Convergence of sequential and asynchronous nonlinear paracontractions

L. Elsner[1,*]**, I. Koltracht**[2,**]**, and M. Neumann**[2,***]
[1] Fakultät für Mathematik, Universität Bielefeld, Postfach 8640,
W-4800-Bielefeld 1, Federal Republic of Germany
[2] Department of Mathematics, University of Connecticut, Storrs, CT 06269-3009, USA

Summary. We establish the convergence of sequential and asynchronous iteration schemes for nonlinear paracontracting operators acting in finite dimensional spaces. Applications to the solution of linear systems of equations with convex constraints are outlined. A first generalization of one of our convergence results to an infinite pool of asymptotically paracontracting operators is also presented.

Mathematics Subject Classifications (1991): MSC 1991, 65F10

1 Introduction

In this paper we study convergence of the iteration

$$(1.1) \qquad x_i = T_{j_i}(x_{i-1}), \quad i = 1, 2, \ldots,$$

where T_{j_i}'s are chosen from a finite pool of nonlinear operators $T_1, T_2, \ldots, T_n$ acting on the Euclidean space $\mathbb{R}^k$, and the convergence of a parallel asynchronous version of (1.1). Such iterations appear in applications to signal processing, system theory, computed tomography and other areas (see, for example, Koltracht and Lancaster [4], and Ortega and Rheinboldt [6], and references contained therein).

In Sect. 2 we introduce a class of operators which we call paracontracting, namely, continuous operators $T: \mathbb{R}^k \to \mathbb{R}^k$ such that for any fixed point $y = T(y)$ and any $x \in \mathbb{R}^k$ either $\| T(x) - y \| < \| x - y \|$ or $T(x) = x$. This class contains, in

particular, all strictly nonexpansive operators studied in [4] and a certain subclass of the strictly nonexpansive operators studied by De Pierro and Iusem in [2]. In the case when T is linear our definition coincides with that of Nelson and Neumann [5]. We also give some examples of useful paracontractions. One such example is $T: \mathbb{R}^k \to \mathbb{R}^k$, where

$$T: x \to \begin{cases} x - \omega \dfrac{S(x) - f}{\|S'(x)\|_2^2} S'(x) & \text{if } S(x) > f \\ \\ x & \text{if } S(x) \leqq f \end{cases}$$

and where $S: \mathbb{R}^k \to \mathbb{R}$ is a convex continuously differentiable function and $\omega \in (0, 2)$. In this case T is the relaxed Gauss-Newton iteration operator for S as defined in Ortega and Rheinboldt, see [6, Sect. 8.5].

In Sect. 3 we prove our main results concerning the convergence of (1.1) for paracontracting operators. First we show in Theorem 1 that the iteration (1.1) converges if and only if there is a common fixed point of those T_i's which appear infinitely often in the sequence $\{T_{j_i}\}_{i=1}^{\infty}$ and that in this case the limit is one such fixed point. In Theorem 2 we prove a similar result for an asynchronous version of the iteration which is useful for practical implementations on multiple instruction multiple data (MIMD) parallel computers.

In Sect. 4 we demonstrate how Theorems 1 and 2 can be used for computing a solution of a linear system of equations which also belongs to an intersection of finitely many convex sets. The convex sets can be defined either by explicitly given projection operators or by the condition $\{x: S(x) \leqq f\}$, where S is a convex continuously differentiable function and f is a number. In Sect. 5 we give our first result (Theorem 3) for the convergence of (1.1) for an infinite pool of operators. We introduce the notion of an asymptotically paracontracting sequence of nonlinear operators for which the statement of Theorem 1 remains valid. Our Theorem 3 is of the same flavor as a recent result of Youla [7], for an infinite pool of relaxed linear projection operators.

We would like to add that our results apply to operators acting in the complex space C^k, and that some of them can be reproduced in the Hilbert space setting. We also indicate in this paper lines for future research.

2 Paracontracting operators

Definition 1. A continuous operator $T: \mathbb{R}^k \to \mathbb{R}^k$ is *paracontracting* (*with respect to a vector norm* $\|\cdot\|$ *on* $\mathbb{R}^k$), if for any fixed point (f.p.) $y \in \mathbb{R}^k$, $y = T(y)$ and any $x \in \mathbb{R}^k$

$$\|T(x) - y\| < \|x - y\| \qquad \text{or} \qquad T(x) = x .$$

We remark that for T linear this coincides with the definition given in [5]. A somewhat stronger notion for nonlinear operators is used in [4]:

Definition 2. An operator $T: \mathbb{R}^k \to \mathbb{R}^k$ is *strictly nonexpansive*, if for any pair $x, y \in \mathbb{R}^k$

$$\|T(x) - T(y)\| < \|x - y\| \qquad \text{or} \qquad T(x) - T(y) = x - y$$

holds.

Obviously, a strictly nonexpansive operator is paracontracting, while the inverse implication does not hold, see Example 4 below. The definition of a paracontracting operator does not imply the existence of a fixed point, so that any continuous operator without fixed points has to be considered paracontracting. A product of two or more paracontracting operators need not be paracontracting itself. On the other hand a product of a finite number of strictly nonexpansive operators is strictly nonexpansive. Let us consider some examples:

Example 1. Let $r, g \in \mathbb{R}^k$, $r \neq 0$, $0 < \omega < 2$, and

$$Tx := \left(I - \omega \frac{rr^\mathsf{T}}{r^\mathsf{T}r}\right) x + g .$$

Then T is strictly nonexpansive and hence paracontracting with respect to the Euclidean vector norm $\|\cdot\|_2$.

Example 2. Let $B \subset \mathbb{R}^k$ be a closed convex set and Q the projection onto B, i.e., for any x:

$$Q(x) \in B, \|x - Q(x)\|_2 < \|x - y\|_2 \quad \text{for all } y \neq Q(x), y \in B .$$

We claim that the relaxed projection operator

$$Q_\omega: x \to (1 - \omega)x + \omega Q(x)$$

is strictly nonexpansive for $\omega \in (0, 2)$. To see this we have to show that for vectors x, y such that $Q_\omega(x) - Q_\omega(y) \neq x - y$, or equivalently

$$(2.1) \qquad y - x - Q(y) + Q(x) \neq 0 ,$$

the inequality

$$(2.2) \qquad \|Q_\omega(y) - Q_\omega(x)\|_2 < \|y - x\|_2$$

holds. An easy calculation shows that

$$(2.3) \qquad \|y - x\|_2^2 - \|Q_\omega(y) - Q_\omega(x)\|_2^2 = \omega(2 - \omega)\alpha + \omega^2\beta ,$$

where $\alpha = (y - x - Q(y) + Q(x), y - x)$ and $\beta = (y - x - Q(y) + Q(x), Q(y) - Q(x))$. For $t \in [0, 1]$, $tQ(x) + (1 - t)Q(y) \in B$ and hence

$$g(t) = \|y - (1 - t)Q(y) - tQ(x)\|_2^2 - \|y - Q(y)\|_2^2 \geqq 0 .$$

From $g(0) = 0$ it follows that

$$0 \leqq g'(0) = 2(y - Q(y), Q(y) - Q(x)) .$$

Interchanging x, y and adding gives $\beta \geqq 0$. By (2.1) we get that $\alpha - \beta > 0$, and hence $\alpha > 0$. The inequality (2.2) follows now from (2.3). In the special case when B is a closed halfspace

$$B = \{z \in \mathbb{R}^k: a^\mathsf{T}z \leqq b\}, \quad a \neq 0, b \in \mathbb{R} ,$$

$Q_\omega(x)$ is given by

$$(2.4) \qquad Q_\omega(x) = x - \omega \frac{(a^\mathsf{T}x - b)_+}{a^\mathsf{T}a} a ,$$

where we use the abbreviation $t_+ = \max(0, t)$. We remark that we have not used the finiteness of the dimension of the underlying space. Therefore the above result holds, in fact, in any Hilbert space (see also [7]).

In the next example we consider the case when the convex set is given implicitly as a level set of a convex function.

Example 3. Let $S: \mathbb{R}^k \to \mathbb{R}$ be a convex continuously differentiable function and let $B = \{x \in \mathbb{R}^k: S(x) \leqq f\} \neq \varnothing$ where f is a number. We claim that the operator T defined by

(2.5)
$$T: x \to \begin{cases} x - \omega \dfrac{S(x) - f}{\|S'(x)\|_2^2} S'(x) & \text{if } S(x) > f \\[4mm] x & \text{if } S(x) \leqq f \end{cases}$$

is paracontracting for any $\omega = (0, 2)$. For notational convenience we consider the gradient $S'(x)$ as a vector in $\mathbb{R}^k$. Since S attains its minimum in B it follows that $S'(x) \neq 0$ for $x \notin B$ and hence B is the set of all fixed points of T. To see that T is paracontracting it suffices to show that $\|T(x) - y\|_2 < \|x - y\|_2$ for any $x \notin B$ and $y \in B$. The inequality

$$(x - \omega \alpha S'(x) - y, x - \omega \alpha S'(x) - y) < (x - y, x - y) ,$$

where

$$\alpha = \frac{S(x) - f}{\|S'(x)\|_2^2} > 0 ,$$

is equivalent to the inequality

$$\omega^2 \alpha^2 \|S'(x)\|_2^2 - 2\omega\alpha(S'(x), x - y) < 0$$

and, after substituting the expression for α to the inequality

$$\omega(S(x) - f) < 2(S'(x), x - y) .$$

Since S is convex and continuously differentiable, it follows that for any $x \notin B$ and any $y \in B$

$$S(x) - S(y) \leqq (S'(x), x - y) .$$

Hence for any $\omega \in (0, 2)$,

$$\omega(S(x) - f) \leqq \omega(S(x) - S(y)) < 2(S(x) - S(y)) \leqq 2(S'(x), x - y) .$$

We remark that if $S(x) > f$ then also $S(T(x)) \geqq f$ for any $\omega \leqq 1$. Indeed, $S(T(x)) = S(x - \omega\alpha S'(x)) \geqq S(x) - (S'(x), \omega\alpha S'(x)) = S(x) - \omega(S(x) - f) = \omega f + (1 - \omega)S(x) \geqq f$. Therefore it follows, for example, from Theorem 1 in the next section that for $\omega \leqq 1$, the iteration $x_i = T(x_{i-1})$ will converge to some point on the boundary of B, $\partial B = \{y \in \mathbb{R}^k: S(y) = f\}$, from any initial x_0. The above argument implies that as long as the iterate x_{i-1} is not on ∂B then the next iterate x_i will be strictly closer to each element of B than x_{i-1}. We remark that the operator (2.5) is, in fact, a relaxed version of the Newton-Gauss iteration operator of S as defined in [6, Sect. 8.5].

As a particular case let, for example, $S(x) = \frac{1}{2}x^{\mathrm{T}}Ax + g^{\mathrm{T}}x$ where A is a positive definite matrix. Then

$$T: x \to \begin{cases} x - \omega \dfrac{\frac{1}{2}x^{\mathrm{T}}Ax + g^{\mathrm{T}}x - f}{(Ax + g)^{\mathrm{T}}(Ax + g)}(Ax + g) & \text{if } \frac{1}{2}x^{\mathrm{T}}Ax + g^{\mathrm{T}}x > f \\ \\ \qquad\qquad x & \text{if } \frac{1}{2}x^{\mathrm{T}}Ax + g^{\mathrm{T}}x \leqq f \end{cases}.$$

Example 4. Let $T: \mathbb{R} \to \mathbb{R}$ be defined by $T(x) = \sin(x^2)$. Since $T(0) = 0$ is the only fixed point of T and $|T(x)| < |x|$, for $x \neq 0$, it follows that T is paracontracting. But as there are points where $|T'(x)| > 1$, T is not strictly nonexpansive.

3 Main results

In this section we follow quite closely the ideas of [3, Sect. 2]. Given n paracontracting operators $T_1, \ldots, T_n$, we want to find a common fixed point (c.f.p.) $u \in \mathbb{R}^k$, i.e.

$$(3.1) \qquad\qquad u = T_j(u), \quad j = 1, \ldots, n .$$

We can think of two possible iterative schemes for finding such a point. The first scheme is

$$(3.2) \qquad\qquad x_i = T_{j_i}(x_{i-1}), \quad i = 1, 2, \ldots ,$$

where x_0 is a given starting vector and the sequence $\{j_i\}_{i=1}^{\infty}$ with $1 \leq j_i \leq n$ is admissible. This means that for any $1 \leq r \leq n$ there are infinitely many integers i such that $j_i = r$. The second scheme is derived from parallel asynchronous computation models, see [3] or [1] for special cases. It is determined by three sequences of integers

$$\{h_i, j_i, r_i\}_{i=T+1}^{\infty} ,$$

where

$$(3.3) \qquad \begin{array}{l} \text{(a)} \ 1 \leq h_i \leq m, \ 1 \leq j_i \leq n, \ 1 \leq r_i \leq T + 1 \\[4pt] \text{(b)} \ \{1, \ldots, n\} = \{j_i, j_{i+1}, \ldots, j_{i+T-1}\} \quad \text{for all } i > T \\[4pt] \text{(c)} \ h_i = 1 \Rightarrow r_i = 1 . \end{array}$$

Here $m \geq 2$ and $T \geq 1$ are given integers. Condition (b) shows that, in the language of [1], $\{j_i\}$ is a regulated sequence on $\{1, \ldots, n\}$ with a computation cycle T. The iteration scheme is now given by

$$(3.4) \qquad\qquad x_i = \begin{cases} x_0 & 0 \leq i \leq T \\ \alpha_{h_i}x_{i-1} + (1 - \alpha_{h_i})T_{j_i}(x_{i-r_i}) & i > T \end{cases} ,$$

where $\alpha_1 = 0$ and $\alpha_i \in (0, 1)$, $i = 2, \ldots, m$, are given numbers.

Theorem 1. *Let T_j, $j = 1, \ldots, n$ be n paracontracting operators with respect to some norm $\|\cdot\|$ in $\mathbb{R}^k$. Let $\{j_i\}_{i=1}^{\infty}$ be an admissible sequence and $x_0 \in \mathbb{R}^k$ be given. Then the sequence $\{x_i\}$ defined by (3.2), namely,*

$$x_i = T_{j_i}(x_{i-1}), \quad i = 1, 2, \ldots ,$$

converges if and only if the operators $T_1, T_2, \ldots, T_n$ *have a common fixed point. Moreover, in this case the limit*

$$(3.5) \qquad \lim_{i \to \infty} x_i = x$$

is one of such common fixed points, namely,

$$(3.6) \qquad T_j(x) = x, \quad j = 1, \ldots, n \, .$$

Proof. Let (3.5) hold and consider some fixed T_k. Since the sequence $\{j_i\}$ is admissible it follows that there exists a subsequence $\{x_{i_p}\}$ such that $x_{i_p} = T_k(x_{i_p - 1})$. As T_k is continuous and the subsequences $\{x_{i_p}\}$ and $\{x_{i_p - 1}\}$ converge to x it follows that $T_k(x) = x$.

Let now y be a common fixed point. As the T_j's are paracontracting, the sequence $\{\|x_i - y\|\}$ satisfies

$$(3.7) \qquad \|x_i - y\| \leqq \|x_{i-1} - y\|, \quad i \geqq 1 \, ,$$

and is hence convergent. Put

$$(3.8) \qquad \lambda = \lim_{i \to \infty} \|x_i - y\| \, .$$

Since $\{x_i\}$ is bounded, it has an accumulation point, say, x. We claim that x is a c.f.p. and that $\lim_{i \to \infty} x_i = x$. There exists a subsequence $\{p_i\}_{i=1}^{\infty}$ such that

$$\lim_{i \to \infty} x_{p_i} = x \, .$$

If x is not a c.f.p., then there is (after eventually renumbering the T_j) an index r, $1 \leqq r \leqq n$ such that

$$(3.9) \qquad T_j(x) = x \quad \text{for } j < r, \ T_j(x) \neq x \quad \text{for } j \geqq r \, .$$

For any i there is a minimal $q_i \geqq p_i$ such that $j_{q_i + 1} \geqq r$ because $\{j_i\}_{i=1}^{\infty}$ is admissible. As $j_v < r$ for $p_i \leqq v \leqq q_i$, it follows that $\|x_{q_i} - x\| = \|T_{j_{q_i}}(x_{q_i - 1}) - T_{j_{q_i}}(x)\| \leqq \ldots \leqq \|x_{p_i} - x\|$ and hence also

$$\lim_{i \to \infty} x_{q_i} = x \, .$$

Now there exists $s \geqq r$ such that $j_{q_i + 1} = s$ for infinitely many i. Let us rename the subsequence of $\{x_{q_i}\}$ defined by those indices by $\{y_i\}$. We observe that $T_s(y_i)$ is an element of the sequence $\{x_v\}_{v=1}^{\infty}$ for all i and hence

$$\lambda = \lim_{i \to \infty} \|T_s(y_i) - y\| = \|T_s(x) - y\| \, .$$

Also

$$\lambda = \lim_{i \to \infty} \|x_{q_i} - y\| = \|x - y\| \, .$$

Since T_s is paracontracting, it follows that $T_s(x) = x$. This contradicts (3.9) so x is a c.f.p. Substituting y by x in (3.7) we see that $\lim x_i = x$. $\square$

Corollary 1. *Let* T_j, $j = 1, \ldots, n$ *be* n *paracontracting operators with a common fixed point. Let* $T = T_1 T_2 \ldots T_n$. *Then* x *is a fixed point of* T *if and only if* x *is a common fixed point of* $T_1, T_2, \ldots, T_n$.

Proof. If $T_i(x) = x$ for $i = 1, \ldots, n$ then clearly $T(x) = x$. Assume now that $T(x) = x$. Consider the iteration $x_0 = x$ and

$$x_i = T_{i(\bmod n)}(x_{i-1}), \quad i = 1, 2, \ldots .$$

Since the sequence of indices $\{i(\bmod n)\}_{i=1}^{\infty}$ is admissible it follows that $\lim x_i$ exists. Since $x_{kn} = T^k(x_0) = x_0 = x$ it follows that the sequence $\{x_i\}$ has a subsequence whose entries are all equal to x. Thus $\lim x_i = x$ which by Theorem 1 must be a common fixed point of $T_1, \ldots, T_n$. $\square$

Remark. The assumption in the above theorem that the sequence $\{j_i\}_{i=1}^{\infty}$ is admissible is not essential. Without this assumption the limit x will be a common fixed point of those operators which appear infinitely often in the sequence. Indeed, if we start with an index i_0 large enough such that each operator which appears for indices larger than i_0 appears there infinitely many times, then Theorem 1 applies to the subset of $T_1, \ldots, T_n$ consisting of such operators.

Remark. It is easy to see that there exist discontinuous operators T which possess a fixed point and which satisfy the condition of paracontractness, namely, that for any fixed point $y = T(y)$ and any point x either $\|T(x) - y\| < \|x - y\|$ or $T(x) = x$. Such an example is given by $T: \mathbb{R} \to \mathbb{R}$, with

$$T: x \to \begin{cases} 1 & \text{if } x > 1 \\ 0 & \text{if } x \leq 1 \end{cases}.$$

It would be interesting to know whether for a finite pool of possibly discontinuous paracontractions with a common fixed point, the iteration (3.2) still converges.

We turn now to the second iteration scheme (3.4). Here we can prove:

Theorem 2. *Let $\|\cdot\|$ be a strictly convex vector norm on $\mathbb{R}^k$, $x_0 \in \mathbb{R}^k$, and T_j, $j = 1, \ldots, n$, paracontracting operators. The sequence $\{x_i\}_{i=0}^{\infty}$ defined by*

$$x_i = \begin{cases} x_0 & \text{if } 0 \leq i \leq T \\ \alpha_{h_i} x_{i-1} + (1 - \alpha_{h_i}) T_{j_i}(x_{i-r_i}) & \text{if } i > T \end{cases}$$

where $\{h_i, j_i, r_i\}_{i=T+1}^{\infty}$ satisfies (3.3), converges if and only if the operators $T_1, T_2, \ldots, T_n$ have a common fixed point. Moreover, in this case the limit

(3.10) $$x = \lim x_i$$

is one of the common fixed points, namely,

(3.11) $$x = T_i(x), \quad i = 1, \ldots, n .$$

Proof. As in the proof of Theorem 1 the existence of $x = \lim x_i$ implies that x is a common fixed point of $T_1, \ldots, T_n$. Suppose now that y is a common fixed point of $T_1, \ldots, T_n$. For $i \geq T$ let

$$\xi_i = \begin{pmatrix} x_i \\ x_{i-1} \\ \vdots \\ x_{i-T} \end{pmatrix} \in \mathbb{R}^{k(T+1)} .$$

Then

(3.12) $\xi_i = B_i(\xi_{i-1})$

where $B_i = B(h_i, j_i, r_i)$ is the operator given by

$$\eta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_{T+1} \end{pmatrix} \rightarrow B_i(\eta) = \begin{pmatrix} \alpha_{h_i}\eta_1 + (1 - \alpha_{h_i})T_{j_i}(\eta_{r_i}) \\ \eta_1 \\ \vdots \\ \eta_T \end{pmatrix} .$$

Set

$$\tilde{y} = \begin{pmatrix} y \\ \vdots \\ y \end{pmatrix} \in \mathbb{R}^{k(T+1)} ,$$

and

$C_i = B_{i+T-1} \circ \ldots \circ B_{i+1} \circ B_i$ and $\tilde{C}_i = B_{i+2T-1} \circ \ldots \circ B_i = C_{i+T} \circ C_i .$

We define a norm on $\mathbb{R}^{k(T+1)}$ by $\|\|\eta\|\| = \max_i \|\eta_i\|$ and show that for a given $\eta \in \mathbb{R}^{k(T+1)}$,

(3.13) $\|\|\tilde{C}_i\eta - \tilde{y}\|\| < \|\|\eta - \tilde{y}\|\|$ or $C_i\eta = \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix}$, x a c.f.p .

First, either

$$\|(B_i(\eta) - \tilde{y})_1\| < \|\|\eta - \tilde{y}\|\|$$

or

$$\|\alpha_{h_i}(\eta_1 - y) + (1 - \alpha_{h_i})[T_{j_i}(\eta_{r_i}) - y]\| = \max_v \|\eta_v - y\| .$$

In the latter case, if $\alpha_{h_i} > 0$, then by the strict convexity of $\| \cdot \|$ and by the fact that T_{j_i} is paracontracting, it follows that

$$\eta_1 - y = T_{j_i}(\eta_{r_i}) - y = \eta_{r_i} - y$$

and hence

$$\eta_1 = \eta_{r_i} = T_{j_i}(\eta_{r_i}), \qquad (B_1(\eta))_1 = \eta_1 .$$

The last equalities are also true if $\alpha_{h_i} = 0$ because then $r_i = 1$. In any case the first two entries of $B_1(\eta)$ are equal to η_1 and $T_{j_i}(\eta_1) = \eta_1$. Proceeding in this manner we have that either there exists $v < T$ such that

$$\|(B_{i+v} \circ \ldots \circ B_i(\eta) - \tilde{y})_1\| < \|\|\eta - \tilde{y}\|\|$$

or

$$C_i\eta = \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix}, \qquad \eta_1 = x, \qquad T_j(x) = x, \quad j = 1, \ldots, n .$$

In the former case after at most T more steps all subvectors have a norm less than $\|\|\eta - \tilde{y}\|\|$ and hence $\|\|\tilde{C}_i(\eta) - \tilde{y}\|\| < \|\|\eta - \tilde{y}\|\|$. Thus (3.13) is proved.

Returning to the sequence (3.12) we see that

(3.14) $$|||\xi_v - \tilde{y}||| \leqq |||\xi_{v-1} - \tilde{y}|||$$

and hence that limit

(3.15) $$\lambda = \lim_{v \to \infty} |||\xi_v - \tilde{y}|||$$

exists. Now consider the sequence $\{\zeta_i\}_{i=1}^{\infty}$ with

$$\zeta_i = \xi_{iT}, \quad i \geqq 1 .$$

It has an accumulation point, say ζ, and by choosing a suitable subsequence $\{s_i\}_{i=1}^{\infty}$, we have that

(3.16) $$\lim_{i \to \infty} \zeta_{s_i} = \zeta, \qquad S_1 \zeta_{s_i} = \zeta_{1+s_i}, \qquad S_2 \zeta_{1+s_i} = \zeta_{2+s_i}, \quad i = 1, 2, \dots ,$$

for some operators S_1 and S_2 which are products of T consecutive B_j's. This is possible as there are only finite many different products of B_j's of length T. But then from (3.16)

$$\lim \zeta_{1+s_i} = S_1 \zeta, \qquad \lim \zeta_{2+s_i} = S_2 S_1 \zeta$$

and by (3.15)

$$\lambda = |||\zeta - \tilde{y}||| = |||S_2 S_1(\zeta) - \tilde{y}||| .$$

By (3.13) we have, as $S_2 S_1 = \tilde{C}_i$ and $S_1 = C_i$ for a suitable i, that

$$S_1(\zeta) = \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix}, \qquad T_j(x) = x, \quad j = 1, \dots, n .$$

In particular

$$\lim |||\zeta_{1+s_i} - S_1(\zeta)||| = 0$$

and, by (3.14), on choosing now $\tilde{y} = S_1(\zeta)$

$$\lim_{r \to \infty} |||\xi_r - S_1(\zeta)||| = 0 .$$

Thus (3.10) has been proved. $\square$

Remark. If $h_i = 1, i = 1, 2 \dots$ then the iteration scheme (3.4) becomes the scheme (3.2). When this is the case, Theorem 1 is stronger than Theorem 2 as it only requires the sequence $\{j_i\}_{i=1}^{\infty}$ to be admissible, but not necessarily regulated.

Remark. The proof of Theorem 2 is similar to the proof of Theorem 2 of [3] with one considerable difference. Whereas in [3] the embedding of the multi-step method (3.4) as a one-step method in higher dimensions is done in the nk-dimensional space, here the embedding is done in the $(T + 1)k$-dimensional space. The analysis we have done is yet another instance that convergence analysis of the parallel asynchronized iteration models can be achieved by transforming them to asynchronized sequential models in higher dimensional spaces.

4 Applications to the solution of linear systems of equations with constraints

We consider a consistent linear system of equations

(4.1) $$Ax = f, \qquad A = \begin{pmatrix} r_1^{\mathrm{T}} \\ \vdots \\ r_n^{\mathrm{T}} \end{pmatrix}, \qquad A \in \mathbb{C}^{n \times k}$$

with infinitely many solutions. Suppose that on the basis of some prior information it is desired to find a solution which belongs to the intersection of given convex sets $B_1, \ldots, B_m$. Theorems 1 and 2 can be used for this purpose as follows. With each row of the matrix A we associate the paracontraction of the Example 1, namely

(4.2) $$T_i = \left(I - \omega_i \frac{r_i r_i^{\mathrm{T}}}{r_i^{\mathrm{T}} r_i} \right) + g_i$$

where

$$g_i = \omega_i \frac{f_i}{r_i^{\mathrm{T}} r_i} r_i, \quad \omega_i \in (0, 2), \qquad \text{and} \qquad f = (f_1, f_2, \ldots, f_n)^{\mathrm{T}} \ .$$

It is well known that the set of solutions of $Ax = f$ coincides with the set of all common fixed points of $T_1, \ldots, T_n$. In order to define paracontracting operators whose sets of fixed points coincide with the convex sets $B_1, \ldots, B_m$, we consider two separate cases.

Case 1: Suppose that the projection operators $Q_1, \ldots, Q_m$ onto the sets $B_1, \ldots, B_m$ are explicitly computable. This is the case, for example, when the B_j's are of the type of a "box",

$$B = \{ x \in \mathbb{R}^k \colon \alpha_i \leq x_i \leq \beta_i, \, i = 1, \ldots, k \}$$

or a ball,

$$B = \{ x \in \mathbb{R}^k \colon \| x - a \| \leq r \} \ .$$

It is obvious that for these types of convex sets the values of the corresponding projection operators Q can be computed explicitly. The corresponding relaxed projection operators as defined in Example 2 are strictly nonexpansive and hence paracontracting. In this case for $i = 1, \ldots, m$, we let

$$T_{n+i} \colon x \to (1 - \omega_{n+i}) x + \omega_{n+i} Q_i(x) \ .$$

Case 2: Suppose that the convex sets $B_1, \ldots, B_m$ are defined by convex continuously differentiable functions $S_1(x), \ldots, S_m(x)$ as in Example 3. Then we associate with each convex set B_i the paracontracting operator defined by (2.5). In this case for $i = 1, \ldots, m$, we let

$$T_{n+i} \colon x \to \begin{cases} x - \omega_{n+i} \dfrac{S_i(x) - f}{\| S_i'(x) \|_2^2} S_i'(x) & \text{if } S_i(x) > f \\[2ex] x & \text{if } S_i(x) \leq f \end{cases} \ .$$

We remark that the combination of both cases when some of the projection operators are given explicitly and some are as in (2.5) is considered in the following algorithm:

Algorithm. Choose an admissible sequence of indices $\{j_i\}_{i=1}^{\infty}$ with $1 \leqq j_i \leqq n + m$ and an arbitrary x_0, and compute recursively

$$x_i = T_{j_i}(x_{i-1}), \quad i = 1, 2, \ldots .$$

   1. If a given convergence criterion is satisfied then accept the last iterate as a common fixed point of the operators $T_1, \ldots, T_{n+m}$, that is as a solution of $Ax = f$ which belongs to the intersection of $B_1, \ldots, B_m$.
   2. If a given divergence criterion is satisfied then conclude that there is no solution of $Ax = f$ which belongs to this intersection.
   For efficient implementations on multiple instruction multiple data (MIMD) parallel computer architectures one can use a similar algorithm with a regulated sequence $\{j_i\}$ and the recursion (3.2) replaced by the recursion (3.4).

Remark. There are many different possibilities for the choice of iteration schemes (3.2) or (3.4). For example, the one considered in [4] is of the form

$$x_i = Q(T(x_{i-1}))$$

where $T = T_n T_{n-1} \ldots T_1$ and $Q = T_{n+1}$ is a projection on a "box", as described earlier in this section. This iteration, called the ART (Algebraic Reconstruction Technique) algorithm with constraints, is useful for tomographic reconstruction from incomplete data. Here are two asynchronous analogues of the last iteration scheme, namely,

(4.3) $$x_i = Q(\alpha_{j_i} x_{i-1} + (1 - \alpha_{j_i}) T_{j_i}(x_{i-r_i}))$$

and

(4.4) $$x_i = \alpha_{j_i} x_{i-1} + (1 - \alpha_{j_i}) Q(T_{j_i}(x_{i-r_i}))$$

where $\{j_i\}_{i=1}^{\infty}$ is a regulated sequence. An asynchronous iteration of this type for solving $Ax = f$ without constraints was studied in [3].
   Since the number of processors of a parallel computer may be much smaller than the number of operators $T_1, \ldots, T_n$, then for the efficient implementation of the processes of the type (4.3) and (4.4) one can group these operators into several disjoint groups whose number coincides with the number of processors. With each group one associates a strictly nonexpansive operator which is the product of all operators in this group taken in some order (recall that the operators $T_1, \ldots, T_n$, given by (4.2), are strictly nonexpansive). The new pool of strictly nonexpansive operators has now the same cardinality as the set of processors and the iteration (4.3) or (4.4) can be applied to this new pool. Since it is assumed that $T_1, \ldots, T_n$ have a common fixed point it follows from Corollary 1 that the limit will be a common fixed point of the original pool of operators.

Remark. Various other iteration schemes for a finite pool of paracontracting operators can be written. Optimal choices of the sequence of indices $\{j_i\}_{i=1}^{\infty}$ and the relaxation parameters ω in the operators of Examples 1, 2, and 3 which would give as fast convergence as possible remain an open problem. Some results in this direction for linear paracontractions can be found in [5]. The effects of round-off errors on the convergence of the proposed iteration schemes also remain to be investigated.

Remark. It is also possible to apply Theorem 1 to the approximate solution of inconsistent systems with constraints. For example, the operator T which is a product of the operators $T_1, \ldots, T_n$ given by (4.2) and taken in some order, will always have a fixed point even if there is no common fixed point of $T_1, \ldots, T_n$. Therefore the iteration $x_i = T(x_{i-1})$ will always converge to some approximate solution of $Ax = f$. If in addition, there is a common fixed point of Q and T, then the iteration $x_i = Q(T(x_{i-1}))$ will converge to an approximate solution of $Ax = b$ which is in the convex set $B = \mathrm{Im}Q$ (see [4] for details).

Remark. The proposed sequential and asynchronous iteration schemes can also be used, at least in principle, for the solution of convex optimization problems, where the pool of paracontractions consists of relaxed projection operators on convex sets and operators of the form (2.5) only. In this case a point in the intersection (on the boundary of the intersection if $\omega \leq 1$) of the corresponding convex sets will be found. Since a linear equation can be replaced by a pair of projections onto corresponding complementary halfspaces, the solution of a linear system of equations with constraints can be viewed similarly.

As in the linear case, the iteration of Theorem 1 can be used to find some approximate solution of the inconsistent convex optimization problem. We have in mind the following situation. Suppose that $T_i = (1 - \omega_i)I + \omega_i P_i$, $i = 1, \ldots, n$, are relaxed projection operators on *bounded* convex sets $B_1, \ldots, B_n$ which may have an empty intersection. Let $T = T_{i_1} T_{i_2} \ldots T_{i_n}$ be a product of the operators $T_1, \ldots, T_n$ taken in some order. It is easy to see that T maps the closed convex hull of $B_1, \ldots, B_n$ into itself and hence by the Brouwer fixed point theorem it must have a fixed point (see, for example, [6, Sect. 6.3]). Since T is strictly nonexpansive it follows from Theorem 1 that the iteration

$$x_i = T(x_{i-1})$$

will converge to a limit point which belongs to the set B_{i_1}. A certain characterization of this limit point for $\omega_i = 1$ can be found in Youla and Velasco [8], where such point is called a minimizer.

A similar argument applies to the case when T_{i_1} is a projection on a bounded convex set and $T_{j_2}, \ldots, T_{j_n}$ are arbitrary strictly nonexpansive operators.

5 Infinite pool of iteration operators

In this section we apply the techniques which were developed in earlier sections to the important, but little investigated, case of an infinite pool of iteration operators (see [7] for discussion). We give here our first result on the convergence of the sequential iteration

$$x_i = S_i(x_{i-1}), \quad i = 1, 2, \ldots,$$

where $\{S_i\}_{i=1}^{\infty}$ is a union of a finite number of convergent subsequences whose limits are paracontracting operators. This result can be viewed as a generalization of a recent result of Youla ([7, Theorem 2]) for an infinite pool of relaxed projection operators.

We say that a nonlinear operator T acting in $\mathbb{R}^k$ is strongly bounded if there exists $\beta > 0$ such that $\| T(x) \| \leqq \beta \max(1, \| x \|)$ for any $x \in \mathbb{R}^k$. The infimum of all such β's is denoted by $\| T \|$ where, clearly,

$$\| T \| = \sup_{x \in \mathbb{R}^k} \frac{\| T(x) \|}{\max(1, \| x \|)} .$$

Definition 3. A sequence $\{ S_i \}_{i=1}^{\infty}$ of continuous operators mapping $\mathbb{R}^k$ into $\mathbb{R}^k$ is called *asymptotically paracontracting*, (*with respect to a vector norm $\| \cdot \|$ on $\mathbb{R}^k$*), if there exist paracontracting operators $T_1, T_2, \ldots, T_n$ and a mapping $J: \mathbb{N} \to \{ 1, 2, \ldots, n \}$ such that $S_i - T_{J(i)}$ is strongly bounded for any $i \geqq M$ where M is a positive integer, and

(5.1)
$$\sum_{i=M}^{\infty} \| S_i - T_{J(i)} \| < \infty .$$

In order to prove our convergence theorem we need the following lemma:

Lemma 1. *Let $\alpha_i \geqq 0$, $\sum_{i=1}^{\infty} \alpha_i < \infty$ and a sequence $\{ e_i \}_{i=0}^{\infty}$ be given. Suppose that*

$$e_i \leqq e_{i-1} + \alpha_i, \quad i = 1, 2, \ldots .$$

Then $\{ e_i \}_{i=1}^{\infty}$ is convergent (and hence bounded).

Proof. Obviously, for $i > j$

$$e_i \leqq e_j + \sum_{k=j+1}^{i} \alpha_k .$$

Hence for $\varepsilon > 0$ there exists $N(\varepsilon)$ such that for $i > j \geqq N(\varepsilon)$

$$e_i \leqq e_j + \varepsilon .$$

This shows that the sequence $\{ e_i \}_{i=1}^{\infty}$ is bounded and each of its accumulation points is less or equal to $\liminf e_i$. Hence there is only one accumulation point and $\{ e_i \}_{i=0}^{\infty}$ is convergent. $\square$

Theorem 3. *Let $\{ S_i \}_{i=1}^{\infty}$ be an asymptotically paracontracting sequence. Then the sequence of iterates*

$$x_i = S_i(x_{i-1}), \quad i = 1, 2, \ldots ,$$

converges if and only if there exists a common fixed point of those operators among $T_1, T_2, \ldots, T_n$ whose indices appear infinitely often in the sequence $\{ J(i) \}_{i=1}^{\infty}$. Moreover, in this case the limit

(5.2)
$$\lim_{i \to \infty} x_i = x$$

is a common fixed point of such T_k's.

Proof. Suppose that $x = \lim_{i \to \infty} x_i$ and let T_k be such that $k = J(i_p)$, $p = 1, 2, \ldots$. Then

$$\sum_{i_p \geqq M} \| S_{i_p} - T_k \| < \infty .$$

Let

$$x_{i_p} = S_{i_p}(x_{i_p-1}) .$$

Both $\{x_{i_p}\}_{p=1}^{\infty}$ and $\{x_{i_p-1}\}_{p=2}^{\infty}$ converge to x and hence are bounded. Therefore for p large enough,

$$\|x_{i_p} - T_k(x)\| \leq \|S_{i_p}(x_{i_p-1}) - T_k(x_{i_p-1})\| + \|T_k(x_{i_p-1}) - T_k(x)\|$$

$$\leq \|S_{i_p} - T_k\| \max(1, \|x_{i_p-1}\|) + \|T_k(x_{i_p-1}) - T_k(x)\| \to 0 ,$$

when $p \to \infty$. Hence $T_k(x) = x$.

Now let y be a common fixed point of those T_k's whose indices appear infinitely often in the sequence $\{J(i)\}_{i=1}^{\infty}$. Let $e_i = \|x_i - y\|$ and

(5.3) $$\alpha_i = \|S_i - T_{J(i)}\| .$$

Then for i large enough,

$$\|x_i - y\| \leq \|S_i(x_{i-1}) - T_{J(i)}(x_{i-1})\| + \|T_{J(i)}(x_{i-1}) - T_{J(i)}(y)\|$$

$$\leq \|S_i - T_{J(i)}\| \max(1, \|x_{i-1}\|) + \|x_{i-1} - y\|$$

$$= \alpha_i \max(1, \|x_{i-1}\|) + e_{i-1} \leq \alpha_i \delta + (1 + \alpha_i)e_{i-1} ,$$

where $\delta = \max(\|y\|, 1)$. Therefore the sequence $\{e_i'\}$, with $e_i' = e_i \prod_{j\leq i}(1 + \alpha_j)$, $i = 0, 1, \ldots$, satisfies the inequality

$$e_i' \leq e_{i-1}' + \delta\alpha_i \prod_{j<i}(1 + \alpha_j) = e_{i-1}' + \alpha_i' .$$

Since $\sum \alpha_i' < \infty$ it follows from the preceeding lemma that $\{e_i'\}$ and hence $\{e_i\}$ is convergent. This in turn implies that the sequence $\{x_i\}$ is bounded. Let us choose a number γ such that for all i, $\|x_i\| \leq \gamma$ and $1 \leq \gamma$. Let x be an accumulation point of $\{x_i\}$, so that

$$\lim_{i \to \infty} x_{p_i} = x$$

for some infinite sequence of positive integers $\{p_i\}_{i=1}^{\infty}$. We can assume (after a possible reordering) that $T_1, T_2, \ldots, T_u$ where $u \leq n$, are those operators whose index appears infinitely often in the sequence $\{J(i)\}$. We claim that $T_i(x) = x$ for $i = 1, \ldots, u$. Assume there is $r \leq u$ such that $T_i(x) = x$ for $i < r$ while $T_i(x) \neq x$ for $i \geq r$ ($r = 1$ is a possibility). Then there is a minimal $q_i \geq p_i$ such that $J(q_i + 1) \geq r$, $J(q_i) < r, \ldots, J(p_i + 1) < r$ ($q_i = p_i$ is a possibility). Then for i large enough,

$$\|x_{q_i} - x\| \leq \|S_{q_i}(x_{q_i-1}) - T_{J(q_i)}(x_{q_i-1})\| + \|T_{J(q_i)}(x_{q_i-1}) - T_{J(q_i)}(x)\|$$

$$\leq \alpha_{q_i}\gamma + \|x_{q_i-1} - x\| \leq \cdots \leq (\alpha_{q_i} + \cdots + \alpha_{p_i+1})\gamma + \|x_{p_i} - x\| .$$

Therefore also $\lim_{i \to \infty} x_{q_i} = x$ exists. There is some index, say $v \geq r$, which appears infinitely often in the sequence $\{J(q_i + 1)\}$ and hence, for this subsequence, denoted by $\{q_i'\}$, the sequence $\{S_{q_i'+1}(x_{q_i'})\}$ is a subsequence of $\{x_{p_i}\}$ and $T_{J(q_i'+1)} = T_v$. Thus

$$\|T_v(x) - y\| = \|\lim T_v(x_{q_i'}) - y\| = \|\lim S_{q_i'+1}(x_{q_i'}) - y\| = \|x - y\| .$$

Since T_v is paracontracting we conclude that $T_v(x) = x$. This contradiction shows that x is a fixed point of $T_1, \ldots, T_u$. Finally, substituting x for y in the definition of

$\{e_i\}$ we see that $\lim e_i = 0$, as a subsequence of $\{e_i\}$ converges to zero. Therefore (5.2) holds. $\quad \square$

Remark. If the boundedness of the sequence $\{x_i\}$ can be established a-priori by some means other than those used in the above proof (for example, if all S_i's are nonexpansive), then it is possible to show that the condition (5.1) can be replaced by weaker conditions. For example, Theorem 3 remains true in this case if the sequence $\{J(i)\}$ is regulated and if the sequence $\{\alpha_i\}$ defined in (5.3) converges to zero. The above theorem should be considered as only one of the starting points for the study of convergence properties of an infinite pool of operators.

References

1. Bru, R., Elsner, L., Neumann, M. (1988): Models of parallel chaotic iteration methods. Linear Algebra Appl. **102**, 175–192
2. De Pierro, A., Iusem, A. (1990): On the asymptotic behavior of some alternate smoothing series expansion iterative methods. Linear Algebra Appl. **130**, 3–24
3. Elsner, L., Koltracht, I., Neumann, M. (1990): On the convergence of asynchronous paracontractions with applications to tomographic reconstruction from incomplete data. Linear Algebra Appl. **130**, 65–82
4. Koltracht, I., Lancaster, P. (1990): Constraining Strategies for linear iterative processes. IMA J. Numer. Anal. **10**, 555–567
5. Nelson, S., Neumann, M. (1987): Generalization of the projection method with applications to SOR method for Hermitian positive semidefinite linear systems. Numer. Math. **51**, 123–141
6. Ortega, J.M., Rheinboldt, W.C. (1970): Iterative solution of nonlinear equations in several variables. Academic Press, New York
7. Youla, D.C. (1990): On deterministic convergence of iterations of relaxed projection operators. J. Visual Comm. Image Rep. **1,1**, 12–20
8. Youla, D.C., Velasco, V. (1986): Extensions of a result on the synthesis of signals in the presence of inconsistent constraints. IEEE Trans. Circuits Syst. CAS-**33**, 455–468