \mathbb{N})$ $T_n \in \mathfrak{T}$ *and* $\operatorname{Fix} T_n = S_{\mathrm{i}(n)}$, *where*

(i) *The index control mapping* $\mathrm{i}\colon \mathbb{N} \to I$ *satisfies*

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

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

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

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

Example 5.2 Take $I = \mathbb{N}$. Then the mapping i defined by $(\forall i \in I)(\forall k \in \mathbb{N})$ $\mathrm{i}(2^i(2k+1) - 1) = i$ 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 2.8 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*

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

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

(ii) *Every orbit of the algorithm*

$$x_{n+1} = Q\big(x_0, x_n, x_n + \lambda_n(T_n x_n - x_n)\big) \quad \textit{where} \quad \lambda_n \in [\varepsilon, 1] \tag{5.5}$$

converges strongly to $P_S x_0$ *if* $S \neq \varnothing$*; if* $S = \varnothing$*, either* (5.5) *terminates or* $\|x_n\| \overset{n}{\to} +\infty$.

Proof. Let us first establish that $(T_n)_{n\geq 0}$ is coherent, i.e., that (4.1) holds. To this end, take a bounded sequence $(y_n)_{n\geq 0} \subset \mathcal{H}$ such that $\sum_{n\geq 0} \|y_{n+1} - y_n\|^2 < +\infty$ and $\sum_{n\geq 0} \|y_n - T_n y_n\|^2 < +\infty$. Next, fix $i \in I$ and $y \in \mathfrak{W}(y_n)_{n\geq 0}$, say $y_{n_k} \overset{k}{\rightharpoonup} y$. Then, since by Assumption 5.1(i)

$$\bigcap_{n\geq 0} \operatorname{Fix} T_n = \bigcap_{n\geq 0} S_{\mathrm{i}(n)} = \bigcap_{i\in I} S_i = S, \tag{5.6}$$

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

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

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

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

Consequently

$$\sum_{n\geq 0} \|y_{n+1} - y_n\|^2 < +\infty \quad \Rightarrow \quad y_{p_k} - y_{n_k} \overset{k}{\to} 0 \quad \Rightarrow \quad y_{p_k} \overset{k}{\rightharpoonup} y. \tag{5.9}$$

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 $(\lambda'_n)_{n\geq 0}$ in $[\varepsilon/(2-\varepsilon), 1]$, $(\mathrm{Id} + \lambda'_n(T_n - \mathrm{Id}))_{n\geq 0}$ is also coherent by virtue of Proposition 4.5. It therefore follows from Theorem 4.2(i) that every orbit of the algorithm

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

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

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 [4], [17], and [27].

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.,

$$\text{Find } x \in \mathcal{H} \text{ such that } (\forall i \in I)\ \ 0 \in A_i x. \tag{6.1}$$

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, $A_{i,\gamma} = (\mathrm{Id} - (\mathrm{Id} + \gamma A_i)^{-1})/\gamma$ denotes the Yosida approximation of A_i of index $\gamma \in]0,+\infty[$, $\operatorname{ran} A_i = \{w \in \mathcal{H} \mid (\exists\, v \in \mathcal{H})\ w \in A_i v\}$ its range, and $\operatorname{gr} A_i = \{(v,w) \in \mathcal{H}^2 \mid w \in A_i v\}$ its graph. We shall exploit the fact that $\operatorname{gr} A_i$ is weakly-strongly closed [9, Prop. 2.5]:

$$(\forall ((v_n,w_n))_{n\geq 0} \subset \operatorname{gr} A_i)\ \ \begin{cases} v_n \overset{n}{\rightharpoonup} v \\ w_n \overset{n}{\to} w \end{cases} \quad \Rightarrow \quad (v,w) \in \operatorname{gr} A_i. \tag{6.2}$$

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

(i) *Every orbit of the algorithm*

$$x_{n+1} = x_n - \gamma_n \lambda_n A_{\mathrm{i}(n),\gamma_n} x_n \ \ \textit{where}\ \ \lambda_n \in [\varepsilon, 2-\varepsilon] \tag{6.3}$$

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

(ii) *Every orbit of the algorithm*

$$x_{n+1} = Q\left(x_0, x_n, x_n - \gamma_n \lambda_n A_{\mathrm{i}(n),\gamma_n} x_n\right) \ \ \textit{where}\ \ \lambda_n \in [\varepsilon, 1] \tag{6.4}$$

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

Proof. Let $(T_n)_{n\geq 0} = (\mathrm{Id} - \gamma_n A_{\mathrm{i}(n),\gamma_n})_{n\geq 0}$. Then, for every $n \in \mathbb{N}$, $\operatorname{Fix} T_n = A_{\mathrm{i}(n)}^{-1}0 = S_{\mathrm{i}(n)}$ and $T_n \in \mathfrak{T}$ by Proposition 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 $y_{n_k} \overset{k}{\rightharpoonup} y$ and $(\forall k \in \mathbb{N})\ \ \mathrm{i}(n_k) = i$. 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 $\sum_{n\geq 0} \gamma_n^2 \|A_{\mathrm{i}(n),\gamma_n} y_n\|^2 < +\infty$. Then $v_k - y_{n_k} \overset{k}{\to} 0$ and therefore $v_k \overset{k}{\rightharpoonup} y$. Moreover, it follows from the assumption $\inf_{k\geq 0} \gamma_{n_k} > 0$ that $A_{i,\gamma_{n_k}} y_{n_k} \overset{k}{\to} 0$. To sum up,

$$\begin{cases} ((v_k, A_{i,\gamma_{n_k}} y_{n_k}))_{k\geq 0} \text{ is in } \operatorname{gr} A_i \\ v_k \overset{k}{\rightharpoonup} y \\ A_{i,\gamma_{n_k}} y_{n_k} \overset{k}{\to} 0. \end{cases} \tag{6.5}$$

Therefore (6.2) implies $0 \in A_i y$ and (5.3) ensues. $\square$

As discussed in [36], a special case of (6.1) of great interest is the problem of finding a zero of a single maximal monotone operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$. 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 $(\gamma_n)_{n\geq 0}$ in $]0,+\infty[$ such that $\inf_{n\geq 0} \gamma_n > 0$. Then:*

(i) *Every orbit of the algorithm*

$$x_{n+1} = (\mathrm{Id} + \gamma_n A)^{-1} x_n \tag{6.6}$$

converges weakly to a zero of A if $0 \in \operatorname{ran} A$ [36, Thm. 1]; if $0 \notin \operatorname{ran} A$ then $\|x_n\| \overset{n}{\to} +\infty$ [35, Thm. 3].

(ii) *Every orbit of the algorithm*

$$x_{n+1} = Q\left(x_0, x_n, (\mathrm{Id} + \gamma_n A)^{-1} x_n\right) \tag{6.7}$$

converges strongly to the projection of x_0 onto $A^{-1}0$ if $0 \in \operatorname{ran} A$; if $0 \notin \operatorname{ran} A$ then $\|x_n\| \overset{n}{\to} +\infty$.

Proof. In light of Corollary 6.1, it is enough to assume $0 \notin \operatorname{ran} A$ and to show that (i) $\|x_n\| \overset{n}{\to} +\infty$ in (6.6); and (ii) (6.7) does not terminate. The former result is already known [35, Thm. 3]. As to the latter, our proof is patterned after that of [37, Thm. 2] and is based on a truncation argument found in [36]. 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

$$(\forall k \in \{0,\ldots,n\})\;\; v_k = (\mathrm{Id} + \gamma_k A)^{-1} x_k \;\text{ and }\; A' = A + N_C,$$
$$\text{where } C = \left\{ v \in \mathcal{H} \;\middle|\; \|v\| \leq 1 + \max_{0\leq k\leq n} \|v_k\| \right\}. \tag{6.8}$$

Recall that N_C is maximal monotone with $\operatorname{dom} N_C = C$ and that $(\forall v \in \operatorname{int}(C))\; N_C v = \{0\}$ [9, Ex. 2.8.2]. Thus, since by construction $(v_k)_{0\leq k\leq n}$ is in $\operatorname{dom} A \cap \operatorname{int}(C)$, we deduce on the one hand that

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

and, on the other hand, that A' is maximal monotone by virtue of Rockafellar's sum theorem [9, Cor. 2.7]. Consequently, its resolvents are single-valued [9, Prop. 2.2] and we derive from (6.9) that $(\forall k \in \{0,\ldots,n\})\; v_k = (\mathrm{Id} + \gamma_k A')^{-1} x_k$. Therefore, up to iteration n, replacing A by A' does not affect the behavior of Algorithm (6.7). However, since $\operatorname{dom} A' \subset C$ is bounded, A' is surjective [9, Cor. 2.2] and it therefore has zeros. Hence, arguing as in the proof of Proposition 3.4(v), we obtain $\varnothing \neq (A')^{-1}0 \subset \bigcap_{k=0}^{n} H(x_0, x_k) \cap H(x_k, v_k)$. We conclude that x_{n+1} is well defined. $\square$

Remark 6.3 When $0 \in \operatorname{ran} A$, 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 $\prod_{n\geq 0} (\mathrm{Id} + \gamma_n A)^{-1} x_0$. Such results go back to the prox-regularization method of [29] (see also [10] and [36] for extensions to inexact iterations and strong convergence conditions). An instance of Corollary 6.2(i) in which strong convergence fails when $0 \in \operatorname{ran} A$ is constructed in [24, Cor. 5.1].

Remark 6.4 A result closely related to Corollary 6.2(ii) was obtained via a different analysis in [37]. 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 $y_n = (\mathrm{Id} + \gamma_n A)^{-1}(x_n + u_n)$ 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

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

Recall that if R_i is nonexpansive then $\mathrm{Id} - R_i$ is maximal monotone and therefore demiclosed [11, Lem. 4]:

$$(\forall (y_n)_{n \geq 0} \subset \mathcal{H}) \;\; \begin{cases} y_n \overset{n}{\rightharpoonup} y \\ y_n - R_i y_n \overset{n}{\to} w \end{cases} \quad \Rightarrow \quad w = y - R_i y. \tag{6.11}$$

Since the fixed point set of a nonexpansive operator is closed and convex [22, 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 $(\forall i \in I) \;\; S_i = \mathrm{Fix}\, R_i$.

Corollary 6.5 *Fix $x_0 \in \mathcal{H}$ and $\varepsilon \in]0,1]$ and suppose that $\mathrm{i}\colon \mathbb{N} \to I$ satisfies Condition (5.2). Then:*

(i) *Every orbit of the algorithm*

$$x_{n+1} = x_n + \lambda_n (R_{\mathrm{i}(n)} x_n - x_n) \;\; \text{where} \;\; \lambda_n \in [\varepsilon, 1 - \varepsilon] \tag{6.12}$$

converges weakly to a point in S if $S \neq \varnothing$ [11, Thm. 5].

(ii) *Every orbit of the algorithm*

$$x_{n+1} = Q\left(x_0, x_n, x_n + \lambda_n (R_{\mathrm{i}(n)} x_n - x_n)\right) \;\; \text{where} \;\; \lambda_n \in [\varepsilon, 1/2] \tag{6.13}$$

converges strongly to $P_S x_0$ if $S \neq \varnothing$; if $S = \varnothing$, either (6.13) terminates or $\|x_n\| \overset{n}{\to} +\infty$.

Proof. Let $(T_n)_{n \geq 0} = ((R_{\mathrm{i}(n)} + \mathrm{Id})/2)_{n \geq 0}$. Then, for every $n \in \mathbb{N}$, $\mathrm{Fix}\, T_n = \mathrm{Fix}\, R_{\mathrm{i}(n)} = S_{\mathrm{i}(n)}$ and T_n is firmly nonexpansive by (2.8). Hence, $T_n \in \mathfrak{T}$ by Proposition 2.3(iii) 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. ☐

In the case of a single nonexpansive operator $R\colon \mathcal{H} \to \mathcal{H}$ 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*

$$x_{n+1} = x_n + \lambda_n (Rx_n - x_n) \quad \textit{where} \quad \lambda_n \in [\varepsilon, 1-\varepsilon] \tag{6.14}$$

converges weakly to a point in $\operatorname{Fix} R$ *if* $\operatorname{Fix} R \neq \varnothing$ *[19, Thm. 8]; if* $\operatorname{Fix} R = \varnothing$ *then* $\|x_n\| \overset{n}{\to} +\infty$ *[7, Cor. 9(b)].*

(ii) *Every orbit of the algorithm*

$$x_{n+1} = Q\big(x_0, x_n, x_n + \lambda_n (Rx_n - x_n)\big) \quad \textit{where} \quad \lambda_n \in [\varepsilon, 1/2] \tag{6.15}$$

converges strongly to the projection of x_0 *onto* $\operatorname{Fix} R$ *if* $\operatorname{Fix} R \neq \varnothing$*; if* $\operatorname{Fix} R = \varnothing$*, then* $\|x_n\| \overset{n}{\to} +\infty$.

Proof. This is an application of Corollary 6.5, and we need only assume that $\operatorname{Fix} R = \varnothing$ and show that (i) $\|x_n\| \overset{n}{\to} +\infty$ in (6.14); and (ii) (6.15) does not terminate. The first result follows from [7, Cor. 9(b)]. To establish the second, suppose, as is true for $n = 1$, that the iterates $(x_k)_{0 \le k \le n}$ are well defined for some $n > 0$. Let

$$R' = P_C \circ R, \quad \text{where} \quad C = \left\{ x \in \mathcal{H} \;\middle|\; \|x\| \le \max_{0 \le k \le n} \|Rx_k\| \right\}. \tag{6.16}$$

Then $(\forall k \in \{0, \ldots, n\})$ $R'x_k = Rx_k$ 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 [22, Thm. 4.1] that $\operatorname{Fix} R' \neq \varnothing$. Hence, we can argue as in the proof of Proposition 3.4(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. ☐

Remark 6.7 Alternatively, Corollary 6.6 can be derived from Corollary 6.2 by using the equivalence between (ii) and (iii) in Proposition 2.3. However, our present proof is more self-contained and relies only on the Browder-Göhde-Kirk theorem.

Remark 6.8 In the case $\operatorname{Fix} R \neq \varnothing$, an instance of (6.14) in which strong convergence fails is constructed in [21].

Remark 6.9 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 $\operatorname{Fix} R' = \operatorname{Fix} R$. 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 $(\operatorname{lev}_{\le 0} f_i)_{i \in I}$ are nonempty. Under consideration is the problem

$$\text{Find } x \in \mathcal{H} \ \text{ such that } \sup_{i \in I} f_i(x) \le 0. \tag{6.17}$$

Upon calling S its set of solutions and setting $(\forall i \in I)\ S_i = \operatorname{lev}_{\leq 0} f_i$, (6.17) is seen to fit into (5.1).

Subsequently, for every $i \in I$ the operator G_i is defined to be

$$G_i : x \mapsto \begin{cases} x - \dfrac{f_i(x)}{\|g_i(x)\|^2} g_i(x), & \text{if } f_i(x) > 0 \\ x & \text{if } f_i(x) \leq 0, \end{cases} \tag{6.18}$$

where g_i is a selection of ∂f_i. We shall say that the subdifferential ∂f_i is bounded if it maps bounded sets to bounded sets (see [6] for a discussion of this property).

Corollary 6.10 *Fix $x_0 \in \mathcal{H}$ and $\varepsilon \in]0,1]$. Suppose that $\mathrm{i} \colon \mathbb{N} \to I$ satisfies Condition (5.2) and that the subdifferentials $(\partial f_i)_{i \in I}$ are bounded. Then:*

(i) *Every orbit of the algorithm*

$$x_{n+1} = x_n + \lambda_n (G_{\mathrm{i}(n)} x_n - x_n) \quad \text{where} \quad \lambda_n \in [\varepsilon, 2-\varepsilon] \tag{6.19}$$

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

(ii) *Every orbit of the algorithm*

$$x_{n+1} = Q\left(x_0, x_n, x_n + \lambda_n (G_{\mathrm{i}(n)} x_n - x_n)\right) \quad \text{where} \quad \lambda_n \in [\varepsilon, 1] \tag{6.20}$$

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

Proof. Let $(T_n)_{n\geq 0} = (G_{\mathrm{i}(n)})_{n \geq 0}$. Then, for every $n \in \mathbb{N}$, $\operatorname{Fix} T_n = \operatorname{lev}_{\leq 0} f_{\mathrm{i}(n)} = S_{\mathrm{i}(n)}$ and $T_n \in \mathfrak{T}$ by Proposition 2.3(iv). Hence, 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 $\sum_{n \geq 0} \|y_n - G_{\mathrm{i}(n)} y_n\|^2 < +\infty$ and containing a subsequence $(y_{n_k})_{k\geq 0}$ such that $y_{n_k} \overset{k}{\rightharpoonup} y$ and $(\forall k \in \mathbb{N})\ \mathrm{i}(n_k) = i$. Then we must show $f_i(y) \leq 0$. Since f_i is weak lower semicontinuous, $f_i(y) \leq \underline{\lim}_k f_i(y_{n_k})$. Passing to a subsequence if necessary, we can assume $(y_{n_k})_{k \geq 0} \subset \complement S_i$ (otherwise the conclusion is immediate). However, since ∂f_i is bounded, $y_{n_k} - G_i y_{n_k} \overset{k}{\to} 0 \Rightarrow f_i(y_{n_k})/\|g_i(y_{n_k})\| \overset{k}{\to} 0 \Rightarrow f_i(y_{n_k}) \overset{k}{\to} 0$. Therefore $f_i(y) \leq 0$. ☐

The above results are applicable to the approximate minimization of a continuous convex function $f \colon \mathcal{H} \to \mathbb{R}$ over a nonempty closed convex set $C \subset \mathcal{H}$. Let us fix $\mu \in \mathbb{R}$ such that $\operatorname{lev}_{\leq \mu} f \neq \emptyset$ and define the approximate solution set as $S = C \cap \operatorname{lev}_{\leq \mu} f$. Then the problem is a special instance of (6.17) with $I = \{1,2\}$, $f_1 = f - \mu$, and $f_2 = d_C$. Furthermore, since $(\forall x \in \complement C)$ $\partial d_C(x) = \{(x - P_C x)/d_C(x)\}$, (6.18) gives $G_2 = P_C$ and Corollary 6.10 with

$$\lambda_{2n} = \alpha_n, \lambda_{2n+1} = 1, \mathrm{i}(2n) = 1, \text{ and } \mathrm{i}(2n+1) = 2 \tag{6.21}$$

yields

Corollary 6.11 *Fix $x_0 \in \mathcal{H}$ and $\varepsilon \in]0,1]$. Suppose that $\overline{\mu} \triangleq \inf_{x\in C} f(x) > -\infty$ and that ∂f is bounded. Then:*

(i) *Every orbit of the algorithm*

$$x_{n+1} = P_C\big(x_n + \alpha_n(G_1 x_n - x_n)\big) \quad \text{where} \quad \alpha_n \in [\varepsilon, 2-\varepsilon] \tag{6.22}$$

converges weakly to a point in S if $S \neq \varnothing$, i.e, if $\mu > \overline{\mu}$ or if f has a minimizer on C and $\mu = \overline{\mu}$.

(ii) *Every orbit of the algorithm*

$$x_{n+1} = Q\left(x_0, z_n, P_C z_n\right), \quad \text{where} \quad z_n = Q\big(x_0, x_n, x_n + \alpha_n(G_1 x_n - x_n)\big) \text{ and } \alpha_n \in [\varepsilon, 1] \tag{6.23}$$

converges strongly to $P_S x_0$ if $S \neq \varnothing$; if $S = \varnothing$, either (6.23) terminates or $\|x_n\| \overset{n}{\to} +\infty$.

Now, suppose that $\inf f(\mathcal{H}) < \min f(C)$ and that ∂f maps the bounded subsets of C to bounded sets. Then we deduce from Corollary 6.11(i) that every orbit of the algorithm

$$\begin{cases} x_0 \in C \\ \varepsilon \in]0,1] \end{cases} \text{and } x_{n+1} = P_C\left(x_n + \alpha_n \frac{\overline{\mu} - f(x_n)}{\|t_n\|^2} t_n\right) \text{ where } \begin{cases} t_n \in \partial f(x_n) \\ \alpha_n \in [\varepsilon, 2-\varepsilon] \end{cases} \tag{6.24}$$

converges weakly to a minimizer of f over C. This classical result is due to Polyak [33, Thm. 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 6.10.

Corollary 6.12 *Let $(S_i)_{1\leq i\leq m}$ be nonempty closed convex subsets of $\mathcal{H}$ with intersection S and fix $x_0 \in \mathcal{H}$. Then:*

(i) *The orbit of Algorithm (1.1) converges weakly to a point in S if $S \neq \varnothing$ [8, Thm. 1].*

(ii) *The orbit of Algorithm (1.3) converges strongly to $P_S x_0$ if $S \neq \varnothing$ [25, Thm. 3-2]; if $S = \varnothing$, either (1.3) terminates or $\|x_n\| \overset{n}{\to} +\infty$.*

Proof. Take $I = \{1, \ldots, m\}$ and i: $n \mapsto n \pmod m + 1$. 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 6.10 with $(\forall i \in I)\ f_i = d_{S_i}$ and $(\forall n \in \mathbb{N})\ \lambda_n = 1$.

□

Acknowledgements. The authors thank Mikhail Solodov for sending them [37] and the two anonymous referees for their comments. The work of P. L. Combettes was supported by the National Science Foundation under grant MIP-9705504.

References

[1] Allgower, E.L., K. Böhmer, F.A. Potra, W.C. Rheinboldt (1986). A mesh-independence principle for operator equations and their discretizations" *SIAM J. Numer. Anal.* **23** 160–169.

[2] Argyros, I.K. (1997). The asymptotic mesh independence principle for inexact Newton-Galerkin-like methods. *Pure Math. Appl.* **8** 169–194.

[3] Bauschke, H.H. (1995). A norm convergence result on random products of relaxed projections in Hilbert space. *Trans. Amer. Math. Soc.* **347** 1365–1373.

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

[5] Bauschke, H.H., J.M. Borwein, A.S. Lewis (1997). The method of cyclic projections for closed convex sets in Hilbert space. *Contemp. Math.* **204** 1–38.

[6] Borwein, J., S. Fitzpatrick, J. Vanderwerff (1994). Examples of convex functions and classifications of Banach spaces. *J. Convex Anal.* **1** 61–73.

[7] Borwein, J., S. Reich, I. Shafrir (1992). Krasnoselski-Mann iterations in normed spaces. *Canad. Math. Bull.* **35** pp 21–28.

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

[9] Brézis, H. (1973). *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert.* North-Holland/Elsevier, New York.

[10] Brézis, H., P.L. Lions (1978). Produits infinis de résolvantes. *Israel J. Math.* **29** 329–345.

[11] Browder, F.E. (1967). Convergence theorems for sequences of nonlinear operators in Banach spaces. *Math. Z.* **100** 201–225.

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

[13] Combettes, P.L. (1995). Construction d'un point fixe commun à une famille de contractions fermes. *C. R. Acad. Sci. Paris Sér. I Math.* **320** 1385–1390.

[14] Combettes, P.L. (1996). The convex feasibility problem in image recovery. P. Hawkes, ed., *Advances in Imaging and Electron Physics* **95**, Academic Press, New York, 155–270.

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

[16] Combettes, P.L. (2000). Strong convergence of block-iterative outer approximation methods for convex optimization. *SIAM J. Control Optim.* **38** 538–565.

[17] Combettes, P.L. (2001). Fejér monotonicity in convex optimization. C.A. Floudas and P.M. Pardalos, eds., *Encyclopedia of Optimization,* Kluwer, Boston, MA.

[18] Dautray, R., J.L. Lions (1988-1993). *Mathematical Analysis and Numerical Methods for Science and Technology* vols. 1–6. Springer-Verlag, New York.

[19] Dotson, W.G. (1970). On the Mann iterative process. *Trans. Amer. Math. Soc.* **149** 65–73.

[20] Fattorini, H.O. (1999). *Infinite–Dimensional Optimization and Control Theory,* Cambridge University Press, Cambridge.

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

[22] Goebel, K., W. A. Kirk (1990). *Topics in Metric Fixed Point Theory,* Cambridge University Press, Cambridge.

[23] Gubin, L.G., B.T. Polyak, E.V. Raik (1967). The method of projections for finding the common point of convex sets. *USSR Comput. Math. Math. Phys.* **7** 1–24.

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

[25] Haugazeau, Y. (1968). *Sur les Inéquations Variationnelles et la Minimisation de Fonctionnelles Convexes.* Thèse, Université de Paris.

[26] Khan, M.A., N.C. Yannelis (Eds.) (1991). *Equilibrium Theory in Infinite Dimensional Spaces,* Springer-Verlag, New York.

[27] Kiwiel, K.C., B. Łopuch (1997). Surrogate projection methods for finding fixed points of firmly nonexpansive mappings. *SIAM J. Optim.* **7** 1084–1102.

[28] Lions, J.L., R. Temam (1966). Une méthode d'éclatement des opérateurs et des contraintes en calcul des variations. *C. R. Acad. Sci. Paris Sér. A Math.* **263** 563–565.

[29] Martinet, B. (1970). Régularisation d'inéquations variationnelles par approximations successives. *Rev. Française Informat. Recherche Opérationnelle* **4** 154–158.

[30] Măruşter, Şt. (1977). The solution by iteration of nonlinear equations in Hilbert spaces. *Proc. Amer. Math. Soc.* **63** 69–73.

[31] Moreau, J.J. (1978). Un cas de convergence des itérées d'une contraction d'un espace hilbertien. *C. R. Acad. Sci. Paris Sér. A Math.* **286** 143–144.

[32] Petryshyn, W.V., T.E. Williamson (1973). Strong and weak convergence of the sequence of successive approximations for quasi-nonexpansive mappings. *J. Math. Anal. Appl.* **43** 459–497.

[33] Polyak, B.T. (1969). Minimization of unsmooth functionals. *USSR Comput. Math. Math. Phys.* **9** 14–29.

[34] Raik, E. (1969). A class of iterative methods with Fejér-monotone sequences. *Eesti NSV Tead. Akad. Toimetised Füüs.-Mat.* **18** 22–26.

[35] Reich, S. (1977). On infinite products of resolvents. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* ser. VIII, **63** 338–340.

[36] Rockafellar, R.T. (1976). Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.* **14** 877–898.

[37] Solodov, M.V., B.F. Svaiter (2000). Forcing strong convergence of proximal point iterations in a Hilbert space. *Math. Programming* **87** 189–202.

[38] Vanderlugt, A. (1992) *Optical Signal Processing,* Wiley, New York.