Computer Math., Vol. 34, pp. 79–94
available directly from the publisher
permitted by license only

ON THE BEHAVIOR OF A BLOCK-ITERATIVE PROJECTION METHOD FOR SOLVING CONVEX FEASIBILITY PROBLEMS

DAN BUTNARIU and YAIR CENSOR*

Department of Mathematics and Computer Science, University of Haifa, 31999 Haifa, Israel

(*Received 9 October 1989; in final form 9 November 1989*)

The behavior of a class of block-iterative projection algorithms for solving convex feasibility problems is studied. A limit characterization theorem and a convergence criterion are proven. Ways of accelerating the computational procedures are pointed out.

KEY WORDS: Convex feasibility problem, relaxation method, orthogonal projection, simultaneous projection operator, transfer operator, block-iterative projection method.

C.R. CATEGORIES: G.1.6, G.1.3.

1. INTRODUCTION

The convex feasibility problem is the problem of computing points laying in the intersection of a finite family of closed convex subsets $\{Q_i | i \in I\}$ of a Hilbert space X. This problem appears in various fields of applied mathematics. Theory of Optimization [4], Image Reconstruction from Projections [7] and Game Theory [5] are some examples. The, so called, "relaxation methods" for solving convex feasibility problems which date back to Kaczmarz [15] and Cimmino [9] are of special interest because of their relatively easy implementation and computational efficiency in solving extremely large and sparse problems. Classical contributions to the study of relaxation methods are surveyed in [6].

Aharoni and Censor have recently discussed a "block-iterative projection method" (called BIP in this note) which incorporates as special cases many of the earlier relaxation techniques (see [2] and the references therein). The *BIP algorithm* is the following procedure for generating a sequence $\{x^k | k \in \mathbb{N}_0\}$ of vectors in the Hilbert space X: Choose an *initial point* x^0 in X and, for any nonnegative integer k, compute x^{k+1} by

$$x^{k+1} = x^k + \lambda_k \cdot (\mathbf{P}_{w_k} x^k - x^k) \tag{1.1}$$

*Work of Yair Censor supported in part by NIH grant HL-28438 while visiting the Medical Image Processing Group (MIPG), Department of Radiology, Hospital of the University of Pennsylvania, Philadelphia, PA.

where, for any $k \in \mathbb{N}_0$, $w_k: I \to \mathbb{R}$ are *weight functions*, that is they satisfy the condition

$$\sum_{i \in I} w_k(i) = 1 \quad \text{and} \quad w_k(i) \geqq 0, \qquad (i \in I), \tag{1.2}$$

$\lambda_k \in \mathbb{R}$ are *relaxation parameters* and $\mathbf{P}_{w_k}: X \to X$ are defined by

$$\mathbf{P}_{w_k} x := \sum_{i \in I} w_k(i) \cdot P_i x \tag{1.3}$$

with $P_i = P_{Q_i}$, $(i \in I)$, being the orthogonal projection operator on the closed convex set Q_i. It is proved in [2] that, if X is the n-dimensional Euclidean space and if $\sum_{k=0}^{\infty} w_k(i) = +\infty$, $(i \in I)$, then the BIP algorithm converges for any initial point x^0, as long as all relaxation parameters λ_k are confined to a closed subinterval of $]0,2[$. Under these conditions, if the given convex feasibility problem is consistent (i.e., $Q := \bigcap_{i \in I} Q_i \neq \emptyset$), then the limit point of the sequence generated by the BIP procedure belongs to Q.

An important novel feature of the BIP scheme is the sequence of variable weight functions $\{w_k | k \in \mathbb{N}_0\}$ and the sequence of variable relaxation parameters $\{\lambda_k | k \in \mathbb{N}_0\}$. Through judicious choice of such sequences a plethora of specific block-iterative algorithms can be obtained. This ability to vary both the sizes of blocks of constraints and the assignment of constraints to blocks throughout iterations (see paragraph 5.3 below) is an indispensable tool for successful parallel implementation of the algorithm. The extra flexibility gained by variable weight functions can then be used to properly spread the computational effort between the parallel processors, thus increasing the overall computational gain.

In this note we study the behavior of "almost simultaneous" BIP procedures, that is BIP procedures in which the sequence of weight functions $\{w_k | k \in \mathbb{N}_0\}$ has a subsequence converging to a weight function w_* with positive values. We characterize the limit points of such BIP procedures (Theorem 3.2) in the case when the convex feasibility problem is consistent as well as in the inconsistent case. In this way we generalize a limit characterization due to Iusem and De Pierro [14] for "simultaneous" BIP procedures (that is almost simultaneous BIP procedures with constant sequences of weight functions and relaxation parameters). Also, we prove a convergence criterion for almost simultaneous BIP procedures in finite dimensional Hilbert spaces (Theorem 4.4) showing that if the BIP procedure is almost simultaneous, then its convergence can be ensured for any sequence of relaxation parameters $\{\lambda_k | k \in \mathbb{N}_0\}$ with $\underline{\lim}_{k \to \infty} \lambda_k > 0$ and $\overline{\lim}_{k \to \infty} \lambda_k < 2$ or under even in weaker conditions if $\mathrm{Int}(Q) \neq \emptyset$. Compared with the main result in [2], our convergence theorem allows a wider range of variation for the relaxation parameters. This is of significance because it was already shown in [14, Section 6] that specific sequences of point-dependent relaxation parameters, which cannot be contained in any closed subinterval of $]0,2[$, have a sensible acceleration effect on BIP procedures even when the weight functions w_k are kept constant (see also the paragraphs 4.6, 4.7, 4.8 and 5.1 below).

The convergence of particular almost simultaneous BIP methods, and especially of "simultaneous" BIP procedures, was studied before and our convergence theorem (Theorem 4.4) can be seen as a generalization of the convergence criteria proved in [4, 10, 11, 12, 14]. Theorem 4.4 improves upon these results by ensuring convergence with variable and less restricted weight functions and relaxation parameters. This becomes relevant when the given convex feasibility problem is defined by a system of convex inequalities, that is, each Q_i is the solution set of an inequality $g_i(x) \leqq 0$ with g_i a convex functional. In such cases, computing orthogonal projections $P_i x^k$, specially when the g_i's are not affine, can be costly and complicated. Almost simultaneous procedures, allowing point-dependent choices of the weight functions and more freedom for the relaxation parameters, can be constructed such that computation of some of the orthogonal projections is systematically avoided by making sure that once the sequence generated by the BIP procedure reaches the interior of a set Q_i at a stage k, then it will not leave this set at any subsequent stage $k+l$. Since, as observed in [11] and [14], the acceleration obtained from the BIP procedure by an adequate choice of the relaxation parameters increases when the set of violated constraints becomes smaller, specific point-dependent choices of the weight functions and of the relaxation parameters may not only avoid computing some of the projections but may also improve the acceleration rate by not allowing the number of nonviolated restrictions to decrease under specific limits (cf. 5.1 below).

The proof of our convergence theorem is made under the assumption that the Hilbert space X has finite dimension. This assumption is employed in order to make sure that bounded sequences of X have convergent subsequences. It can be easily seen that the proof of Theorem 4.4 still holds when the finite dimension requirement is dropped and "convergence" is replaced by "weak convergence". In this way, a generalization of Theorem 1 in [13] and of Theorem 6 in [11] can be obtained.

2. PRELIMINARIES

2.1 In this section we establish several properties of operators and functions which are involved in our study. We recall that each operator P_i is continuous on X and thus, for any weight function w, the *simultaneous projection operator* $\mathbf{P}_w$ is also continuous on X. It follows from the differentiability of the function $\phi_i(x) = \|P_i x - x\|^2$ (see [3, p. 24]) that the function $f_w: X \to \mathbb{R}$ defined by

$$f_w(x) = \sum_{i \in I} w(i) \cdot \|P_i x - x\|^2, \tag{2.1}$$

where $\|\cdot\|$ denotes the norm induced by the inner product $\langle \cdot, \cdot \rangle$ of X, is continuously differentiable on X and its gradient is

$$\nabla f_w(x) = 2 \cdot (x - \mathbf{P}_w x). \tag{2.2}$$

Note that f_w is convex because each ϕ_i is convex.

2.2 The following Proposition summarizes a series of results of Iusem and De Pierro [14].

PROPOSITION *(i) Let w be any weight function. Then the simultaneous projection operator* $\mathbf{P}_w$ *has the following properties:*

$$z \in Q \text{ implies } \mathbf{P}_w z = z; \tag{2.3}$$

$$\|\mathbf{P}_w x - \mathbf{P}_w y\| \leqq \|x - y\|, \qquad (x, y \in X); \tag{2.4}$$

$$x \in X \text{ and } z \in Q \text{ implies } \|\mathbf{P}_w x - z\| \leqq \|x - z\|; \tag{2.5}$$

$$x \in X \text{ and } z \in Q \text{ implies } \langle \mathbf{P}_w x - x, \mathbf{P}_w x - z \rangle \leqq 0; \tag{2.6}$$

$$\langle \mathbf{P}_w x - \mathbf{P}_w y, x - y \rangle \geqq 0, \qquad (x, y \in X). \tag{2.7}$$

ii) If w is a positive weight function, i.e., if $w(i) > 0$ *for all* $i \in I$, *then the set of fixed points of the simultaneous projection operator* $\mathbf{P}_w$ *and the set of minimizers of* f_w *coincide. These sets coincide with Q provided that* $Q \neq \emptyset$.

2.3 Let w be a weight function and $\lambda \in \mathbb{R}$. The *transfer operator* (w.r.t. w and λ) is the function $\mathbf{T}_w^\lambda: X \to X$ defined by

$$\mathbf{T}_w^\lambda x = x + \lambda \cdot (\mathbf{P}_w x - x). \tag{2.8}$$

The transfer operator $\mathbf{T}_{\delta_i}^\lambda$ with respect to the weight function δ_i, having the values $\delta_i(j) = 1$ when $i = j$ and $\delta_i(j) = 0$ when $i \neq j$, is denoted T_i^λ. It is clear that

$$\mathbf{T}_w^\lambda x = \sum_{i \in I} w(i) \cdot T_i^\lambda x. \tag{2.9}$$

2.4 Several properties of the transfer operators are useful for our subsequent analysis.

THEOREM *If w is a positive weight function and if* $\lambda \in]0, 1]$, *then*

a) The transfer operator $\mathbf{T}_w^\lambda$ *is nonexpansive, that is*

$$\|\mathbf{T}_w^\lambda x - \mathbf{T}_w^\lambda y\| \leqq \|x - y\|, \qquad (x, y \in X). \tag{2.10}$$

b) If $x, y \in X$ *then*

$$\|\mathbf{T}_w^\lambda x - \mathbf{T}_w^\lambda y\| = \|x - y\| \Rightarrow \mathbf{T}_w^\lambda x - \mathbf{T}_w^\lambda y = x - y; \tag{2.11}$$

c) If for some z in Q

$$\|\mathbf{T}_w^\lambda x - z\| = \|x - z\|, \tag{2.12}$$

then $x \in Q$.

Proof (a) Since the operator P_i is nonexpansive, we have

$$\|T_i^\lambda x - T_i^\lambda y\| \leqq (1 - \lambda) \cdot \|x - y\| + \lambda \cdot \|P_i x - P_i y\| \leqq \|x - y\|. \tag{2.13}$$

Now, from (2.9) and (2.13) formula (2.10) follows.

b) Assume that $\|\mathbf{T}_w^\lambda x - \mathbf{T}_w^\lambda y\| = \|x-y\|$. According to (2.9) and (2.10) we have

$$\|x-y\| = \|\mathbf{T}_w^\lambda x - \mathbf{T}_w^\lambda y\| \leqq \sum_{i\in I} w(i)\cdot\|T_i^\lambda x - T_i^\lambda y\| \leqq \|x-y\|.$$

This implies, using again (2.10) for T_i^λ, that $\|T_i^\lambda x - T_i^\lambda y\| = \|x-y\|$ and, therefore, that

$$\|x-y\|^2 = (1-\lambda)^2\cdot\|x-y\|^2 + \lambda^2\cdot\|P_i x - P_i y\|^2 + 2\lambda(1-\lambda)\cdot\langle x-y, P_i x - P_i y\rangle \quad (2.14)$$

for each i in I. According to (2.12), the last term on the right-hand side of (2.14) is nonnegative and not greater than $2\cdot\lambda\cdot(1-\lambda)\cdot\|x-y\|\cdot\|P_i x - P_i y\|$. Thus, from (2.14),

$$\|x-y\|^2 \leqq [(1-\lambda)\cdot\|x-y\| + \lambda\cdot\|P_i x - P_i y\|]^2, \qquad (i\in I).$$

Since each P_i is nonexpansive and $\lambda\in]0,1]$, this implies

$$\|x-y\| \leqq (1-\lambda)\cdot\|x-y\| + \lambda\cdot\|P_i x - P_i y\| \leqq \|x-y\|, \qquad (i\in I),$$

that is $\|P_i x - P_i y\| = \|x-y\|$, $(i\in I)$. Consequently, for each $i\in I$, there exists a real number $\alpha_i \geqq 0$ such that $P_i x - P_i y = \alpha_i\cdot(x-y)$ and each α_i is either 0 or 1. In all cases one obtains

$$P_i x - P_i y = x - y, \qquad (i\in I).$$

Summing up all these equalities, each of them weighted by the corresponding $w(i)$, one deduces $\mathbf{P}_w x - \mathbf{P}_w y = x-y$, which implies $\mathbf{T}_w^\lambda x - \mathbf{T}_w^\lambda y = x - y$.

c) Assume that for some $z\in Q$ the equality (2.12) holds. Then, by (2.3), z is a fixed point of $\mathbf{T}_w^\lambda$ and $\|\mathbf{T}_w^\lambda x - \mathbf{T}_w^\lambda z\| = \|x-z\|$, which implies (according to (b)) that $\mathbf{T}_w^\lambda x - \mathbf{T}_w^\lambda z = x - z$ and this is equivalent to $\mathbf{T}_w^\lambda x = x$ because $\mathbf{T}_w^\lambda z = z$. By (2.8) this means that $\mathbf{P}_w x = x$ since $\lambda \neq 0$. Hence, x is a fixed point of $\mathbf{P}_w$ where w is a positive weight function. Therefore, by Proposition 2.2(ii), $x\in Q$. ■

3. A CHARACTERIZATION OF THE LIMIT OF CONVERGENT BIP PROCEDURES

3.1 In this section we show that the BIP algorithm, whenever it converges, leads to a point which is a convex combination of its own projections onto the sets involved in the convex feasibility problem. This point is necessarily a solution of the convex feasibility problem if such a solution exists.

3.2 THEOREM *If the sequence $\{\lambda_k | k \in \mathbb{N}_0\}$ of relaxation parameters of an almost simultaneous BIP procedure satisfies*

$$\inf\{|\lambda_k| \,|k \in \mathbb{N}_0\} = \alpha > 0, \tag{3.1}$$

then for each $x^0 \in X$ for which the sequence $\{x^k | k \in \mathbb{N}\}$ generated by the BIP algorithm converges we have:

a) The limit x^* of $\{x^k | k \in \mathbb{N}\}$ is a fixed point of $\mathbf{P}_{w_*}$.

b) The limit x^* of $\{x^k | k \in \mathbb{N}\}$ is a minimizer of the function f_{w_*}.

c) If $Q \neq \emptyset$, then the limit x^* of the sequence $\{x^k | k \in \mathbb{N}\}$ is contained in Q.

Proof Because the BIP procedure is almost simultaneous the sequence $\{w_k | k \in \mathbb{N}_0\}$ has a subsequence $\{w_{k_p} | p \in \mathbb{N}_0\}$ converging to a positive weight function w_*.

a) According to (2.2) and (3.1),

$$\|\nabla f_{w_k}(x^k)\| = \frac{2}{|\lambda_k|} \cdot \|x^{k+1} - x^k\| \leqq \frac{2}{\alpha} \cdot \|x^{k+1} - x^k\|. \tag{3.2}$$

Letting $k \to \infty$ gives

$$\lim_{k \to \infty} \nabla f_{w_k}(x^k) = 0, \tag{3.3}$$

because the sequence $\{x^k | k \in \mathbb{N}_0\}$ converges. Note that

$$\|\nabla f_{w_k}(x^k) - \nabla f_{w_*}(x^*)\| = 2 \cdot \|(x^k - x^*) + (\mathbf{P}_{w_k}(x^k) - \mathbf{P}_{w_*}(x^*)\|$$

$$\leqq 2 \cdot (\|x^k - x^*\| + \|\mathbf{P}_{w_k}(x^k) - \mathbf{P}_{w_*}(x^*)\|) \tag{3.4}$$

because of (2.2). For each $k \in \mathbb{N}$, we have that

$$\|\mathbf{P}_{w_k}(x^k) - \mathbf{P}_{w_*}(x^*)\| \leqq \|\mathbf{P}_{w_*}(x^k) - \mathbf{P}_{w_*}(x^*)\| + \sum_{i \in I} |w_k(i) - w_*(i)| \, \|P_i x^k\|. \tag{3.5}$$

This shows that $\mathbf{P}_w x$ is continuous as function of (w, x). Therefore, replacing k by k_p in (3.4) and letting $p \to \infty$, we deduce that

$$\lim_{p \to \infty} \nabla f_{w_{k_p}}(x^{k_p}) = \nabla f_{w_*}(x^*). \tag{3.6}$$

From (3.3) and (3.6) we then have that x^* is a stationary point of f_{w_*}. Therefore, x^* is a fixed point of $\mathbf{P}_{w_*}$, by (2.2)

b) From Proposition 2.2(ii), the set of the fixed points of $\mathbf{P}_{w_*}$ coincides with the set of minimizers of f_{w_*}. Hence, by (a), x^* is a minimizer of f_{w_*}.

c) According to Proposition 2.2(ii), Q coincides with the set of fixed points of $\mathbf{P}_{w_*}$ when $Q \neq \emptyset$. Therefore, if $Q \neq \emptyset$, then $x^* \in Q$. This completes the proof of the Theorem. ■

3.3 Under the conditions of Theorem 3.2 it is interesting, for computational purposes, to know how far a point x^k is from the set Q. An estimate of the distance $d(x^k, Q) = \inf\{\|x^k - z\| \mid z \in Q\}$ of x^k from Q may be used to formulate stopping rules for computer programs implementing the BIP algorithm. The following result provides such an estimate.

Proposition *For any sequence of weight functions $\{w_k \mid k \in \mathbb{N}_0\}$ and for any sequence $\{\lambda_k \mid k \in \mathbb{N}_0\}$ of nonnegative relaxation parameters, we have*

$$d^2(x^{k+1}, Q) \leqq d^2(x^0, Q) - \sum_{i=1}^{k} \lambda_i \cdot (2 - \lambda_i) \cdot \|\mathbf{P}_{w_i} x^i - x^i\|^2. \tag{3.7}$$

Proof Let x be any point in X. Denote $y = \mathbf{T}_w^\lambda x$. Then for any z in Q we have

$$\begin{aligned} 0 \leqq \|y - z\|^2 &= \|(x - z) + \lambda \cdot (\mathbf{P}_w x - x)\|^2 \\ &= \|x - z\|^2 + \lambda^2 \cdot \|\mathbf{P}_w x - x\|^2 + 2\lambda \cdot [\langle \mathbf{P}_w x - z, \mathbf{P}_w x - x\rangle - \|x - \mathbf{P}_w x\|^2] \\ &= \|x - z\|^2 + \lambda(\lambda - 2) \cdot \|\mathbf{P}_w x - x\|^2 + 2 \cdot \lambda \cdot \langle \mathbf{P}_w x - z, \mathbf{P}_w x - x\rangle. \end{aligned} \tag{3.8}$$

This implies

$$d^2(y, Q) \leqq d^2(x, Q) - \lambda \cdot (2 - \lambda) \cdot \|\mathbf{P}_w(x) - x\|^2 \tag{3.9}$$

because the last inner product in (3.8) is nonpositive (cf. (2.6)). Writing (3.9) for each x^i with i from 1 to k and summing up the resulting inequalities yields (3.7). ■

4. A CONVERGENCE ANALYSIS

4.1 The main differences among various convergence theorems for special cases of BIP procedures discussed in the literature are the different ways in which the weight functions and the relaxation parameters are restricted. A commonly used restriction concerning the relaxation parameters is that the whole sequence $\{\lambda_k \mid k \in \mathbb{N}_0\}$ is contained in a closed subinterval of $]0,2[$. It appears first in [1]. It was observed in [18] that allowing the relaxation parameters to converge to 2 may accelerate the convergence of sequential relaxation algorithms. In what follows we prove that something similar happens for almost simultaneous BIP procedures.

4.2 The next result will be of use:

Proposition *Let w be a weight function. If $x \in X$ and if $z \in Q$, then we have:*

i) The function

$$h_w(\alpha, x, z) = \|\mathbf{T}_w^\alpha x - z\|^2 \tag{4.1}$$

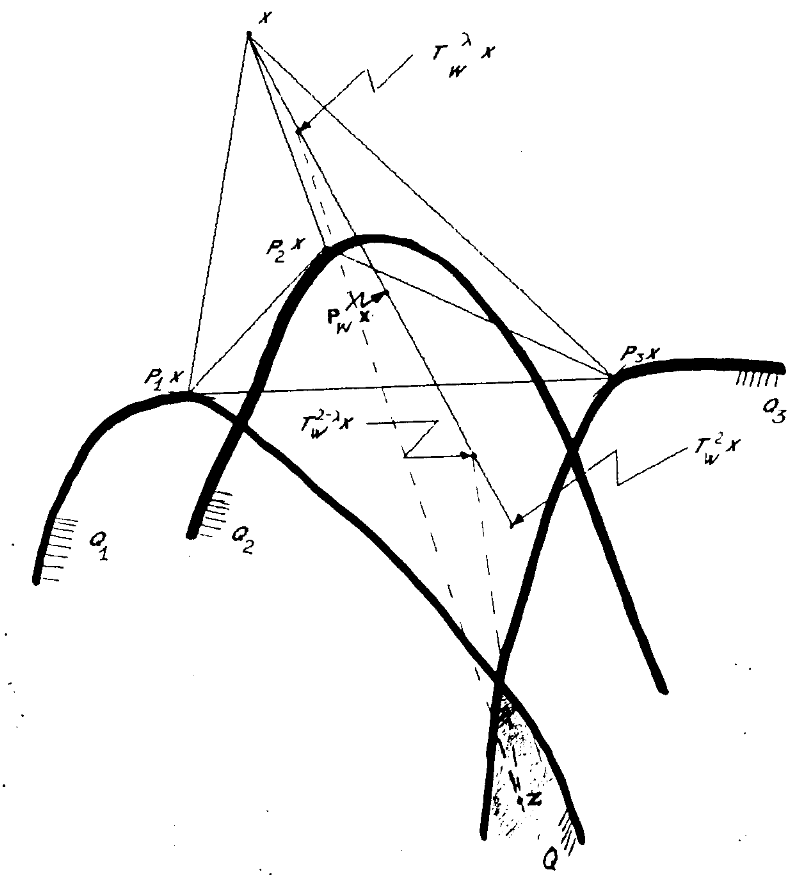

Figure 1

is nondecreasing with respect to α *on the interval* $[0,1]$.

ii) If $\lambda\in[1,2]$, *then*

$$\|\mathbf{T}_w^\lambda x-z\|\leqq\|\mathbf{T}_w^{2-\lambda}x-z\|\leqq\|x-z\|. \tag{4.2}$$

iii) If $\lambda\in[0,1]$, *then*

$$\|\mathbf{T}_w^{2-\lambda}x-z\|\leqq\|\mathbf{T}_w^\lambda x-z\|. \tag{4.3}$$

Proof (i) Using (2.8) and (4.1) we obtain that

$$\begin{aligned} &h_w(\alpha,x,z)-h_w(\beta,x,z)\\ &=(\alpha-\beta)\cdot[(\alpha+\beta-2)\cdot\|x-z\|^2+(\alpha+\beta)\cdot\|\mathbf{P}_wx-z\|^2+2\cdot\langle x-z,\mathbf{P}_wx-z\rangle]\\ &=(\alpha-\beta)\cdot[(\alpha+\beta)\cdot\|\mathbf{P}_wx-x\|^2+2\cdot\langle x-z,\mathbf{P}_wx-z\rangle-2\cdot\|x-z\|^2]\\ &=(\alpha-\beta)\cdot[(\alpha+\beta-1)\cdot\|\mathbf{P}_wx-x\|^2+\|\mathbf{P}_wx-z\|^2-\|x-z\|^2], \end{aligned} \tag{4.4}$$

for any $\alpha, \beta \in \mathbb{R}$, $x \in X$ and $z \in Q$. If $1 \geqq \alpha > \beta \geqq 0$, then (4.4) implies

$$h_w(\alpha, x, z) - h_w(\beta, x, z) \leqq (\alpha - \beta) \cdot [\|\mathbf{P}_w x - x\|^2 + \|\mathbf{P}_w x - z\|^2 - \|x - z\|^2]$$

where, according to Auslender [4, Theorem V.1.1], we have

$$\|x - z\|^2 \geqq \|\mathbf{P}_w x - z\|^2 + \|\mathbf{P}_w x - x\|^2, \tag{4.5}$$

i.e., the right hand side is nonpositive.

ii) Let $z \in Q$. Then, according to Theorem 2.4(a), we have

$$\|\mathbf{T}_w^{2-\lambda} x - z\| = \|\mathbf{T}_w^{2-\lambda} x - \mathbf{T}_w^{2-\lambda} z\| \leqq \|x - z\|,$$

Note that for $\alpha = \lambda$ and $\beta = 2 - \lambda$ formula (4.4) still holds. Using this fact and (4.5), the first inequality in (4.2) follows. This situation is demonstrated by the drawing in Figure 1.

The proof of (iii) is similar to that of (ii). ■

4.3 Combining Proposition 4.2(ii) and Theorem 2.4(a) we obtain the following well-known result ([11, 2]).

COROLLARY *Let w be a weight function. If $\lambda \in [0, 2]$, then*

$$\|\mathbf{T}_w^{\lambda} x - z\| \leqq \|x - z\|, \qquad (x \in X, z \in Q). \tag{4.6}$$

4.4 The convergence theorem now follows.

THEOREM *Suppose that X is a finite-dimensional Hilbert space and $Q \neq \emptyset$. If the sequence $\{\lambda_k | k \in \mathbb{N}_0\}$ has $0 < \underline{\lim}_{k \to \infty} \lambda_k$ and $\overline{\lim}_{k \to \infty} \lambda_k < 2$, then any almost simultaneous BIP procedure with these relaxation parameters converges to a point in Q, regardless of the choice of the initial point x^0. Moreover, if we have also $Int(Q) \neq \emptyset$, then the result still holds when $\overline{\lim}_{k \to \infty} \lambda_k = 2$, provided that all but finitely many λ_k's are in the interval $]0, 2]$.*

Proof According to the hypothesis, the sequence $\{\lambda_k | k \in \mathbb{N}_0\}$ has all but finitely many of its terms in $[0, 2]$. Therefore, according to Corollary 4.3, there exists a $k^* \in \mathbb{N}_0$ such that

$$\|x^{k+1} - z\| \leqq \|x^k - z\|, \qquad (k \geqq k^*), \tag{4.7}$$

for any $z \in Q$. Since $Q \neq \emptyset$, this implies that the sequence $\{x^k | k \in \mathbb{N}_0\}$ is bounded. Consequently, it has a convergent subsequence $\{x^{k_p} | p \in \mathbb{N}_0\}$ such that the subsequence $\{w_{k_p} | p \in \mathbb{N}_0\}$ of $\{w_k | k \in \mathbb{N}_0\}$ converges to the positive weight function w_* and the subsequence $\{\lambda_{k_p} | p \in \mathbb{N}_0\}$ of $\{\lambda_k | p \in \mathbb{N}_0\}$ converges to a number $\lambda_* \in]0, 2[$. Let x^* be the limit of $\{x^{k_p} | p \in \mathbb{N}_0\}$. Without loss of generality, we may assume that all the weight functions w_{k_p} are positive. Observe that, for any z in Q, the sequence $\{\|x^k - z\| \mid k \in \mathbb{N}_0\}$ is nonincreasing and bounded, hence convergent. Thus, it converges necessarily to $\|x^* - z\|$. Two cases have to be distinguished according to the value of λ_*.

Case 1 Assume that $\lambda_* \in]0,1]$. Fix z arbitrarily in Q. Note that $\mathbf{T}_w^\lambda x$ is continuous as a function of (λ, w, x). Therefore,

$$\lim_{p\to\infty} \mathbf{T}_{w_{k_p}}^{\lambda_{k_p}}(x^{k_p}) = \mathbf{T}_{w_*}^{\lambda_*}(x^*). \tag{4.8}$$

Since $\{\|x^k - z\| \mid k \in \mathbb{N}_0\}$ is convergent any of its subsequences has the same limit $\|x^* - z\|$; thus

$$\|x^* - z\| = \lim_{p\to\infty} \|x^{k_p+1} - z\| = \lim_{p\to\infty} \|\mathbf{T}_{w_{k_p}}^{\lambda_{k_p}}(x^{k_p}) - z\| = \|\mathbf{T}_{w_*}^{\lambda_*}(x^*) - z\|,$$

because of (4.8). The point z is a fixed point of $\mathbf{T}_{w_*}^{\lambda_*}$. Thus, we have $\|\mathbf{T}_{w_*}^{\lambda_*}(x^*) - \mathbf{T}_{w_*}^{\lambda_*}(z)\| = \|x^* - z\|$. According to Theorem 2.4(c), this implies $x^* \in Q$, because $\lambda_* \in]0,1]$. This means that (4.7) still holds for $z = x^*$. Hence, the sequence $\{\|x^k - x^*\| \mid k \in \mathbb{N}_0\}$ converges to the same limit as its subsequence $\{\|x^{k_p} - x^*\| \mid p \in \mathbb{N}_0\}$. This shows that $x^k \to x^*$, as $k \to \infty$, and the proof is complete in this case.

Case 2 Assume that $\lambda_* \in]1,2[$. Then, we assume, without loss of generality, that all relaxation parameters λ_{k_p} belong to the interval $]1,2[$. According to (4.7) and to Proposition 4.2(ii) we have

$$\begin{aligned} \|x^{k_{p+1}} - z\| \leqq \|x^{k_p+1} - z\| &= \|\mathbf{T}_{w_p}^{\lambda_p}(x^{k_p}) - z\| \\ &\leqq \|\mathbf{T}_{w_p}^{2-\lambda_p}(x^{k_p}) - z\| \leqq \|x^{k_p} - z\|, \end{aligned} \tag{4.9}$$

for any $z \in Q$ and for any $p \in \mathbb{N}_0$. Fix z arbitrarily in Q. Taking $p \to \infty$ in (4.9), we obtain

$$\|\mathbf{T}_{w_*}^{2-\lambda_*}(x^*) - \mathbf{T}_{w_*}^{2-\lambda_*}(z)\| = \|x^* - z\|, \tag{4.10}$$

because $\mathbf{T}_{w_p}^{2-\lambda_p}(x^{k_p}) \to \mathbf{T}_{w_*}^{2-\lambda_*}(x^*)$, as $p \to \infty$, for the same reasons as in Case 1. Since $2 - \lambda_* \in]0,1[$, we apply Theorem 2.4(c) to $\mathbf{T}_{w_*}^{2-\lambda_*}$. Together with (4.10), it implies that $x^* \in Q$. Using (4.7) with $z = x^*$, we deduce that $\{\|x^k - x^*\| \mid k \in \mathbb{N}_0\}$ converges and it must have the same limit as its subsequence $\{\|x^{k_p} - x^*\| \mid p \in \mathbb{N}_0\}$. This shows that $x^k \to x^* \in Q$, as $p \to \infty$. The proof of the Theorem is complete in this case too.

Now assume that $\mathrm{Int}(Q) \neq \emptyset$ and $\overline{\lim}_{k\to\infty} \lambda_k = 2$ and $\lambda_k \in]0,2]$, $(k \in \mathbb{N}_0)$. Furthermore, we assume without loss of generality that there exists a convergent subsequence $\{x^{k_p} \mid p \in \mathbb{N}_0\}$ such that the corresponding subsequences $\{w_{k_p} \mid k \in \mathbb{N}_0\}$ and $\{\lambda_{k_p} \mid k \in \mathbb{N}_0\}$ converge to w_* and $\lambda_* = 2$, respectively, and such that the latter subsequence is contained in $]1,2]$. Using again (4.7), we obtain

$$\|x^{k_{p+1}} - z\| \leqq \|x^{k_p+1} - z\| \leqq \|x^{k_p} - z\|, \tag{4.11}$$

for any $p \in \mathbb{N}_0$ and for any $z \in Q$. Let x^* be the limit of the subsequence $\{x^{k_p} | k \in \mathbb{N}_0\}$. Taking $p \to \infty$ in (4.11) we deduce

$$\lim_{p \to \infty} \|x^{k_p+1} - z\| = \|x^* - z\|, \tag{4.12}$$

for any $z \in Q$. Since $\{x^{k_p+1} | p \in \mathbb{N}_0\}$ is bounded (cf. (4.7)) and the Hilbert space X is finite-dimensional, it follows that the sequence $\{x^{k_p+1} | k \in \mathbb{N}_0\}$ has a convergent subsequence. We denote this subsequence by $\{x^{k_p+1} | k \in \mathbb{N}_0\}$ too. Let x^{**} be its limit. According to (4.12) we have

$$\|x^{**} - z\| = \|x^* - z\|, \qquad (z \in Q).$$

This implies that Q is contained in the hyperplane defined by

$$2 \cdot \langle x^{**} - x^*, x \rangle = \|x^{**} - x^*\|^2.$$

According to the hypothesis this is not possible except when $x^{**} = x^*$. This means that the sequences $\{x^{k_p+1} | k \in \mathbb{N}_0\}$ and $\{x^{k_p} | k \in \mathbb{N}_0\}$ have the same limit x^*. Note that, for any $p \in \mathbb{N}_0$, we have

$$0 \leqq \|\mathbf{P}_{w_{k_p}} x^{k_p} - x^{k_p}\| = (\lambda_{k_p})^{-1} \cdot \|x^{k_p+1} - x^{k_p}\| \leqq \|x^{k_p+1} - x^{k_p}\|,$$

because all the relaxation parameters λ_{k_p} belong to $]1,2]$. Thus, taking $p \to \infty$ in the last inequality, it yields that $\|\mathbf{P}_{w_{k_p}} x^{k_p} - x^{k_p}\|$ converges to zero, as $p \to \infty$. But $\mathbf{P}_{w_{k_p}} x^{k_p} \to \mathbf{P}_{w_*} x^*$ as $p \to \infty$, because $w_{k_p} \to w_*$ and $x^{k_p} \to x^*$, as $p \to \infty$, and because (3.5) still holds when k is replaced in it by k_p. Consequently, we have that $\|\mathbf{P}_{w_*} x^* - x^*\| = 0$, i.e., x^* is a fixed point of $\mathbf{P}_{w_*}$ where w_* is a positive weight function. Then, according to Proposition 2.2(ii), we have that $x^* \in Q$. Hence, (4.7) still holds for $z = x^*$. This shows that the sequence $\{\|x^k - x^*\| \,|\, k \in \mathbb{N}_0\}$ converges. Therefore, it must have the same limit as its subsequence $\{\|x^{k_p} - x^*\| \,|\, p \in \mathbb{N}_0\}$, that is $\|x^k - x^*\| \to 0$, as $k \to \infty$. Hence, $\{x^k | k \in \mathbb{N}_0\}$ converges to the point $x^* \in Q$ and the proof of Theorem is complete. ■

4.5 Observe that even if $\mathrm{Int}(Q) \neq \emptyset$, if $\{\lambda_k | k \in \mathbb{N}_0\}$ converges to 0, then the BIP algorithm may not converge or it may converge to a point which is not a solution of the given convex feasibility problem when such solutions exist.

Example Consider $X = \mathbb{R}^2$. Let Q_1 be the rectangle $ABCD$, let Q_2 be the rectangle $A'B'C'D'$ and let $x^0 = x$ be a point as shown in Figure 2. Choose $w_k = w = (0.5, 0.5)$ for all $k \in \mathbb{N}_0$. If at each step k the number λ_k is chosen such that $0 < \lambda_k \cdot \|\mathbf{P}_{w_k} x^k - x^k\| < \|x^k - P_1 x^k\|$, then each x^{k+1} will be a point x' on the line segment $[x, P_1 x]$ and $\mathbf{P}_w x^k$ will be invariably equal to $\mathbf{P}_w x^0$. Since the line segment $[x, P_1 x]$ is closed, any accumulation point of the sequence $\{x^k | k \in \mathbb{N}_0\}$ is contained

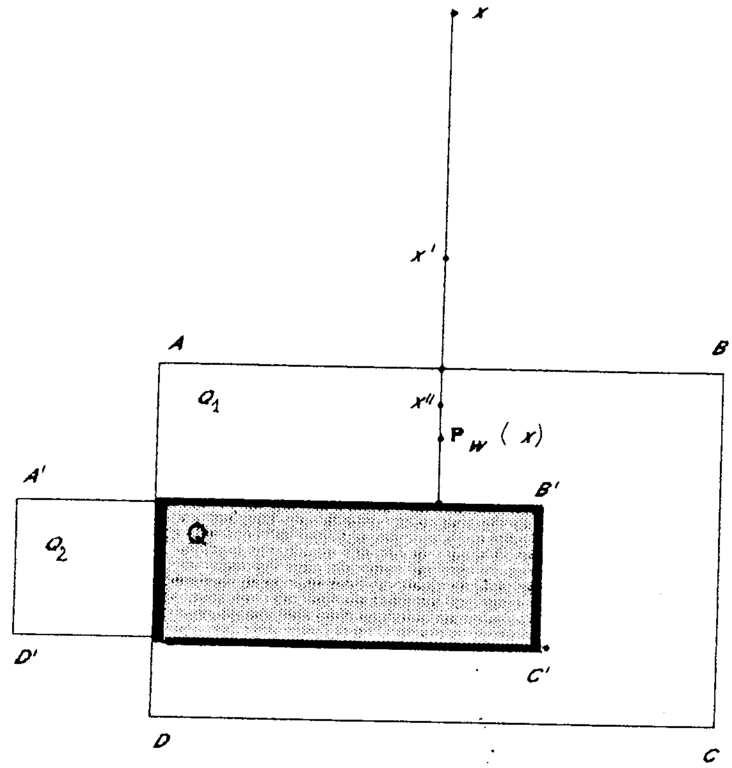

Figure 2

in it. Hence, the sequence generated by the BIP algorithm in such a situation does not converges to a point in Q.

4.6 In specific situations, replacing the sequence of relaxation parameters $\{\lambda_k | k \in \mathbb{N}_0\}$ by another sequence $\{\tilde{\lambda}_k | k \in \mathbb{N}_0\}$ with $\lambda_k \leqq \tilde{\lambda}_k$, $(k \in \mathbb{N}_0)$, may improve considerably *the initial behavior* of the BIP method, that is the number of iterations needed in order to reach a point within a ε-neighborhood of Q may be much smaller when the relaxation parameters $\tilde{\lambda}_k$ are used than with λ_k. If $0 \leqq \lambda_k < \tilde{\lambda}_k \leqq 1$, for all $k \in \mathbb{N}_0$, then this clearly follows from Proposition 4.2(i). The following example shows how the initial behavior of the BIP method may vary when the sequence of relaxation parameters involved is changed.

Example Let $X = \mathbb{R}^2$. Consider the convex feasibility problem of computing a point contained in the set $Q = Q_1 \cap Q_2 \cap Q_3$ where $Q_1 := \{x | 3 \cdot x_1 - 4 \cdot x_2 + 12 \leqq 0\}$, $Q_2 := \{x | 5 \cdot x_1 + 12 \cdot x_2 + 20 \leqq 0\}$ and $Q_3 := \{x | x_1 \leqq -5\}$. We choose $x^0 = (0, 5)$ and all $w_k = (1/3, 1/3, 1/3)$. We use the BIP algorithm with different constant sequences of relaxation parameters and stop the procedure when the distance of x^{k+1} to Q is less than 10^{-6}. In this way, we observe that for $\lambda_k = \alpha$, $(k \in \mathbb{N}_0)$, the procedure achieves a point x^s whose distance to Q is less than 10^{-6} after a number of iterations s which strongly depends on α. The results are summarized in the following table:

$\alpha=$	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80	2.00
$s=$	182	88	56	41	35	25	20	17	14	11
$x_1^s=$	−5.07	−5.08	−5.09	−5.10	−5.12	−5.13	−5.15	−5.18	−5.21	−5.26
$x_2^s=$	0.449	0.452	0.456	0.460	0.465	0.472	0.480	0.492	0.508	0.523

The values of the coordinates x_1^s and x_2^s in this table are truncated since this precision is enough to observe the variation of x^s. A similar dependence of the number of steps s required to reach (with a given error) a point in Q on the relaxation parameters can be observed when various nonconstant sequences $\{\lambda_k | k \in \mathbb{N}_0\}$ are employed.

4.7 According to Theorem 4.4, if $\mathrm{Int}(Q) \neq \emptyset$, then the BIP algorithm generates a sequence which converges to a point in Q, even if the sequence of relaxation parameters converges to 2. In this case, the condition $\mathrm{Int}(Q) \neq \emptyset$ is essential. This fact is illustrated by the next.

Example Let $X = \mathbb{R}^3$ and let Q_1 and Q_2 be two rectangles situated in the same plane as shown in Figure 3. Let $\{w_k | k \in \mathbb{N}_0\}$ be any sequence of weight functions and let all relaxation parameters λ_k be equal to 2. Then, if the initial point x^0 is outside of the plane of Q_1 and Q_2 but on a line which is orthogonal to this plane at a point in the set $Q_1 \cap Q_2$, then the sequence generated by the BIP algorithm oscillates between x^0 and x^1 but does not converge.

4.8 Proposition 4.2(iii) combined with Corollary 4.3 show that the initial behavior of any BIP procedure with all relaxation parameters in $[0,2]$ can be eventually improved by replacing the relaxation parameters λ_k contained in $[0,1]$ by relaxation parameters $\tilde{\lambda}_k = 2 - \lambda_k$ which are closer to 2. This explains the phenomenon observed in Example 4.6.

5. PRACTICAL CONSIDERATIONS

5.1 Consider the convex feasibility problem in which each Q_i is the set of solutions of an inequality $g_i(x) \leqq 0$ where g_i is a convex functional on X. Using Theorems 4.4 and 3.2 we can prove the convergence of the BIP procedure having an eventually unbounded sequence of relaxation parameters when the weight functions are defined as follows. For each $k \in \mathbb{N}_0$ denote $I_k = \{i \in I | g_i(x^k) < 0\}$, $L_k = \{i \in I | g_i(x^k) \leqq 0\}$ and $J_k = I \setminus L_k$ and define

$$
\begin{aligned}
v_k &= \max_{(i,j) \in I_k \times J_k} \{g_i(P_j x^k)/g_i(x^k)\}, \quad \text{if } I_k \times J_k \neq \emptyset, \\
&= 1 \qquad\qquad\qquad\qquad\qquad\qquad \text{otherwise.}
\end{aligned}
$$

Choose $\alpha_k \in]0,1]$ such that $(\alpha_k - 1)/\alpha_k \in]-\infty, v_k[$ and $\lambda_k = \alpha_k^{-1}$. Define the weight

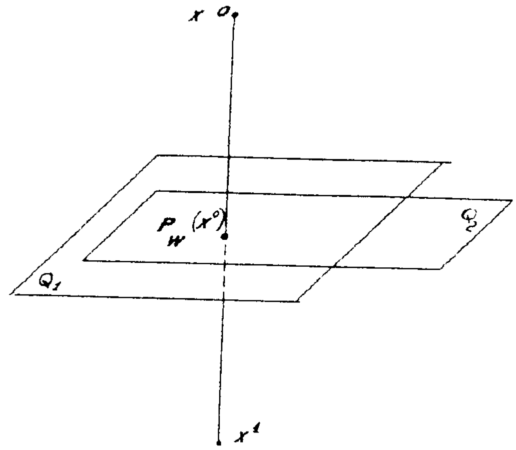

Figure 3

function w_k as follows. For any $j \in L_k$ let $w_k(j) = (1-\alpha_k)/|J_k|$. For the remaining indices j two cases have to be distinguished. If there exists $(i_0, j_0) \in (L_k \setminus I_k) \times J_k$ such that $g_{i_0}(P_{j_0} x^k) < 0$, then let $w_k(j_0) = \alpha_k$ and $w_k(j) = 0$ for $j \in J_k \setminus \{j_0\}$; otherwise, let $w_k(j) = \alpha_k / |J_k|$ for all $j \in J_k$.

For any $k \in \mathbb{N}_0$, $I_k \subseteq I_{k+1}$ and, when the situation described in the first case above holds, then $I_{k+1} \supset I_k$. If one denotes

$$L^* = \{i \in I \,|\, \exists\, k(i) \in \mathbb{N}_0 : g_i(x^k) \leqq 0,\ (k \geqq k(i)\},$$

then either $L_k = L^* = I$ for some k and then finite convergence occurs, or there exists $k^* \in \mathbb{N}_0$ such that for any $k \geqq k^*$ the term x^k generated by the procedure coincides with the corresponding iterate generated by the BIP algorithm $\mathbf{T}^1_{w_k'}$ applied to the convex feasibility problem $\{Q_i | i \in J^*\}$ with $J^* = I \setminus L^*$, where $w_k' : J^* \to \mathbb{R}$ is defined by $w_k'(j) = w_k(j)/\alpha_k$, if $j \in J_k$, and $w_k'(j) = 0$, otherwise. For the BIP algorithm $\mathbf{T}^1_{w_k'}$ Theorem 4.4 holds, therefore the sequence $\{x^k | k \in \mathbb{N}_0\}$ generated by it (and implicitly the sequence generated by the BIP procedure $\mathbf{T}^{\lambda_k}_{w_k}$ described above) converges. Since Theorem 3.2 applies to the BIP procedure $\mathbf{T}^{\lambda_k}_{w_k}$, it follows that its limit has to be in Q, provided that $Q \neq \emptyset$.

The BIP procedure $\mathbf{T}^{\lambda_k}_{w_k}$ described above consists of determining the barycenter of the vectors $\{P_i x^k | i \in J_k\}$ at each iterative step at which no new element can be added to I_k. If $i \in I_k$, then the set Q_i is practically eliminated from the computational process in all subsequent steps. Note that, when the functions g_i are continuously differentiable, $v_k \to 1$ and that, according to (3.9), $d^2(x^{k+1}, Q)$ is smaller than $d^2(x^k, Q)$ by at least the quantity

$$R_k = \left\| \sum_{j \in J_k} (P_j x^k - x^k) \right\| \cdot (2 - \lambda_k)/|J_k|.$$

Therefore, when k is large enough, α_k can be chosen close to 1 and this will sensibly accelerate the convergence by the combined effect of a small λ_k and of a small $|J_k|$.

5.2 Various convergence criteria for the BIP algorithm (our above results included) ensure that the procedure generates a sequence whose limit (or, when X is not finite-dimensional, weak limit) exists and belongs to Q (if Q is nonempty). However, when the BIP algorithm is used in order to solve real-world convex feasibility problems one has to stop the computational procedure after finitely many steps and the question arises how far the last computed x^{k+1} is from the set Q. From the user's point of view it is interesting to know an upper-bound of $d(x^{k+1}, Q)$. Proposition 3.3 shows how such an upper-bound can be computed *if an estimate of the distance of* x^0 *to* Q *is already known.* It is even better, for practical purposes, to have an upper-bound for the distance between x^{k+1} and the limit x^* of the sequence generated by the iterative projection algorithm but, as far as we know, *determining such an "error upper-bound" for the iterative projection algorithm is an open problem* (*even in very particular cases as those concerning sets* Q_i *defined by linear inequalities*)—see [16].

5.3 The BIP algorithm can be programmed for block-parallel or block-serial processing (cf. [10, 14, 7]). Such implementations can be designed by taking advantage of the variable weight functions. Assume that, at step k of the BIP procedure $\alpha_1 < \cdots < \alpha_q$ are all the values taken by w_k. Denote $J_h^k = \{i \in I \mid w_k(i) = \alpha_h\}$ *the block of constraints having weight* α_h at step k. The sums $\sum_{i \in J_h^k} P_i x^k$ can be processed in parallel and x^{k+1} is a linear combination of them and of x^k. The blocks may be dynamically changed from one step to another by appropriate choices of the weight functions. Under the conditions of Theorem 4.4 for the relaxation parameters, the cyclic BIP procedure with $w_k(j) = \delta_{k \bmod (m+1)}(j)$ if $k \bmod (m+1) \neq 0$ and $w_k(j) = 1/m$ otherwise, where $m = |I|$ can be efficiently implemented on small computers.

References

[1] S. Agmon, The relaxation method for linear inequalities, *Canadian Journal of Mathematics* **6** (1954), 382–392.

[2] R. Aharoni and Y. Censor, Block-iterative projection methods for parallel computation of solutions to convex feasibility problems, *Linear Algebra And Its Applications* **120** (1988), 165–175.

[3] J. P. Aubin and A. Cellina, *Differential Inclusions*, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo (1984).

[4] A. Auslander, *Optimization—Méthodes Numériques*, Masson, Paris, New York, Barcelona, Milan (1976).

[5] D. Butnariu and Y. Censor, On a class of bargaining schemes for points in the cores of n-person cooperative games, Preprint, June (1989).

[6] Y. Censor, Row-action methods for huge and sparse systems and their applications, *SIAM Review* **23** (1981), 444–466.

[7] Y. Censor, Parallel applications of block-iterative methods in medical imaging and radiation therapy, *Mathematical Programming* **42** (1988), 307–325.

[8] Y. Censor and T. Elfving, New methods for linear inequalities, *Linear Algebra And Its Applications* **42** (1981), 199–211.

[9] G. Cimmino, Calcolo approsimato per le solutzioni di sistemi di equazioni lineari, *La Riserca Scientifica, Roma*, XVI, Anno IX **2** (1938), 326–333.

[10] A. R. De Pierro and A. N. Iusem, A simultaneous projection method for linear inequalities, *Linear Algebra And Its Applications* **64** (1985), 243–253.

[11] A. R. De Pierro and A. N. Iusem, A parallel projection method of finding common points of a family of convex sets, *Pesquisa Operacional* **5** (1985), 1–20.

[12] J. L. Goffin, The relaxation method for solving systems of linear inequalities, *Mathematics of Operation Research* **5** (1980), 388–414.

[13] C. 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. Computational Mathematics and Mathematical Physics* **7** (1966), 1–24.

[14] A. N. Iusem and A. R. De Pierro, Convergence results for an accelerated nonlinear Cimmino algorithm, *Numerische Mathematik* **49** (1986), 347–368.

[15] S. Kaczmarz, Angenaherte auflösung von systemen linearer gleichungen, *Bull. Acad. Polon. Sci. Lett.* **A35** (1937), 355–357.

[16] S. Kayalar and H. L. Weinert, Error bounds for the method of alternating projections, *Mathematics of Control, Signals and Systems* **1** (1988), 43–59.

[17] J. Mandel, Convergence for the cyclical relaxation method for linear inequalities, *Mathematical Programming* **30** (1984), 218–228.

[18] T. S. Motzkin and I. J. Scheönberg, The relaxation method for linear inequalities, *Canadian Journal of Mathematics* **6** (1954), 393–404.