Publ. Math. Debrecen
58 / 1-2 (2001), 29–48

Improving the speed of convergence in the method of projections onto convex sets

By G. CROMBEZ (Gent)

Abstract. A serious drawback of the method of projections onto convex sets to find a point in the intersection of a finite number of closed convex sets in an Euclidean space, is its often very slow convergence. This bad behaviour, sometimes called the "tunneling effect", seems a.o. to be connected with the monotone behaviour of the usual algorithms. We present a new algorithm that can interrupt at different steps this monotone behaviour; this can substantially improve the speed of convergence.

1. Introduction

Finding a point in the intersection of a finite number of closed convex sets $\{C_j\}_{j=1}^r$ in an m-dimensional Euclidean space $\mathbb{R}^m$ ($\mathbb{R}$ the set of reals), with $C^* \equiv \bigcap_{j=1}^r C_j \neq \emptyset$, is a problem that often arises in applied mathematics; we refer to [7] and [9] for a short overview of some general applications, and to [15] and [16] for specific applications in image processing. The stated problem is also known as "solving the convex feasibility problem".

The method of projections onto convex sets (abbreviated as POCS), seems to be very well suited to solve this problem. Originating from the work by J. von Neumann [14] for alternating orthogonal projections onto closed linear subspaces, it has been generalized and extended to cases involving closed convex subsets and to parallel methods for the algorithm; e.g., see [1], [3], [4] and the references in these papers. Acceleration schemes for POCS-like methods for finding the shortest-distance projection of a

Mathematics Subject Classification: 65B99.
Key words and phrases: projections onto convex sets, convex feasibility problem, speed of convergence.

given point onto the intersection of a number of closed linear varieties have been given in [8]; in [10], an acceleration scheme for Kaczmarz's method is presented.

The POCS method, however, has one serious drawback: it often converges very slowly. This fact, sometimes called the "tunneling effect" [11], is in an intuitive manner explained as follows: in order to reach the intersection, the constructed iteration sequence enters a narrow corridor (a "tunnel") formed by two of the involved sets, and at each iteration step the progress towards the intersection set, caused by projecting the current iteration point onto the involved sets, is very small.

This explanation, however, is only half of the truth. Indeed, the decisive factor concerning the speed of convergence seems to be the combination of the structure of the involved sets together with the starting point of the iteration sequence (this fact makes it difficult to compare the rate of convergence of different algorithms in practical applications on the base of theoretical results). By way of example, we consider the problem of finding, with the POCS method, a point in the intersection of the following twelve disks $\{C_j\}_{j=1}^{12}$

$$(1) \qquad \left(x - \cos\left(\frac{j\pi}{12}\right)\right)^2 + \left(y - \sin\left(\frac{j\pi}{12}\right)\right)^2 \leq 1, \quad \text{for } j = 1, 2, \ldots, 12$$

in the plane (clearly, $(0,0)$ is a point in the intersection).

Denoting by P_j the projection onto the corresponding disk C_j, denoting momentarily the points in $\mathbb{R}^m$ by boldface letters, the iteration sequence as $\{\boldsymbol{x}_k\}_{k=0}^{+\infty}$ with $\boldsymbol{x}_k \equiv (x_k, y_k)$, and using the iteration $\boldsymbol{x}_{k+1} = T\boldsymbol{x}_k$ with $T \equiv P_{12}P_{11}\ldots P_2P_1$, we obtain very different convergence results depending on the choice of the starting point $\boldsymbol{x_o} \equiv (x_0, y_0)$ in the plane. For $(x_0, y_0) \equiv (-3, 0)$, a point in the intersection is obtained after one iteration; on the other hand, for $(x_0, y_0) \equiv (3, 4)$, the sum of the distances of the current iteration point to the sets C_j is 3.661634×10^{-3}, 5.49556×10^{-4} and 1.66893×10^{-5} after 25, 50 and 100 iterations respectively. A behaviour comparable to this one is observed when using different existing adaptations of the method of pure projections, as for instance for the parallel method

$$(2) \qquad \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_{k+1} \sum_{j=1}^{r} \mu_{k+1}(j)(P_j\boldsymbol{x}_k - \boldsymbol{x}_k),$$

with one variable relaxation coefficient λ_{k+1}, and with nonnegative weights $\mu_{k+1}(j)$ $(j = 1, \dots, r)$ such that $\sum_{j=1}^{r} \mu_{k+1}(j) = 1$.

Due to the fact that, when convex sets C_j are given, there is usually no clear indication of how to choose the starting point of the iteration in order to avoid the tunneling effect, it seems that, in order to circumvent this effect, it might be favourable to interrupt at different iteration steps during the procedure the "monotone" way of converging. By this we mean that, contrary to the existing algorithms where it is true that for the iteration sequence $\{\boldsymbol{x}_n\}_{n=0}^{+\infty}$ in $\mathbb{R}^m$ we have that

$$\|\boldsymbol{x}_{k+1} - \boldsymbol{w}\| \leq \|\boldsymbol{x}_k - \boldsymbol{w}\|, \ \forall \boldsymbol{w} \in C^*, \quad \text{for all } k, \tag{3}$$

the speed of convergence may be improved by an algorithm that automatically chooses, at some instances, the update iteration point $\boldsymbol{x}_{k+1}$ such that (3) is not necessarily true. At first glance, this may seem rather awkward, but in this way we may approach the tunnel from another direction, leading to an improvement of the convergence speed in those cases where the tunneling effect is strong, while keeping an acceptable speed of convergence in the other cases. With the algorithm presented in this paper, convergence to a point in the intersection of the given example is obtained after 9 iterations when starting from $(x_0, y_0) \equiv (-3, 0)$, and after 9 iterations when starting from $(x_0, y_0) \equiv (3, 4)$.

The method we present is an adaptation of the parallel method (2) with fixed weights $\mu_{k+1}(j)$. As has been shown by Pierra [12], the parallel method (2) in $\mathbb{R}^m$ may then be seen as a sequential method in the product space $(\mathbb{R}^m)^r$ (extensions of this with variable weights have been given in [5]). According to the Pierra method, in $(\mathbb{R}^m)^r$ there are two important subsets: the linear subspace $\mathcal{D}$, which is the canonical imbedding $q(\mathbb{R}^m)$ of $\mathbb{R}^m$ into $(\mathbb{R}^m)^r$, and the closed convex set $\mathcal{F} \equiv C_1 \times C_2 \times \cdots \times C_r$. Denoting for the moment the points in $(\mathbb{R}^m)^r$ by capital letters (in order to distinguish them from the points in $\mathbb{R}^m$), the ordinary parallel projection method (2) in $\mathbb{R}^m$ with fixed equal weights $\mu_{k+1}(j) \equiv \frac{1}{r}$ at each iteration step is then equivalent with the following procedure in $(\mathbb{R}^m)^r$ starting from some point $X_0 \equiv q(\boldsymbol{x}_0)$ in $\mathcal{D}$: whenever X_k in $\mathcal{D}$ has been obtained, the next iteration point X_{k+1} in $\mathcal{D}$ is obtained as a result of the following two steps: put

$$Y_{k+1} = X_k + \lambda_{k+1}(P_{\mathcal{F}} X_k - X_k), \tag{4}$$

where $P_{\mathcal{F}}X_k$ is the projection of X_k onto $\mathcal{F}$; then put

$$X_{k+1} = P_{\mathcal{D}}(Y_{k+1}) = X_k + \lambda_{k+1}(P_{\mathcal{D}}(P_{\mathcal{F}}X_k) - X_k). \tag{5}$$

Then, for suitable λ_{k+1} the sequence $\{X_k\}_{k=0}^{+\infty}$ in $\mathcal{D}$ is convergent to a point of $\mathcal{F} \cap \mathcal{D}$, and for each point W in $\mathcal{F} \cap \mathcal{D}$ the Fejér-monotone property ([13]) analogous to (3) is fulfilled (e.g., see [6]).

Formula (5), valid in $\mathcal{D} \subset (\mathbb{R}^m)^r$, has the typical structure of an iteration process: the update X_{k+1} of X_k is found by starting from X_k and choosing a direction in $\mathcal{D}$ (determined by $P_{\mathcal{D}}(P_{\mathcal{F}}X_k)$) and a step-length λ_{k+1}. This update process in (5) is derived from the update given in (4) which, however, is a process in $(\mathbb{R}^m)^r$ and as such gives much more flexibility. Indeed, while in (4) the computation of Y_{k+1} is started from $X_k \in \mathcal{D}$ in order to obtain (5), there is certainly no need to restrict such starting point of (4) to $\mathcal{D}$. When taking as starting point for (4) a point $Z_k \in (\mathbb{R}^m)^r$ that does not necessarily belong to $\mathcal{D}$ but is such that $P_{\mathcal{D}}Z_k = X_k$, then formulas (4) and (5) are changed into

$$Y_{k+1} = Z_k + \lambda_{k+1}(P_{\mathcal{F}}Z_k - Z_k), \tag{4'}$$

$$X_{k+1} = P_{\mathcal{D}}(Y_{k+1}) = X_k + \lambda_{k+1}(P_{\mathcal{D}}(P_{\mathcal{F}}Z_k) - X_k), \tag{5'}$$

and so also (5') is a typical iteration process in $\mathcal{D}$, for a suitable step-length λ_{k+1} and a direction determined by $P_{\mathcal{D}}(P_{\mathcal{F}}Z_k)$.

The basic idea in our paper is precisely to use the freedom of choosing Z_k in (4'), which itself entails more possibilities for the value of λ_{k+1}. Just for illustrative purposes we sketch the first two steps of our iteration algorithm. As concerns notation, when $Z_k \in (\mathbb{R}^m)^r$ but $Z_k \notin \mathcal{F}$, we denote by $\mathcal{P}_{k+1}$ the hyperplane supporting $\mathcal{F}$ at $P_{\mathcal{F}}Z_k$.

Let $Z_0 \equiv X_0$ be a starting point in $\mathcal{D}$. Put

$$Y_1 = Z_0 + \lambda_1(P_{\mathcal{F}}Z_0 - Z_0),$$

$$X_1 = P_{\mathcal{D}}Y_1,$$

where λ_1 is determined such that $X_1 \in \mathcal{P}_1 \cap \mathcal{D}$ (such λ_1 always exists). This is the first iteration point of our sequence $\{X_k\}_{k=0}^{+\infty}$ in $\mathcal{D}$. In order to obtain X_2, we construct Z_1 by using a relaxed projection of Y_1 onto $\mathcal{D}$ by means of

$$Z_1 = Y_1 + (1 + \gamma_1)(X_1 - Y_1) = X_1 + \gamma_1(X_1 - Y_1) \tag{6}$$

for some $\gamma_1 > 0$.

For this Z_1 we can verify if there is a real number λ_2 such that the projection onto $\mathcal{D}$ of the point $Z_1 + \lambda_2(P_{\mathcal{F}} Z_1 - Z_1)$ belongs to $\mathcal{P}_2 \cap \mathcal{D}$. If λ_2 exists and is bigger than 1, we put

$$Y_2 = Z_1 + \lambda_2(P_{\mathcal{F}} Z_1 - Z_1), \tag{7}$$

$$X_2 = P_{\mathcal{D}} Y_2 \tag{8}$$

(which is the second iteration point in this case),

$$Z_2 = Y_2 + (1 + \gamma_2)(X_2 - Y_2) \quad \text{for some } \gamma_2 > 0 \tag{9}$$

(which is the starting point for the next projection onto $\mathcal{F}$ in this case).

On the other hand, if $\lambda_2 \leq 1$, we put

$$Y_2 = P_{\mathcal{F}} Z_1, \tag{10}$$

$$Z_2 \equiv X_2 = P_{\mathcal{D}} Y_2 \tag{11}$$

(which is now at the same time the second iteration point X_2 of the approximation sequence, and the starting point for the next projection onto $\mathcal{F}$).

Finally, if λ_2 does not exist, then X_1 is chosen as the starting point to obtain X_2, in the same way as X_1 was constructed from X_0.

So, from X_1 on, each next iteration point X_{k+1} of the sequence $\{X_k\}_{k=0}^{+\infty}$ can be obtained on the discriminating base of the value of some parameter λ_{k+1}, and this may cause a different behaviour with respect to Fejér-monotony; in particular, the (monotone but) slow convergence may be interrupted at some iteration points.

The values of λ_{k+1} involved in this procedure play the role of relaxation coefficients connected with projections onto $\mathcal{F}$ when they are bigger than 1; as they may also be bigger than 2, the corresponding relaxed projection onto $\mathcal{F}$ has no longer the non-expansivity property.

The rest of the paper is organized as follows. In Section 2 we introduce the results that are needed to show that the algorithm is meaningful, and we describe the algorithm. In Section 3 we show that the sequence $\{X_k\}_{k=0}^{+\infty}$ in $\mathcal{D} \subset (\mathbb{R}^m)^r$, constructed according to the algorithm, is convergent to a point of $\mathcal{F} \cap \mathcal{D}$, leading to convergence of the corresponding sequence $\{\boldsymbol{x}_k\}_{k=0}^{+\infty}$ in $\mathbb{R}^m$ to a point in C^*. Finally, in Section 4 we illustrate the new algorithm for some examples , and we compare its behaviour for different starting points to the method of pure projections and to the parallel method (2).

2. Description of the algorithm

Let $\mathbb{R}^m$ be the m-dimensional Euclidean space with standard inner product $\langle\, ,\, \rangle$ and norm $|\quad|$ derived from $\langle\, ,\, \rangle$; denote $(\mathbb{R}^m, \langle\, ,\, \rangle, |\quad|)$ for short by H. Elements of H are denoted by boldface letters.

Suppose that in H, r closed convex subsets $\{C_j\}_{j=1}^r$ are given, having nonempty intersection $C^* \equiv \bigcap_{j=1}^r C_j \neq \emptyset$. Projection onto C_j is denoted as P_j. We want to obtain, in an iterative manner, a point in C^*.

Consider the r-fold product $(\mathbb{R}^m)^r$ of $\mathbb{R}^m$; elements of $(\mathbb{R}^m)^r$ are denoted by capital letters. We introduce an inner product $\langle\!\langle\ ,\ \rangle\!\rangle$ and norm $\|\quad\|$ as follows: when $V \equiv (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots \boldsymbol{v}_r) \in (\mathbb{R}^m)^r$ and $W \equiv (\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_r) \in (\mathbb{R}^m)^r$, put $\langle\!\langle V, W\rangle\!\rangle = \sum\limits_{j=1}^r \langle \boldsymbol{v}_j, \boldsymbol{w}_j\rangle, \|V\|^2 = \sum\limits_{j=1}^r |\boldsymbol{v}_j|^2$.

We denote $((\mathbb{R}^m)^r, \langle\!\langle\, ,\, \rangle\!\rangle, \|\quad\|)$ for short by $\mathcal{H}$.

In $\mathcal{H}$ we consider the subsets $\mathcal{D}$ and $\mathcal{F}$, defined as follows. $\mathcal{D}$ is the set of all r-tuples with equal components, i.e., for $\boldsymbol{v} \in H$ we have that $(\boldsymbol{v}, \boldsymbol{v}, \dots, \boldsymbol{v}) \in \mathcal{D} \subset \mathcal{H}$. $\mathcal{D}$ is the image of H under the canonical imbedding $q : H \to \mathcal{H}$, where for $\boldsymbol{v} \in H$ we put $q(\boldsymbol{v}) \equiv (\boldsymbol{v}, \boldsymbol{v}, \dots, \boldsymbol{v})$. $\mathcal{D}$ is a closed linear subspace of $\mathcal{H}$. Projection onto D is denoted as $P_{\mathcal{D}}$.

The subset $\mathcal{F}$ of $\mathcal{H}$ is defined as the Cartesian product of the convex sets $\{C_j\}_{j=1}^r$ in H, i.e., $\mathcal{F} = C_1 \times C_2 \times \cdots \times C_r$. It is a closed convex subset of $\mathcal{H}$. Projection onto $\mathcal{F}$ is denoted as $P_{\mathcal{F}}$. Clearly, as $C^* \neq \emptyset$ we also have that $\mathcal{F} \cap \mathcal{D} \neq \emptyset$, and, moreover, $q(C^*) = \mathcal{F} \cap \mathcal{D}$. Hence, obtaining a point in $C^* \subset H$ is equivalent to obtaining a point in $\mathcal{F} \cap \mathcal{D} \subset \mathcal{H}$. So, our aim is to construct a sequence $\{X_k\}_{k=0}^{+\infty}$ in $\mathcal{D} \subset \mathcal{H}$, starting from some given point $X_0 \in \mathcal{D}$, that converges in $\mathcal{H}$ to a point in $\mathcal{F} \cap \mathcal{D}$; the corresponding sequence $\{\boldsymbol{x}_k\}_{k=0}^{+\infty}$ in H with $q(\boldsymbol{x}_k) = X_k$ will then be convergent to a point in C^*.

In order to construct the sequence $\{X_k\}_{k=0}^{+\infty}$ in $\mathcal{H}$, we need some preliminary results.

Lemma 1. *Let* $V \equiv (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_r) \in \mathcal{H}$. *Then*

(i) $P_{\mathcal{F}} V = (P_1 \boldsymbol{v}_1, P_2 \boldsymbol{v}_2, \dots, P_r \boldsymbol{v}_r)$

(ii) $P_{\mathcal{D}} V = q(\frac{1}{r} \sum_{k=1}^r \boldsymbol{v}_k)$

Proof. (i) For arbitrary $S \equiv (\boldsymbol{s}_1, \boldsymbol{s}_2, \dots, \boldsymbol{s}_r) \in \mathcal{F}$ we have

$$\|V - S\|^2 = \sum_{j=1}^r |\boldsymbol{v}_j - \boldsymbol{s}_j|^2,$$

and this is minimal when $\boldsymbol{s}_j = P_j \boldsymbol{v}_j$.

(ii) See [12, Lemma 1.1]. □

For a given point Z in $\mathcal{H}$ with projection $P_{\mathcal{F}}Z$ onto $\mathcal{F}$ (assume $Z \neq P_{\mathcal{F}}Z$), consider the hyperplane $\mathcal{P}$ going through $P_{\mathcal{F}}Z$ which is orthogonal onto $Z - P_{\mathcal{F}}Z$, i.e.,

$$\mathcal{P} = \{X \in \mathcal{H} : \langle\langle X - P_{\mathcal{F}}Z, P_{\mathcal{F}}Z - Z\rangle\rangle = 0\}. \tag{12}$$

On the straight line through Z and $P_{\mathcal{F}}Z$ we consider the points Y given by the following expression

$$Y = Z + \lambda(P_{\mathcal{F}}Z - Z), \tag{13}$$

where λ is a real parameter.

Lemma 2. *Let $Z \in \mathcal{H}$, $Z \neq P_{\mathcal{F}}Z$, and put $Y = Z + \lambda(P_{\mathcal{F}}Z - Z)$. Then $P_{\mathcal{D}}Y$ belongs to $\mathcal{P}$ if and only if λ is given by*

$$\lambda = \frac{\langle\langle P_{\mathcal{D}}Z - P_{\mathcal{F}}Z, Z - P_{\mathcal{F}}Z\rangle\rangle}{\|P_{\mathcal{D}}(P_{\mathcal{F}}Z) - P_{\mathcal{D}}Z\|^2}. \tag{14}$$

PROOF. Assume that there exists a real number λ such that $P_{\mathcal{D}}Y$ belongs to $\mathcal{P}$.

As $P_{\mathcal{D}}$ is linear, $P_{\mathcal{D}}Y$ is given by

$$P_{\mathcal{D}}Y = P_{\mathcal{D}}Z + \lambda(P_{\mathcal{D}}(P_{\mathcal{F}}Z) - P_{\mathcal{D}}Z). \tag{15}$$

In the expression (12) of $\mathcal{P}$ we write the first part in the inner product sign as $X - P_{\mathcal{D}}Z + P_{\mathcal{D}}Z - P_{\mathcal{F}}Z$, and replace X by $P_{\mathcal{D}}Y$ as given by its expression on the right-hand-side of (15). The condition to determine λ then becomes

$$\lambda\langle\langle P_{\mathcal{D}}(P_{\mathcal{F}}Z) - P_{\mathcal{D}}Z, P_{\mathcal{F}}Z - Z\rangle\rangle + \langle\langle P_{\mathcal{D}}Z - P_{\mathcal{F}}Z, P_{\mathcal{F}}Z - Z\rangle\rangle = 0.$$

The first inner-product term on the left-hand-side may be written as

$$\langle\langle P_{\mathcal{D}}(P_{\mathcal{F}}Z) - P_{\mathcal{D}}Z, P_{\mathcal{D}}(P_{\mathcal{F}}Z) - Z\rangle\rangle,$$

due to the fact that $P_{\mathcal{D}}$ is an orthogonal projection onto $\mathcal{D}$. Hence,

$$\lambda = \frac{\langle\langle P_{\mathcal{D}}Z - P_{\mathcal{F}}Z, Z - P_{\mathcal{F}}Z\rangle\rangle}{\langle\langle P_{\mathcal{D}}(P_{\mathcal{F}}Z) - P_{\mathcal{D}}Z, P_{\mathcal{D}}(P_{\mathcal{F}}Z) - Z\rangle\rangle}.$$

The expression (14) is then obtained by writing in the denominator

$$P_{\mathcal{D}}(P_{\mathcal{F}}Z) - Z = P_{\mathcal{D}}(P_{\mathcal{F}}Z) - P_{\mathcal{D}}Z + P_{\mathcal{D}}Z - Z$$

and computing the inner product, using the fact that $P_{\mathcal{D}}Z - Z$ is orthogonal onto $\mathcal{D}$.

Conversely, when for a given $Z \in \mathcal{H}$, $Z \neq P_{\mathcal{F}}Z$, the number λ given by (14) exists, it follows by a routine control that, when that value of λ is substituted into (13), then $P_{\mathcal{D}}Y \in \mathcal{P}$. □

In the special case that the point Z belongs to $\mathcal{D}$ we obtain the following interesting result about the corresponding value of λ. Denoting now Z as X, we have $P_{\mathcal{D}}X = X, P_{\mathcal{F}}X \neq X$, and

$$\lambda = \frac{\|X - P_{\mathcal{F}}X\|^2}{\|P_{\mathcal{D}}(P_{\mathcal{F}}X) - X\|^2} = 1 + \frac{\|P_{\mathcal{F}}X - P_{\mathcal{D}}(P_{\mathcal{F}}X)\|^2}{\|P_{D}(P_{\mathcal{F}}X) - X\|^2} \tag{16}$$

where the last equality is obtained by using the Pythagorean theorem on the triangle with vertices X, $P_{\mathcal{F}}X$, $P_{\mathcal{D}}(P_{\mathcal{F}}X)$. Hence, when $X \in D$, $X \notin \mathcal{F}$, and the corresponding value of λ exists, then this value is never smaller than 1.

Lemma 3. *When $X \in \mathcal{D}$, $X \notin \mathcal{F}$, then λ is well-defined.*

Proof. We have to show that the denominator appearing in the expressions of (16) can not be zero. To this end, use is made of the following formulas concerning distances in H and $\mathcal{H}$ related to a point $X \equiv q(\boldsymbol{x})$ in $\mathcal{D} \subset \mathcal{H}$ with $\boldsymbol{x} \in H$:

$$\|X - P_{\mathcal{F}}X\|^2 = \sum_{j=1}^{r} |\boldsymbol{x} - P_j\boldsymbol{x}|^2, \tag{17}$$

$$\|X - P_{\mathcal{D}}(P_{\mathcal{F}}X)\|^2 = r\left|\sum_{j=1}^{r} \frac{1}{r}(P_j\boldsymbol{x} - \boldsymbol{x})\right|^2. \tag{18}$$

For $\boldsymbol{c} \in C^* \equiv \bigcap_{j=1}^{r} C_j \subset H$ we have

$$\left\langle \boldsymbol{x}-\boldsymbol{c}, \sum_{j=1}^{r}\frac{1}{r}(P_j\boldsymbol{x}-\boldsymbol{x})\right\rangle = \sum_{j=1}^{r}\frac{1}{r}\langle \boldsymbol{x}-P_j\boldsymbol{x}, P_j\boldsymbol{x}-\boldsymbol{x}\rangle + \sum_{j=1}^{r}\frac{1}{r}\langle P_j\boldsymbol{x}-\boldsymbol{c}, P_j\boldsymbol{x}-\boldsymbol{x}\rangle,$$

and each term in the last sum is non-positive due to a well-known property of projections. Hence,

$$\left\langle \boldsymbol{x} - \boldsymbol{c},\ \sum_{j=1}^{r} \frac{1}{r}(P_j\boldsymbol{x} - \boldsymbol{x}) \right\rangle \le -\frac{1}{r}\sum_{j=1}^{r} |\boldsymbol{x} - P_j\boldsymbol{x}|^2. \tag{19}$$

Assuming now that the denominator of (16) is zero, it follows from (18) that $\sum_{j=1}^{r} \frac{1}{r}(P_j\boldsymbol{x} - \boldsymbol{x}) = 0$, and from (19) it follows also that $\sum_{j=1}^{r} |\boldsymbol{x} - P_j\boldsymbol{x}|^2 = 0$; using (17) we conclude that then $X = P_{\mathcal{F}}X$, contrary to our assumptions. □

The foregoing result is interesting for our algorithm in the sense that, when an iteration point X_k in $\mathcal{D}$ has been obtained that not yet belongs to $\mathcal{F}$, and when that iteration point X_k is used as the starting point for the next iteration step in $\mathcal{H}$, then we are sure that the corresponding λ is well-defined; moreover, it is not smaller than 1.

Lemma 4. *Let $X \in \mathcal{D}$, $X \notin \mathcal{F}$, and put $Y = X + \lambda(P_{\mathcal{F}}X - X)$ with λ as given by* (16)*. Then $\lambda = 1$ if and only if $Y \in \mathcal{D}$.*

Proof. Suppose $\lambda = 1$. Then $Y = P_{\mathcal{F}}X$ and $P_{\mathcal{D}}Y = P_{\mathcal{D}}(P_{\mathcal{F}}X)$. Moreover, from the last expression in (16) we also conclude that then $P_{\mathcal{D}}(P_{\mathcal{F}}X) = P_{\mathcal{F}}X$. Hence, $Y = P_{\mathcal{D}}Y$, and so $Y \in \mathcal{D}$.

Conversely, suppose that $Y \in \mathcal{D}$, which means that $P_{\mathcal{D}}Y = Y$. Equating the expressions of Y and $P_{\mathcal{D}}Y$ leads to $\lambda(P_{\mathcal{F}}X - P_{\mathcal{D}}(P_{\mathcal{F}}X)) = 0$. As $\lambda \ge 1$, there results that $P_{\mathcal{F}}X = P_{\mathcal{D}}(P_{\mathcal{F}}X)$ and, again from (16), we conclude that $\lambda = 1$. □

In particular, when $X \in \mathcal{D}$, $X \notin \mathcal{F}$, and $\lambda = 1$, then $P_{\mathcal{D}}Y = P_{\mathcal{F}}X \in \mathcal{F} \cap \mathcal{D}$. For our algorithm this means that, when taking an iteration point X_k in $\mathcal{D}$ as the starting point for the next iteration step in $\mathcal{H}$, and when λ corresponding to X_k has the value 1, then X_{k+1} will belong to $\mathcal{F} \cap \mathcal{D}$ when $P_{\mathcal{D}}Y$ is taken as the next iteration point X_{k+1}.

For a point $Z \in \mathcal{H}$ with $Z \notin \mathcal{D}$ and $Z \notin \mathcal{F}$, Lemma 4 can be adapted as follows.

Lemma 5. *Let $Z \notin \mathcal{D}$, $Z \notin \mathcal{F}$, and put $Y = Z + \lambda(P_{\mathcal{F}}Z - Z)$ with λ as given by* (14)*. When $Y \in \mathcal{D}$, then necessarily $\lambda = 1$.*

Proof. By assumption, $P_{\mathcal{D}}Y = Y$ and $P_{\mathcal{D}}Y \in \mathcal{P}$. Hence,

$$\begin{aligned} 0 &= \langle\!\langle Z - P_{\mathcal{F}}Z, Y - P_{\mathcal{F}}Z \rangle\!\rangle = \|Z - P_{\mathcal{F}}Z\|^2 + \lambda \langle\!\langle Z - P_{\mathcal{F}}Z, P_{\mathcal{F}}Z - Z \rangle\!\rangle \\ &= (1-\lambda)\|Z - P_{\mathcal{F}}Z\|^2, \end{aligned}$$

which leads to the result, as $Z \neq P_{\mathcal{F}}Z$. $\square$

Let Z_0 be an element of $\mathcal{H}$, whether or not in $\mathcal{D}$, but $Z_0 \notin \mathcal{F}$. Put $Y_1 = Z_0 + \lambda_1(P_{\mathcal{F}}Z_0 - Z_0)$, and suppose that the corresponding λ_1-value in order that $P_{\mathcal{D}}Y_1$ should belong to the supporting hyperplane $\mathcal{P}_1$ at $P_{\mathcal{F}}Z_0$, exists and is bigger than 1. Denote $P_{\mathcal{D}}Y_1$ as X_1, and put, for some $\gamma > 0$

$$Z_1 = Y_1 + (1+\gamma)(X_1 - Y_1). \tag{20}$$

In view of Lemmas 4 and 5 we know that $X_1 \neq Y_1$. Hence, Z_1 is the result of a relaxed projection of Y_1 onto D. We also note that $P_{\mathcal{D}}Z_1 = X_1$.

We are interested in the fact whether or not Z_1 can belong to $\mathcal{F}$.

Lemma 6. *When $Z_0 \notin F$, $\lambda_1 > 1$, $\gamma > 0$, then $Z_1 \notin \mathcal{F}$.*

PROOF. Replacing in the following inner product Z_1 by its expression given in (20) we obtain

$$\begin{aligned}\langle\langle Z_0 - P_{\mathcal{F}}Z_0, Z_1 - P_{\mathcal{F}}Z_0\rangle\rangle = \langle\langle Z_0 - P_{\mathcal{F}}Z_0, X_1 - P_{\mathcal{F}}Z_0\rangle\rangle \\ + \gamma\langle\langle Z_0 - P_{\mathcal{F}}Z_0, X_1 - Y_1\rangle\rangle.\end{aligned}$$

The first inner product on the right-hand-side is zero due to the definition of λ_1. Replacing in the second inner product Y_1 by its expression in Z_0 and $P_{\mathcal{F}}Z_0$ we obtain

$$\langle\langle Z_0 - P_{\mathcal{F}}Z_0, Z_1 - P_{\mathcal{F}}Z_0\rangle\rangle = \gamma\lambda_1\|Z_0 - P_{\mathcal{F}}Z_0\|^2 + \gamma\langle\langle Z_0 - P_{\mathcal{F}}Z_0, X_1 - Z_0\rangle\rangle.$$

To obtain the inner product on the right-hand-side we write

$$\begin{aligned}\|X_1 - P_{\mathcal{F}}Z_0\|^2 &= \|X_1 - Z_0\|^2 + \|Z_0 - P_{\mathcal{F}}Z_0\|^2 + 2\langle\langle X_1 - Z_0, Z_0 - P_{\mathcal{F}}Z_0\rangle\rangle \\ &= \|X_1 - P_{\mathcal{F}}Z_0\|^2 + 2\|Z_0 - P_{\mathcal{F}}Z_0\|^2 + 2\langle\langle X_1 - Z_0, Z_0 - P_{\mathcal{F}}Z_0\rangle\rangle,\end{aligned}$$

due to the Pythagorean theorem for the right triangle with vertices Z_0, $P_{\mathcal{F}}Z_0$ and X_1. Hence,

$$\langle\langle X_1 - Z_0, Z_0 - P_{\mathcal{F}}Z_0\rangle\rangle = -\|Z_0 - P_{\mathcal{F}}Z_0\|^2.$$

Using this result in the former equality we obtain

$$\langle\langle Z_0 - P_{\mathcal{F}}Z_0, Z_1 - P_{\mathcal{F}}Z_0\rangle\rangle = \gamma(\lambda_1 - 1)\|Z_0 - P_{\mathcal{F}}Z_0\|^2.$$

Hence, $\langle\langle Z_0 - P_{\mathcal{F}}Z_0, Z_1 - P_{\mathcal{F}}Z_0\rangle\rangle$ is strictly positive, from which we conclude, due to a property of projections, that $Z_1 \notin \mathcal{F}$. $\square$

Now we are ready to describe our algorithm to construct the approximating sequence $\{X_k\}_{k=0}^{+\infty}$ in $\mathcal{D}$ in order to obtain a point in $\mathcal{D} \cap \mathcal{F}$.

Algorithm

Starting from some point $Z_0 \equiv X_0 \in \mathcal{D}$, suppose that points $X_1, \ldots$ $\ldots, X_k$ in $\mathcal{D}$ and $Z_1, \ldots, Z_k$ in $\mathcal{H}$ (whether or not in $\mathcal{D}$) have been obtained, with $P_{\mathcal{D}} Z_k = X_k$, $X_k \notin \mathcal{F} \cap \mathcal{D}$, $Z_k \notin \mathcal{F}$. Then:

(i) IF $P_{\mathcal{D}}(P_{\mathcal{F}} Z_k) = X_k$, then take X_k as new Z_k and go to (ii)

(ii) ELSE, compute λ_{k+1} by (14)(with Z replaced by Z_k)

a. IF $\lambda_{k+1} > 1$, let

$$\begin{aligned}
Y_{k+1} &= Z_k + \lambda_{k+1}(P_{\mathcal{F}} Z_k - Z_k), \\
X_{k+1} &= P_{\mathcal{D}} Y_{k+1}, \\
Z_{k+1} &= Y_{k+1} + (1 + \gamma_{k+1})(X_{k+1} - Y_{k+1}),
\end{aligned}$$

(where $\gamma_{k+1} > 0$ will be determined further on)

b. ELSE, let $X_{k+1} = P_{\mathcal{D}}(P_{\mathcal{F}} Z_k)$,
$Z_{k+1} = X_{k+1}$. □

As has been shown in Lemma 3 and in the remark following (16), whenever $Z_k \equiv X_k \in \mathcal{D}$ but $X_k \notin \mathcal{F}$ then λ_{k+1} is well-defined and is never smaller than 1. Hence, in that case the steps (i) and (ii)a. of the algorithm need not be investigated; moreover, when $\lambda_{k+1} = 1$, then $X_{k+1} \in \mathcal{D} \cap \mathcal{F}$ according to Lemma 4.

Concerning a suitable value for γ_{k+1} needed in (ii)b., we remember that taking Z_{k+1} instead of X_{k+1} as a new starting point in $\mathcal{H}$ for the computation of the next update X_{k+2} has the purpose of possibly interrupting, at some steps, the monotone behaviour of the approximating sequence $\{X_k\}_{k=0}^{+\infty}$; this needs a γ_{k+1} that is different from zero; we choose it strictly positive. On the other hand, however, we need convergence of the sequence $\{X_k\}_{k=0}^{+\infty}$, and so we want an interruption that somehow takes into consideration part of the approximation already obtained. These facts seem to imply that, in any case, the distance between X_{k+1} and Z_{k+1} should tend to zero with growing k. To this end, for M a fixed positive number we let γ_{k+1} be dependent on $\frac{M}{k+1}$. In order to take care of the fact that $\{\lambda_{k+1}\}_{k=0}^{+\infty}$ may be unbounded, and hence also $\{\|X_{k+1} - Y_{k+1}\|\}_{k=0}^{+\infty}$, let B be a (big) fixed positive number; we then put

$$\alpha_{k+1} = \min\left(\frac{1}{\lambda_{k+1}}, \frac{M}{k+1}\right), \qquad \beta_{k+1} = \min\left(1, \frac{B}{\|X_{k+1} Y_{k+1}\|}\right),$$

$$\gamma_{k+1} = \alpha_{k+1}\beta_{k+1}.$$

3. Convergence of the constructed sequence

In this section we show that the sequence $\{X_k\}_{k=0}^{+\infty}$ in $\mathcal{D}$ is convergent to a point of $\mathcal{D} \cap \mathcal{F}$. Due to the fact that the construction of the points X_{k+1} in $\mathcal{D}$ is different according to the corresponding values of λ_{k+1}, we have to consider separately the properties of elements of the sequence corresponding to values of λ_{k+1} that are bigger than 1, or not.

A. Properties corresponding to $\lambda_{k+1} > 1$

Let us suppose that we obtained X_k in $\mathcal{D}$, Z_k in $\mathcal{H}$ with $P_{\mathcal{D}}(Z_k) = X_k$ (it is possible that $X_k = Z_k$), and that λ_{k+1}, given by (14) with Z replaced by Z_k, is well-defined and is bigger than 1. According to (ii)b. of the algorithm, we have

$$X_{k+1} = P_D Y_{k+1} = X_k + \lambda_{k+1}(P_{\mathcal{D}}(P_{\mathcal{F}} Z_k) - X_k), \tag{21}$$

and X_{k+1} belongs to the hyperplane $\mathcal{P}_{k+1}$ supporting $\mathcal{F}$ at $P_{\mathcal{F}} Z_k$; its equation is

$$\mathcal{P}_{k+1} = \{X \in \mathcal{H} : \langle\!\langle X - P_{\mathcal{F}} Z_k, P_{\mathcal{F}} Z_k - Z_k \rangle\!\rangle = 0\}. \tag{22}$$

Lemma 7.

$$\langle\!\langle X_k - X_{k+1}, V - X_{k+1} \rangle\!\rangle \leq 0, \quad \forall V \in \mathcal{D} \cap \mathcal{F}. \tag{23}$$

Proof. Using the facts that $X_k - Z_k$ and $Y_{k+1} - X_{k+1}$ are orthogonal onto D, and that $V - X_{k+1} \in \mathcal{D}$, we have

$$\begin{aligned} \langle\!\langle X_k - X_{k+1}, V - X_{k+1} \rangle\!\rangle &= \langle\!\langle Z_k - X_{k+1}, V - X_{k+1} \rangle\!\rangle \\ &= \langle\!\langle Z_k - Y_{k+1}, V - X_{k+1} \rangle\!\rangle. \end{aligned}$$

Replacing Y_{k+1} by its expression in Z_k and $P_{\mathcal{F}} Z_k$ leads to

$$\begin{aligned} &\langle\!\langle X_k - X_{k+1}, V - X_{k+1} \rangle\!\rangle = -\lambda_{k+1} \langle\!\langle P_{\mathcal{F}} Z_k - Z_k, V - X_{k+1} \rangle\!\rangle \\ &\quad = \lambda_{k+1} \langle\!\langle Z_k - P_{\mathcal{F}} Z_k, V - P_{\mathcal{F}} Z_k \rangle\!\rangle + \lambda_{k+1} \langle\!\langle Z_k - P_{\mathcal{F}} Z_k, P_{\mathcal{F}} Z_k - X_{k+1} \rangle\!\rangle. \end{aligned}$$

In the right-hand-side, the last term is zero because $X_{k+1} \in \mathcal{P}_{k+1}$, while the first term is nonpositive due to a well-known property of projections.

Hence the result follows. □

For points $V \in \mathcal{D} \cap \mathcal{F}$ we further have

$$\|X_k - V\|^2 = \|X_k - X_{k+1}\|^2 + \|X_{k+1} - V\|^2 + 2\langle\langle X_k - X_{k+1}, X_{k+1} - V\rangle\rangle,$$

and taking (23) into account we derive that

$$\|X_k - V\|^2 \geq \|X_k - X_{k+1}\|^2 + \|X_{k+1} - V\|^2, \quad \forall V \in \mathcal{D} \cap \mathcal{F}. \tag{24}$$

In particular, we conclude that, when X_{k+1} is obtained using a value of λ_{k+1} that is bigger than 1, then

$$\|X_k - V\| \geq \|X_{k+1} - V\|, \quad \forall V \in \mathcal{D} \cap \mathcal{F},$$

which expresses the monotone approach to $D \cap F$ in that case.

Let us suppose now that, except at a finite number of steps, all values of λ_{k+1} are bigger than 1; let us restrict to such values of k. Then from (24) we derive that

$$\|X_k - X_{k+1}\| \to 0 \quad \text{when} \quad k \to +\infty. \tag{25}$$

Making use of the Pythagorean theorem for the right triangles with vertices Z_k, X_k, X_{k+1} and Z_k, $P_{\mathcal{F}}Z_k$, X_{k+1} respectively, we obtain

$$\begin{aligned}\|X_{k+1} - X_k\|^2 &= \|Z_k - X_{k+1}\|^2 - \|Z_k - X_k\|^2 \\ &= \|Z_k - P_{\mathcal{F}}Z_k\|^2 + \|P_{\mathcal{F}}Z_k - X_{k+1}\|^2 - \|Z_k - X_k\|^2.\end{aligned}$$

Due to our choice of γ_{k+1} in the algorithm we know that $\|Z_k - X_k\| \to 0$ when $k \to +\infty$. Hence, we obtain that also $\|Z_k - P_{\mathcal{F}}Z_k\| \to 0$ and $\|P_{\mathcal{F}}Z_k - X_{k+1}\| \to 0$ when $k \to +\infty$. As

$$\|X_k - P_{\mathcal{F}}X_k\| \leq \|X_k - P_{\mathcal{F}}Z_k\| \leq \|X_k - X_{k+1}\| + \|X_{k+1} - P_{\mathcal{F}}Z_k\|,$$

there results that also

$$\|X_k - P_{\mathcal{F}}X_k\| \to 0 \quad \text{when} \quad k \to +\infty. \tag{26}$$

Combining the fact that the sequence $\{X_k\}_{k=0}^{+\infty}$ is bounded with the relations (24), (25) and (26), we have the necessary information needed to prove that the sequence $\{X_k\}_{k=0}^{+\infty}$ is weakly (and hence in norm) convergent to a point of $\mathcal{F} \cap \mathcal{D}$. We do not represent the proof here, but refer instead to [6] where it has been shown (in another situation) that from the mentioned properties convergence may be obtained.

B. Properties corresponding to $\lambda_{k+1} \le 1$

We know that, when $Z_k \equiv X_k \in \mathcal{D}$, then $\lambda_{k+1} \ge 1$, and when $\lambda_{k+1}=1$ then $X_{k+1} \equiv P_{\mathcal{F}} Z_k$ is a point of $\mathcal{D} \cap \mathcal{F}$ and the procedure is finished; so, when $Z_k \equiv X_k$ and the next X_{k+1} is not yet in $\mathcal{D} \cap \mathcal{F}$, we always have that $\lambda_{k+1} > 1$; we also started with $Z_0 \equiv X_0 \in \mathcal{D}$.

On the other hand, when $Z_k \neq X_k$ and $\lambda_{k+1} \le 1$, then we know from (ii)c. of the algorithm that $X_{k+1} = P_{\mathcal{D}}(P_{\mathcal{F}} Z_k) = Z_{k+1}$. It is at such point X_{k+1} that the monotone behaviour of the sequence $\{X_k\}_{k=0}^{+\infty}$ may be (but is not necessarily) interrupted, i.e., it is not necessarily true that $\|X_{k+1} - V\| \le \|X_k - V\|$, $\forall V \in \mathcal{F} \cap \mathcal{D}$, but, as $Z_{k+1} \equiv X_{k+1} \in \mathcal{D}$, we always have that $\lambda_{k+2} \ge 1$.

Hence, in the case that we are investigating now, we may suppose that the sequence $\{X_k\}_{k=0}^{+\infty}$ has an (infinite) subsequence $\{X_{n_j}\}_{j=0}^{+\infty}$ with $X_{n_0} \equiv Z_{n_0}$, $X_{n_1} \equiv Z_{n_1}, \ldots, X_{n_j} \equiv Z_{n_j}, \ldots$, with $n_0 = 0$, and those points were obtained from a corresponding λ_{n_j}-value not bigger than 1 (except for the starting point X_0). The λ-values for points between two consecutive points X_{n_p} and X_{n_q} of the subsequence are strictly bigger than 1; i.e., we have that $\lambda_{n_p} \le 1$ (except for $n_p = 0$), $\lambda_{n_p+1} > 1, \ldots, \lambda_{n_q-1} > 1, \lambda_{n_q} \le 1$.

We show that, due to our choice of γ_{k+1}, a weaker form of monotony is still available, which will be sufficient to guarantee convergence of the sequence. We first prove the following lemma, valid for λ-values that are strictly bigger than 1.

Lemma 8. *Suppose, with the usual notations, that for some $k \in \mathbb{Z}^+$ we have that $X_k \in \mathcal{D}$, $Z_k \notin \mathcal{D}$, $\lambda_{k+1} > 1$, giving rise to Y_{k+1}, X_{k+1} and Z_{k+1}. Then*

$$\|Z_{k+1} - X_{k+1}\|^2 \le \|Z_k - X_k\|^2 + \|X_k - X_{k+1}\|^2. \tag{27}$$

Proof. The straight line through $P_{\mathcal{F}} Z_k$ that is parallel to the straight line through Y_{k+1} and Z_{k+1}, intersects the straight line $Z_k X_{k+1}$ in a point S_k. From the similarity of the triangles with vertices $Z_k, P_{\mathcal{F}} Z_k, S_k$ and Z_k, Y_{k+1}, X_{k+1} we deduce that $\|Y_{k+1} - X_{k+1}\| = \lambda_{k+1} \|P_{\mathcal{F}} Z_k - S_k\|$. But $Z_{k+1} = X_{k+1} + \gamma_{k+1}(X_{k+1} - Y_{k+1})$, and so

$$\|Z_{k+1} - X_{k+1}\| \le \gamma_{k+1}\lambda_{k+1}\|P_{\mathcal{F}} Z_k - S_k\| \le \|P_{\mathcal{F}} Z_k - S_k\|.$$

Moreover, $\|P_{\mathcal{F}}Z_k - S_k\| \leq \|Z_k - X_{k+1}\|$, which follows from inspection of the triangle with vertices Z_k, $P_{\mathcal{F}}Z_k$, X_{k+1}. Finally, $Z_k - X_k$ is orthogonal onto $X_k - X_{k+1}$, and so we obtain

$$\|Z_{k+1} - X_{k+1}\|^2 \leq \|Z_k - X_{k+1}\|^2 = \|Z_k - X_k\|^2 + \|X_k - X_{k+1}\|^2. \quad \square$$

Using (27) and the Fejér-monotony property (24) a number of times leads to the following weaker form of monotony in the case where an interrupting subsequence $\{X_{n_j}\}_{j=0}^{+\infty}$ may exist.

Lemma 9. *Suppose that $X_{n_p} \equiv Z_{n_p}$ and $X_{n_q} \equiv Z_{n_q}$ are consecutive points of the subsequence $\{X_{n_j}\}_{j=0}^{+\infty}$. Then, $\forall V \in \mathcal{D} \cap \mathcal{F}$ we have*

$$\|X_{n_p} - V\|^2 \geq \|X_{n_q} - V\|^2 + \frac{3}{4}\|X_{n_p} - X_{n_p+1}\|^2. \tag{28}$$

Proof. In order not to overload the proof with cumbersome indexes, we present a proof with concrete small numbers. So, let $n_p = 0$, $n_q = 4$, $V \in \mathcal{F} \cap \mathcal{D}$. Then we first have

$$\|V - X_4\|^2 \leq \|V - P_{\mathcal{F}}Z_3\|^2 \leq \|V - Z_3\|^2 = \|V - X_3\|^2 + \|X_3 - Z_3\|^2, \tag{29}$$

by using the non-expansivity property of projections and the fact that $Z_3 - X_3$ is orthogonal onto $\mathcal{D}$. Using (27) for $k = 2$ leads to

$$\|V - X_4\|^2 \leq \|V - X_3\|^2 + \|Z_2 - X_2\|^2 + \|X_2 - X_3\|^2.$$

As $\lambda_3 > 1$, inequality (24) may be used for $k = 2$, giving

$$\begin{aligned}\|V - X_4\|^2 &\leq \|V - X_2\|^2 - \|X_2 - X_3\|^2 + \|Z_2 - X_2\|^2 + \|X_2 - X_3\|^2 \\ &= \|V - X_2\|^2 + \|X_2 - Z_2\|^2,\end{aligned}$$

and it is instructive to compare this result with (29). Hence, again applying (24) and (27), now for $k = 1$, leads to

$$\begin{aligned}\|V - X_4\|^2 &\leq \|V - X_1\|^2 - \|X_1 - X_2\|^2 + \|Z_1 - X_1\|^2 + \|X_1 - X_2\|^2 \\ &= \|V - X_1\|^2 + \|Z_1 - X_1\|^2.\end{aligned}$$

For the final step we use two different things. First, according to (24) we have that $\|V - X_1\|^2 \leq \|V - X_0\|^2 - \|X_0 - X_1\|^2$. For the term $\|Z_1 - X_1\|^2$, however, we remark that now $X_0 \in \mathcal{D}$ (remember that this is true for the general index n_p). Hence, the triangle with vertices X_0, $P_{\mathcal{F}}X_0$ and X_1 has a right angle in $P_{\mathcal{F}}X_0$. From some elementary geometry it follows then easily that

$$\|Z_1 - X_1\|^2 \leq \|P_{\mathcal{F}}X_0 - P_{\mathcal{D}}(P_{\mathcal{F}}X_0)\|^2 \leq \frac{1}{4}\|X_0 - X_1\|^2.$$

So, we finally obtain

$$\|V - X_4\|^2 \leq \|V - X_0\|^2 - \|X_0 - X_1\|^2 + \frac{1}{4}\|X_0 - X_1\|^2,$$

which is precisely formula (28) in our case. □

Inequality (28) gives an idea about the measure of interrupting the monotone behaviour at the points of the subsequence: although it may be true that $\|X_{n_q} - W\| > \|X_{n_q - 1} - W\|$ for some points W in $\mathcal{F} \cap \mathcal{D}$, the monotony is repaired with respect to the foregoing interruption point.

Due to inequalities (28) and (24) we are able to prove that, also in case that the sequence $\{X_k\}_{k=0}^{+\infty}$ has an infinite interrupting subsequence $\{X_{n_p}\}_{p=0}^{+\infty}$, the whole sequence $\{X_k\}_{k=0}^{+\infty}$ is convergent to a point $A \in \mathcal{D} \cap \mathcal{F}$. Again, the method described in [6] may be used. So we just sketch the main ideas. The subsequence $\{X_{n_p}\}_{p=0}^{+\infty}$ has a subsequence, say $\{X'_q\}_{q=0}^{+\infty}$, that weakly converges to a point $A \in \mathcal{D}$; this point also belongs to $\mathcal{F}$, as $\|P_{\mathcal{F}}X_{n_p} - X_{n_p}\| \to 0$. Moreover, each subsequence of $\{X_{n_p}\}_{p=0}^{+\infty}$ that weakly converges has the same point A as its weak limit; hence, $\{X_{n_p}\}_{p=0}^{+\infty}$ is (weakly and in) norm convergent to $A \in \mathcal{D} \cap \mathcal{F}$. Convergence of the whole sequence $\{X_k\}_{k=0}^{+\infty}$ to A then follows from the fact that, for $k > n_q$ we have that $\|X_k - A\| \leq \|X_{n_q} - A\|$.

We summarise the results in the following theorem.

Theorem. *Suppose that in $\mathbb{R}^m$, r closed convex sets $\{C_j\}_{j=1}^{r}$ with nonempty intersection $\bigcap_{j=1}^{r} C_j$ are given. Let q be the natural imbedding of $\mathbb{R}^m$ in $(\mathbb{R}^m)^r$, and put $\mathcal{D} = q(\mathbb{R}^m)$, $\mathcal{F} = C_1 \times C_2 \times \cdots \times C_r$. For a given point $\boldsymbol{x}_0 \in \mathbb{R}^m$, put $X_0 \equiv q(\boldsymbol{x}_0)$. Then the sequence $\{X_k\}_{k=0}^{+\infty}$, constructed according to the algorithm mentioned in Section 2, is convergent to a point of $\mathcal{F} \cap \mathcal{D}$. Hence, the sequence $\{\boldsymbol{x}_k\}_{k=0}^{+\infty}$ in $\mathbb{R}^m$ with $\boldsymbol{x}_k = q^{-1}(X_k)$ is convergent to a point in $\bigcap_{j=1}^{r} C_j$.*

4. Examples and concluding remarks

In this last section we illustrate the algorithm given in Section 2, and we compare the results with the ones corresponding to two other methods. The following algorithms have been used:

PP: The method of pure projections in a sequential manner: when $\boldsymbol{x}_k$ is the current iteration point, $\boldsymbol{x}_{k+1}$ is obtained by $\boldsymbol{x}_{k+1} = T\boldsymbol{x}_k$, with $T \equiv P_r \dots P_2 P_1$, and with P_j denoting the projection operator onto C_j.

PAR: The parallel projection given in (2) with fixed equal weights (i.e., $\mu_{k+1}(j) = \frac{1}{r}$ for each k), and with λ_{k+1} determined as in [6], i.e.,

$$\lambda_{k+1} = \frac{\sum_{j=1}^{r} \frac{1}{r} |\boldsymbol{x}_k - P_j \boldsymbol{x}_k|^2}{|\boldsymbol{x}_k - \sum_{j=1}^{r} \frac{1}{r} P_j \boldsymbol{x}_k|^2}.$$

NEW0 and NEW3: The algorithm presented in this paper, with $M = 1$ and $M = 10^3$ respectively, and with $B = 10^6$ in both cases.

As a first example, we take as closed convex sets the twelve disks $\{C_j\}_{j=1}^{12}$ mentioned in formula (1), Section 1; their intersection (in fact determined by C_1 and C_{12}) contains more than one point. Starting from some given point in the plane, we want to obtain a point in their intersection; explicit expressions for P_j may be found in [2].

In Table 1 we have mentioned, for the given eight starting points $((-3, 0), \dots)$, either the number of iterations needed to obtain a point in the intersection (this is a positive integer), or the sum of the distances of the current iteration point to the twelve sets C_j, after 25 and 50 iterations respectively. The results in Table 1 suggest that the convergence behaviour of the algorithm presented in this paper is much less dependent on the starting point than in the other cases; extremely slow convergence did not appear, for a lot of starting points.

In our second example we considered in $\mathbb{R}^3$ with variables x, y, z the intersection of the ball given by $x^2 + y^2 + z^2 \leq R^2$ (with $\frac{1}{6} \leq R \leq 1$), and the three half-spaces given by $x+y+4z \leq 1$, $x+y-4z \leq 1$, $-x+y-8z \leq 1$. The four algorithms mentioned in example 1 all gave rather comparable results: a point in the intersection was obtained in at most 4 iterations, for different starting points and for different values of the radius R of the ball.

The third example, again in $\mathbb{R}^3$, investigates the case of (hyper)planes making small angles with each other. It is known that such case leads to very slow convergence by using the existing POCS algorithm. We considered four planes through the z-axis and four planes through the y-axis, having the origin as intersection point. The equations of the planes in $\mathbb{R}^3$ were given by $y = x$, $y = 1.4x$, $y = 1.7x$, $y = 2x$, $z = 4x$, $z = 4.4x$, $z = 4.7x$, $z = 5x$ respectively.

Starting Point	$(-3, 0)$	$(10, -10)$	$(3, 4)$
PP	1	3.279208×10^{-3} 5.000838×10^{-4}	3.661634×10^{-3} 5.49556×10^{-4}
PAR	9.972098×10^{-3} 3.128052×10^{-3}	4	1.129448×10^{-2} 3.427267×10^{-3}
NEW0	9	5	9
NEW3	20	5	8

Starting Point	$(-17, 12)$	$(-2, 1)$	$(-100, -50)$
PP	3.601907×10^{-3} 5.419265×10^{-4}	3.202676×10^{-3} 4.89951×10^{-4}	1
PAR	1.185358×10^{-2} 3.548027×10^{-3}	9.768488×10^{-3} 3.080129×10^{-3}	8.859039×10^{-3} 2.859947×10^{-3}
NEW0	10	10	10
NEW3	8	9	42

Starting Point	$(2, -4)$	$(0, 2)$
PP	3.005983×10^{-3} 4.637248×10^{-4}	3.694175×10^{-3} 5.537283×10^{-4}
PAR	5	9.757404×10^{-3} 3.077506×10^{-3}
NEW0	5	9
NEW3	5	10

Table 1.

In Table 2 we give, again for the four algorithms mentioned in the first example, the result either of the sum of the distances of the current iteration point to the eight planes after 1000 iterations, or the number of iterations needed to stop the procedure (this happened when the sum of the distances was less than 1×10^{-8}). Of course, this example is not intended to present an alternative for existing algebraic methods in this case, but only to compare the behaviour of the mentioned algorithms.

Starting Point	(0.1, 0.2, 0.3)	(−1, 2, −3)	(3, −1, 2)
PP	4.846649×10^{-6}	3.737408×10^{-5}	3.23111×10^{-5}
PAR	7.679005×10^{-3}	7.220158×10^{-2}	4.867536×10^{-3}
NEW0	9.224179×10^{-3}	0.067403	6.159196×10^{-2}
NEW3	374	372	430

Table 2.

The less good behaviour of NEW0 when compared to NEW3 may be explained by the fact that, with the relaxation used in NEW0, the monotoneous way of iteration has never been interrupted, contrary to what happens when using NEW3. In any case, this last example seems to suggest that interrupting the monotony by using suitable relaxations with respect to $\mathcal{D}$ can substantially improve the speed of convergence.

References

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

[2] D. Butnariu and S. D. Flam, Strong convergence of expected-projection methods in Hilbert spaces, *Numer. Funct. Anal. and Optimiz.* **16** (1995), 601–636.

[3] P. L. Combettes, The foundations of set theoretic estimation, *Proc. IEEE* **81** (1993), 182–208.

[4] G. Crombez, Weak and norm convergence of a parallel projection method in Hilbert spaces, *Applied Mathem. and Computation* **56** (1993), 35–48.

[5] G. Crombez, Viewing parallel projection methods as sequential ones in convex feasibility problems, *Trans. Amer. Math. Soc.* **347** (1995), 2575–2583.

[6] G. Crombez, Solving convex feasibility problems by a parallel projection method with geometrically-defined parameters, *Applicable Analysis* **64** (1997), 277–290.

[7] F. Deutsch, The method of alternating orthogonal projections, Approximation theory, spline functions and applications, *Kluwer Academic Publishers*, 1992, 105–121.

[8] W. Gearhart and M. Koshy, Acceleration schemes for the method of alternating projections, *J. Comp. Appl. Math.* **26** (1989), 235–249.

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

[10] M. Hanke and W Niethammer, On the acceleration of Kaczmarz's method for inconsistent linear systems, *Linear Algebra Appl.* **130** (1990), 83–98.

[11] A. Levi and H. Stark, Restoration from phase and magnitude by generalized projections, Image recovery: Theory and application, *Academic Press*, *New York*, 1987, 277–320.

[12] G. Pierra, Decomposition through formalization in a product space, *Math. Programming* **28** (1984), 96–115.

[13] D. Schott, Iterative solution of convex problems by Fejér-monotone methods, *Numer. Funct. Anal. and Optimiz.* **16** (1995), 1323–1357.

[14] J. von Neumann, Functional operators, Vol. II, The geometry of orthogonal spaces, Annals of Math. Studies # 22, *Princeton University Press*, 1950.

[15] Y. Yang, N. Galatsanos and A. Katsaggelos, Projection-based spatially adaptive reconstruction of block-transform compressed images, *IEEE Trans. Image Processing* **4** (1995), 896–908.

[16] D. C. Youla, Mathematical theory of image restoration by the method of convex projections, Image recovery: Theory and Application, *Academic Press*, *New York*, 1987, 29–77.

G. CROMBEZ
DEPARTMENT OF APPLIED MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF GHENT
KRIJGSLAAN 281/S9
B–9000 GENT
BELGIUM

E-mail: Gilbert.Crombez@rug.ac.be

(Received November 2, 1998; revised November 2, 1999;
file arrived February 10, 2000)