Electronic Transactions on Numerical Analysis.
Volume 13, pp. 38-55, 2002.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

PERTURBATION OF PARALLEL ASYNCHRONOUS LINEAR ITERATIONS BY FLOATING POINT ERRORS *

PIERRE SPITERI †, JEAN-CLAUDE MIELLOU ‡, AND DIDIER EL BAZ §

Abstract. This paper deals with parallel asynchronous linear iterations perturbed by errors in floating point arithmetic. An original result is presented which permits one to localize the limits of perturbed parallel asynchronous linear iterations. The result is established by using the approximate contraction concept. Simple examples are studied.

Key words. approximate contraction, parallel algorithms, asynchronous iterations.

AMS subject classifications. 65F10, 65G05, 65Y05, 68Q22, 68Q10.

1. Introduction. Major contributions on the accuracy and stability of numerical computations have been developped in numerous works (see [3] , [9], [10] and [18]). Computations carried out on machines imply approximate representation of real numbers. When approximate computations are performed, one obtains an exact solution of a perturbed problem. Generally speaking, many iterative algorithms lead to the solution of fixed point problems for which the previous observation is still valid.

From a mathematical point of view, the concept of approximate contraction is a very classical tool; the reader is refered to the books of N.J. Higham [9] (see also its references), M.A. Krasnosel'skii et al. [11] and J. Ortega and W.C. Rheinboldt [15] for the analysis of the effect of perturbations by using the approximate contraction concept. This concept was extended in an abstract context by J.C. Miellou et al. (see [13]) in order to study the behaviour of various perturbed fixed point algorithms. In the above mentionned reference, the authors are rather interested in the localization of the limits of iterate vector subsequences in the sequential and parallel asynchronous context than in convergence properties; the latter issue is nevertheless considered in references [1], [5] and [12]. Indeed, under a general approximate contraction assumption, it can be shown that the limits of subsequences of iterate vectors belong to a ball of center $u^\star$, solution of the exact problem, and finite radius. In the case of iterative methods with linear convergence, the radius of the ball is a function of the contraction constant associated with the fixed point mapping on the one hand and the perturbations on the other hand.

The case of linear iterations is very important in practice for the solution of algebraic systems. For example, when one uses Newton-like methods, each step finally consists in the solution of a linear problem.

The objective of the present study is to analyze the effect of roundoff errors on the precision of computations in the context of parallel asynchronous iterations for linear fixed point systems. Using technical results, we can show that the approximate contraction assumption is satisfied for the linear fixed point problem $u^\star = Bu^\star + c$, where B is an (m, m) square matrix and c is a vector in the m-dimensional space. In particular, it is shown that the norm of matrix B, subordinate to the vector p-norm, where p is any given integer, is less than or equal to a number λ which bounds the spectral radius of matrix B. If the matrix B is irreducible, then λ is equal to the spectral radius of matrix B. The previous result is a generalization of the

*Received April 24, 2001. Accepted for publication August 12, 2002. Recommended by A. Frommer.
†ENSEEIHT-IRIT/LIMA, 2 rue Camichel, B.P. 7122, F-31071 Toulouse-Cedex, France. E-mail: Pierre.Spiteri@enseeiht.fr
‡Département de Mathématiques, UFR Sciences et Techniques, 16 route de Gray, F-25030 Besançon-Cedex, France. E-mail: miellou@erdos.univ-fcomte.fr
§LAAS du CNRS, 7 avenue du Colonel Roche, F-31077 Toulouse-Cedex 4, France. E-mail: elbaz@laas.fr

ETNA
Kent State University
etna@mcs.kent.edu

classical convexity theorem of Riesz (see [16]). Furthermore we note that a linear iteration is carried out on a machine, as a dot product of m vectors on the one hand and a sum of two real numbers on the other hand. Thus, these arithmetic operations can be viewed as a particular situation of the context studied in the books of N.J. Higham (see [9]) and G.H. Golub and C.F. Van Loan (see [10]). This allows us to verify the approximate contraction assumption. As a consequence, the effect of roundoff errors can be studied in the case of parallel asynchronous iterations for linear fixed point problems by using the concept of approximate contraction.

Section 2 deals with the perturbation of general fixed point iterative methods in connection with the notion of approximate contraction. Section 3 presents localization results of subsequences of perturbed parallel asynchronous linear iterations. Section 4 proposes original and technical results which allow us to obtain an upper bound of the p-norm of the matrix B for $p \in [1, \infty]$, generalizing the algebraic matricial formulation of the little theorem of Riesz (see [16]); note that this result can be used only for successive approximation methods in the sequential context. In the case of parallel asynchronous iterations only the weighted maximum norm permits one to analyze the behaviour of the algorithm (see [6]). Thus, only the case where $p = \infty$ must be considered in the parallel asynchronous context. The previous characterization allows us to obtain, in section 5, the property of approximate contraction for affine mappings. The link with errors arising in floating point arithmetic is developped in section 6 in the case of rounding and chopping and in the context of parallel asynchronous linear iterations. In the last section, various examples for which the matrices B have nonnormality property are studied (see also [4]).

2. Perturbation of fixed point iterative methods in the classical context.

2.1. Approximate contraction with respect to an element. We consider a formalism similar to the nested sets pattern described by D.P. Bertsekas (see [2]) and adapted to perturbed fixed point problems in J.C. Miellou et al. (see [13]). Let E be a normed vector space. Let us denote by $v \to \|v\|_E$ the norm defined on E and let $\mathbb{N}$ be the set of natural numbers. Consider a sequence $\{E^n\}_{n\in\mathbb{N}}$ of nested closed subsets of E, such that

$$E^{n+1} \subset E^n, \forall n \in \mathbb{N}. \tag{2.1}$$

We denote by H the intersection of the subsets E^n; we note that H is also closed.

Let $\{u^n\}$ be a sequence of E; we denote by $a(\{u^n\})$ the set, which is possibly empty, of all accumulation points of $\{u^n\}$. Let $T : D(T) \subset E \to E$ be a mapping such that

$$E^0 \subset D(T) \text{ and } T(E^n) \subset E^{n+1}, \forall n \in \mathbb{N}. \tag{2.2}$$

Let also $R(T)$ be the range of T and assume the relative compactness of $E^0 \cap R(T)$.

$$\text{If } u^n \in \overline{E^0 \cap R(T)}, \forall n \text{ then we can extract a convergent subsequence in } E, \tag{2.3}$$

where $\overline{E^0 \cap R(T)}$ is the closure of the set $E^0 \cap R(T)$. We will denote by $u^\star$ the limit of the subsequence. Moreover, assume that

$$\overset{\circ}{D}(T) \neq \emptyset, \tag{2.4}$$

where $\overset{\circ}{D}(T)$ denotes the interior of $D(T)$. Assume that

$$u^\star \in \overset{\circ}{D}(T) \text{ and } \exists\, \delta > 0 \text{ such that the closed ball } B_E(u^\star; \delta) \subset \overset{\circ}{D}(T). \tag{2.5}$$

We recall the important concept of an approximately contracting mapping (see [11]).

DEFINITION 2.1. *The mapping T is approximately contracting (in brief a-contracting) with respect to $u^\star$ on $B_E(u^\star;\delta)$, if there exists a nonnegative real number θ and a real constant $l \in [0,1[$ such that*

$$\|u^\star - Tv\|_E \le l\|u^\star - v\|_E + \theta, \forall v \in B_E(u^\star;\delta), \tag{2.6}$$

where θ, l and δ satisfy

$$\theta \le (1-l)\delta. \tag{2.7}$$

REMARK 1. *The number θ is the approximation constant related to the perturbation of the fixed point mapping and l is the contraction constant.*

We now recall the following important result (see [11], [13] and [15]).

THEOREM 2.2. *Let assumptions (2.1) to (2.5) hold and consider the successive approximation method*

$$u^{n+1} = T(u^n), n = 0, 1, ..., \ u^0 \in E^0. \tag{2.8}$$

Assume that the mapping T is a-contracting with respect to $u^\star$ on $B_E(u^\star;\delta)$. Then the iteration (2.8) with $u^0 \in E^0 = B_E(u^\star;\delta)$ generates a sequence $\{u^n\}$ such that the set $a(\{u^n\})$ satisfies

$$a(\{u^n\}) \ne \emptyset\ ,\ a(\{u^n\}) \subset B_E(u^\star;\delta_\star), \tag{2.9}$$

where

$$\delta_\star = \frac{\theta}{1-l}.$$

COROLLARY 2.3. *Assume that the assumptions of Theorem 2.2 hold and consider the particular case where $H = \cap_{n\in\mathbb{N}} E^n = \{u^\star\}$. Then $a(\{u^n\}) = \{u^\star\}$ is the unique fixed point of the iteration (2.8).*

REMARK 2. *The result of Theorem 2.2 measures the maximum distance between the limit of the iterate vector subsequence and the exact solution. The measure depends on the constant $(1-l)$. In particular if $(1-l)$ is small, then $\frac{\theta}{1-l}$ may not necessarily be small.*

REMARK 3. *In the case where T is a classical contracting fixed point mapping defined on E, i.e. $\theta = 0$, the set E^n can be naturally chosen as the closed ball of center $u^\star$ and radius $l^n\|u^\star - u^0\|_E$, where $\|u^\star - u^0\|_E$ denotes the distance between $u^\star$ and u^0 in the space E. Let us note that in this case we have $\lim_{n\to\infty}(diam(E_n)) = 0$, where $diam(E_n)$ is the diameter of E_n and the assumption (2.3) is not necessary in the statement of Corollary 2.3.*

Let us consider the product space $E = \mathbb{R}^m$ and denote by $v \to \|v\|_E$, the norm defined on E. Assume that the approximate contraction assumption (2.6) holds for the mapping T. Suppose that in (2.6) $v = u^n$, where $\{u^n\}$ is the sequence produced by the successive approximation method (2.8). Then we easily obtain the following estimation

$$\|u^\star - u^n\|_E \le l^n\ \|u^\star - u^0\|_E + (l^{n-1} + ... + l + 1)\,\theta,$$

which can be written as follows

$$\|u^\star - u^n\|_E \le l^n\ \|u^\star - u^0\|_E + \left(\frac{1-l^n}{1-l}\right)\ \theta. \tag{2.10}$$

Thus, we can define a sequence $\{E^n\}_{n\in\mathbb{N}}$ of nested closed subsets

$$E^n = B_E\left(u^\star; l^n\ \|u^\star - u^0\|_E + \left(\frac{1-l^n}{1-l}\right)\ \theta\right).$$

ETNA
Kent State University
etna@mcs.kent.edu

2.2. Perturbation of a fixed point mapping. From a numerical point of view, the above mentioned mapping T generally results from a perturbation or an approximation, of a mapping $\bar{T} : D(\bar{T}) \subset E \to E$, with fixed point $u^\star$. Assume that the exact mapping $\bar{T}$ is contracting with respect to $u^\star$ in E, i.e.,

$$\|\bar{T}v - u^\star\|_E \le l\|u^\star - v\|_E \, , \forall v \in E. \tag{2.11}$$

If we make the additional assumption that the error of approximation $\|Tv - \bar{T}v\|_E$ is proportional to $\|\bar{T}v\|_E$, i.e.,

$$\|Tv - \bar{T}v\|_E \le \tau\|\bar{T}v\|_E \, , \tau > 0 \, , \forall v \in E, \tag{2.12}$$

then by the triangle inequality we have

$$\|Tv - \bar{T}v\|_E \le \tau\|\bar{T}v - u^\star\|_E + \tau\|u^\star\|_E \, , \forall v \in E,$$

and it follows from (2.11) that

$$\|Tv - \bar{T}v\|_E \le \tau l\|u^\star - v\|_E + \theta_\star \, , \forall v \in E, \tag{2.13}$$

where $\theta_\star = \tau\|u^\star\|_E$. Thus, by the triangle inequality, we obtain

$$\|Tv - u^\star\|_E \le \|Tv - \bar{T}v\|_E + \|\bar{T}v - u^\star\|_E, \forall v \in E,$$

and it follows from (2.11) and (2.13) that we have

$$\|Tv - u^\star\|_E \le (1+\tau)l\|u^\star - v\|_E + \theta_\star \, , \forall v \in E, \tag{2.14}$$

and if $(1+\tau)l < 1$, then the mapping T is a-contracting with respect to $u^\star$ in E.

REMARK 4. *It follows that the constant τ must satisfy*

$$\tau < \frac{1-l}{l}. \tag{2.15}$$

REMARK 5. *In some contexts such as the study of roundoff errors, the inequality (2.12) is not easy to use. In the sequel, we will introduce an analogous assumption well adapted to our context of study by using a vectorial norm concept (see Remark 7).*

2.3. The case where $R(T)$ is finite. Let us consider the situation where the range of T, $R(T)$ is finite. Such a situation can occur, for example, when one uses computers since the set of floating point numbers is then finite. In this case, Theorem 2.2 leads to the following result (see [13]).

THEOREM 2.4. *Let assumptions (2.1) to (2.5) hold, assume that the range of T is finite and consider the successive approximation method (2.8). Then, there exists an index n_0 such that for $n \ge n_0$, $u^n \in H$. Moreover if $H = \{u^\star\}$, then for $n \ge n_0$, $u^n = u^\star$.*

More generally we obtain the following result.

THEOREM 2.5. *Let assumptions (2.1) to (2.5) hold. If $R(T)$ is finite, then there exists an integer n_0 such that $\forall n \ge n_0$, $u^n \in H = B_E(u^\star; \frac{\theta}{1-l})$.*

ETNA
Kent State University
etna@mcs.kent.edu

3. Perturbation of parallel asynchronous iterations. For simplicity, let us consider a positive integer α and α normed vector spaces E_i, $i = 1,..,\alpha$. Let $|\ .\ |_i$ be the norm defined on E_i and let us consider also the product space E such that

$$E = \prod_{i=1}^{\alpha} E_i.$$

E is also a normed vector space. In the sequel we will consider the weighted maximum norm (see [1], [6], [8] and [12]):

$$\|u\|_{e,\infty} = \max_{1\leq i\leq\alpha}\left(\frac{|u_i|_i}{e_i}\right),$$

where e is a positive vector, with $e_i > 0$, $i = 1, 2, ..., \alpha$ and $u \in E$ is decomposed as follows

$$u = (u_1, ..., u_\alpha), \text{ with } u_i \in E_i, i = 1, 2, ..., \alpha.$$

Then T being a mapping from $D(T) \subset E$ into E, we decompose this mapping accordingly

$$T(u) = (T_1(u), ..., T_\alpha(u)), \text{ with } T_i(u) \in E_i, i = 1, 2, ..., \alpha.$$

Let us consider an initial guess $u^0 \in D(T)$ and the asynchronous iterative sequence $\{u^n\}$ defined by

$$u_i^{n+1} = \begin{cases} T_i(...., u_j^{s_j(n)}, ...), \forall i \in J(n), \\ u_i^n, \forall i \notin J(n), \end{cases} \tag{3.1}$$

where $J = \{J(n)\}_{n\in\mathbb{N}}$ is a sequence of non empty subsets of $\{1, 2, ..., \alpha\}$ denoting the subsets of indices of the components updated at the n-th iteration,

$$S = \{s_1(n), s_2(n), ..., s_\alpha(n)\}_{n\in\mathbb{N}},$$

is a sequence of elements of $\mathbb{N}^\alpha$, and J, S satisfy

$$\forall i \in \{1, 2, ..., \alpha\}, \text{ the set } \{n \in \mathbb{N} \mid i \in J(n)\} \text{ is infinite}, \tag{3.2}$$

$$\forall i \in \{1, 2, ..., \alpha\}, \forall n \in \mathbb{N},\ s_i(n) \leq n, \tag{3.3}$$

$$\forall i \in \{1, 2, ..., \alpha\},\ \lim_{n\to\infty} s_i(n) = +\infty. \tag{3.4}$$

According to a result of J.C. Miellou, P. Cortey-Dumont, and M. Boulbrachêne (see [13]), we can deduce the following result

THEOREM 3.1. *The assumptions and notations being the same as in Theorem 2.2, let E be normed by $\|.\|_{e,\infty}$. Let assumption (2.6) hold with respect to the previous norm. If assumptions (3.2) to (3.4) are satisfied, then*
1- for all $u^0 \in B_E(u^\star; \delta)$, the asynchronous iterations (3.1) are well defined,
2- with $a(\{u^n\})$ being the set of the limits of the subsequences of $\{u^n\}$, one has $a(\{u^n\}) \neq \emptyset$ and $a(\{u^n\}) \subset H = B_E(u^\star; \delta_\star)$, defined by

$$B_E(u^\star; \delta_\star) = B_E\left(u^\star; \frac{\theta}{1-l}\right) = \prod_{i=1}^{\alpha} B_{E_i}\left(u_i^\star; e_i\frac{\theta}{1-l}\right), \tag{3.5}$$

where l is the contraction constant and θ is the approximation constant.
3- If morever $R(T)$ is finite, then there exists an integer n^0, such that for any $n \geq n^0$, $u^n \in B_E(u^\star;\delta_\star)$.

The case of perturbation of iterative methods with linear convergence is one of the most important cases. This situation occurs, particularly, in the case of linear asynchronous iterations where $\bar{T}(u) = Bu + c$. With respect to linear convergence, we can deduce the following result.

COROLLARY 3.2. *The assumptions and notations being the same as in Theorem 3.1, with $\bar{T}(u) = Bu + c$ and the perturbation T of $\bar{T}$ being a-contractant with respect to an element, in the sense of Definition 2.1. Then assertions 1 to 3 of Theorem 3.1 are true.*

4. Preliminary mathematical results.

4.1. A finite dimensional weighted norm form of Riesz's convexity theorem. Let B be a nonnegative (m, m) real matrix. Assume that there exist a nonnegative constant λ and two vectors denoted by e and $e^\star$, both of which have all their components strictly positive and such that

$$Be \leq \lambda e \text{ and } B^t e^\star \leq \lambda e^\star. \tag{4.1}$$

Assume that the space $\mathbb{R}^m$ is normed by

$$\|x\|_{ee^\star,p} = \left[\sum_{i=1}^{m} e_i e_i^\star \frac{|x_i|^p}{e_i^p}\right]^{\frac{1}{p}}. \tag{4.2}$$

REMARK 6. *Consider also the weighted maximum norm defined by*

$$\|x\|_{e,\infty} = \max_{1\leq i\leq m}\left(\frac{|x_i|}{e_i}\right). \tag{4.3}$$

Let j be the index such that

$$\frac{|x_j|}{e_j} = \|x\|_{e,\infty}.$$

Then, it can be noted that

$$(e_j e_j^\star)^{\frac{1}{p}} \|x\|_{e,\infty} \leq \|x\|_{ee^\star,p} \leq \left(\sum_{i=1}^{m} e_i e_i^\star\right)^{\frac{1}{p}} \|x\|_{e,\infty}, \forall x \in \mathbb{R}^m.$$

Thus

$$\lim_{p\to\infty} \|x\|_{ee^\star,p} = \|x\|_{e,\infty}, \forall x \in \mathbb{R}^m.$$

Consequently, in the sequel, the use of norms (4.3) and (4.2) for every $p \in [1, \infty[$ leads us to consider the set $[1, \infty]$ for simplicity of presentation.

LEMMA 4.1. *Assume that the space $\mathbb{R}^m$ is normed by the weighted maximum norm (4.3) and that assumption (4.1) holds. Then the subordinate matrix norm associated with the scalar norm (4.3) satisfies:*

$$]|B|[_{e,\infty} \leq \lambda.$$

Proof. See [7], Lemma 2.1 and Proposition 1 of [12] in a general context of vectorial norms. ☐

LEMMA 4.2. *Under assumption (4.1), the subordinate matrix norm associated with the scalar norm (4.2) satisfies*

$$]|B|[_{ee^\star,p} \ \leq \ \lambda, \forall p \in [1, \infty].$$

Proof. Let us note first that the assumption (4.1) implies

$$Be \leq (\lambda + \epsilon)e \text{ and } B^t e^\star \leq (\lambda + \epsilon)e^\star, \forall \epsilon > 0. \tag{4.4}$$

Then, whatever $\epsilon > 0$, there exists $\beta(\epsilon) > 0$, such that

$$B_\beta e = (B + \beta A)e \leq (\lambda + \epsilon)e, B_\beta^t e^\star = (B^t + \beta A)e^\star \leq (\lambda + \epsilon)e^\star, \forall \beta \in]0, \beta(\epsilon)], \tag{4.5}$$

where the entries of the matrix A are all equal to one; indeed by (4.1) and (4.4), we can take $\beta(\epsilon) = \frac{\epsilon}{m}$. Let us denote by $\tilde{b}_{ij} = b_{ij} + \beta$, the entries of the matrix $B_\beta, \forall \beta \in]0, \beta(\epsilon)]$. Note that $\tilde{b}_{ij} > 0$. Let μ_i be defined by

$$\mu_i = \frac{\sum_{j=1}^m \tilde{b}_{ij} e_j}{e_i}.$$

For all i and j, the real numbers $\tilde{b}_{ij}$ being strictly positive, then μ_i is also strictly positive. Then (4.5) implies $0 < \mu_i < \lambda + \epsilon$. Let t_{ij} be defined by

$$t_{ij} = \frac{\tilde{b}_{ij} e_j}{\mu_i e_i}.$$

Then

$$\sum_{j=1}^m t_{ij} = 1 \text{ and } 0 < t_{ij}.$$

We have

$$e_i \left| \frac{\sum_{j=1}^m \tilde{b}_{ij} x_j}{\mu_i e_i} \right|^p \leq e_i \left| \sum_{j=1}^m t_{ij} \frac{|x_j|}{e_j} \right|^p, \forall x \in \mathbb{R}^n.$$

$|x_j|$ being ≥ 0, by the convexity of the mapping $y \to y^p$, $(y \geq 0)$ we finally obtain

$$e_i \left| \frac{\sum_{j=1}^m \tilde{b}_{ij} x_j}{\mu_i e_i} \right|^p \leq e_i \left| \sum_{j=1}^m t_{ij} \frac{|x_j|}{e_j} \right|^p \leq e_i \sum_{j=1}^m t_{ij} \frac{|x_j|^p}{e_j^p} = \frac{1}{\mu_i} \sum_{j=1}^m \tilde{b}_{ij} e_j \frac{|x_j|^p}{e_j^p}.$$

The previous inequality implies

$$e_i \frac{|\sum_{j=1}^m \tilde{b}_{ij} x_j|^p}{e_i^p} \leq \mu_i^{p-1} \sum_{j=1}^m \tilde{b}_{ij} e_j \frac{|x_j|^p}{e_j^p}. \tag{4.6}$$

It follows from (4.5) that

$$\sum_{i=1}^{m} \tilde{b}_{ij} e_i^\star \leq (\lambda + \epsilon) e_j^\star \ .$$

Thus, by multiplying the first term of inequality (4.6) by $e_i^\star$, and adding for all i and also by taking into account that $\mu_i^{p-1} \leq (\lambda+\epsilon)^{p-1}$ we finally obtain for all β, such that $0 < \beta \leq \beta(\epsilon) < \epsilon$,

$$\sum_{i=1}^{m} e_i^\star e_i \frac{|\sum_{j=1}^{m} \tilde{b}_{ij} x_j|^p}{e_i^p} = \sum_{i=1}^{m} e_i^\star e_i \frac{|\sum_{j=1}^{m} (b_{ij}+\beta) x_j|^p}{e_i^p} \leq (\lambda+\epsilon)^p \sum_{j=1}^{m} e_j^\star e_j \frac{|x_j|^p}{e_j^p}.$$

If now the real number $\epsilon \to 0$, then $\lim_{\epsilon \to 0} \beta(\epsilon) = 0$. Thus, by the continuity with respect to ϵ of $\beta(\epsilon)$, we can pass to the limit and finally get

$$\sum_{i=1}^{m} e_i^\star e_i \frac{|\sum_{j=1}^{m} b_{ij} x_j|^p}{e_i^p} \leq \lambda^p \sum_{j=1}^{m} e_j^\star e_j \frac{|x_j|^p}{e_j^p},$$

and the lemma is true. □

4.2. Some results related to Perron-Frobenius theory. The matrix B being reducible or irreducible, for all real number β let us associate the irreducible matrix $B_\beta = B + \beta A$. Let us also consider the respective strictly positive eigenvectors e_β and $e_\beta^\star$, respectively, of the irreducible matrices B_β and B_β^t, respectively, associated with the eigenvalue equal to the spectral radius $\rho(B_\beta)$ of B_β. The matrix B_β being an irreducible positive matrix, then according to the Perron-Frobenius Theorem

$$\forall \epsilon > 0, \exists \beta(\epsilon) > 0, \text{ such that } \forall \beta \in]0, \beta(\epsilon)], \rho(B) \leq \rho(B_\beta) \leq \rho(B) + \epsilon. \tag{4.7}$$

We have the following result

PROPOSITION 4.3. *For all $\epsilon > 0$ there exists $\beta(\epsilon)$ such that for all $\beta \in]0, \beta(\epsilon)]$, if we consider the weighted norms $\| \, . \, \|_{e_\beta e_\beta^\star, p}$, defined in (4.2) in which we substitute the vectors e and $e^\star$, respectively by e_β and $e_\beta^\star$; then for all $p \in [1, \infty]$, the nonnegative matrix B satisfies*

$$]|B|[_{e_\beta e_\beta^\star, p} \leq \rho(B) + \epsilon, \forall p \in [1, \infty].$$

Proof. As previously said, for any real number $\beta > 0$, the matrix B_β is positive and irreducible; thus, whatever the real number $\beta > 0$,

$$B_\beta e_\beta = B e_\beta + \beta A e_\beta = \rho(B_\beta) e_\beta, \text{ with for example } ||e_\beta||_2 = 1, \tag{4.8}$$

where $||e_\beta||_2$ denotes the Euclidean norm. According to the Perron-Frobenius Theorem

$$e_\beta > 0 \text{ and moreover } 0 < \beta_1 < \beta_2 \text{ involves } 0 \leq \rho(B) < \rho(B_{\beta_1}) < \rho(B_{\beta_2}). \tag{4.9}$$

By (4.8) and the compactness of the unit sphere of $\mathbb{R}^m$, let us extract a convergent sequence $e_{\beta_i} \to e_0$ when $\beta_i \searrow 0$, where $\beta_i \searrow 0$ denotes a decreasing sequence of scalars which converges to zero; by the closedness of the cone of vectors of $\mathbb{R}^m$ with nonnegative components and also by the closedness of the unit sphere of $\mathbb{R}^m$, $e_0 > 0$, $||e_0||_2 = 1$ and $B_{\beta_i} e_{\beta_i} = B e_{\beta_i} + \beta_i A e_{\beta_i} \to B e_0$. Moreover (4.9) implies that

$$\rho(B_{\beta_i}) \searrow \rho^\star \geq \rho(B) \text{ if } i \to \infty, \tag{4.10}$$

and also $\rho(B_\beta) \searrow \rho^\star \geq \rho(B)$ when $\beta \searrow 0$. Thus, the corresponding right hand side of (4.8) satisfies the convergence property $\rho(B_{\beta_i})e_{\beta_i} \to \rho^\star e_0$, when $\beta_i \searrow 0$. Therefore, passing to the limit in (4.8) is possible in order to obtain $Be_0 = \rho^\star e_0$, which means that $e_0 \neq 0$ is an eigenvector of the matrix B, which implies that $\rho^\star \leq \rho(B)$. Thus, according to (4.10) we obtain

$$\rho^\star = \rho(B). \tag{4.11}$$

Moreover from (4.8) we get

$$Be_\beta \leq \rho(B_\beta)e_\beta. \tag{4.12}$$

Let us consider now the transpose B^t_β of the matrix B_β. The matrix B^t_β is also positive and irreducible. By a similar way, a strictly positive eigenvector $e^\star_\beta$ is associated with B^t_β, such that $B^t_\beta e^\star_\beta = (B^t + \beta A)e^\star_\beta = \rho(B_\beta)e^\star_\beta$. Therefore $B^t e^\star_\beta \leq \rho(B_\beta)e^\star_\beta$, which together with (4.12), allows us to use the result of Lemma 4.2 with $\lambda = \rho(B_\beta)$. Since by (4.10), (4.11) and (4.12), assumption (4.1) is satisfied, the proof is complete. □

The following result follows from the Perron-Frobenius Theorem.

COROLLARY 4.4. *Assume that the matrix B is non-negative and irreducible; then*

$$]|B|[_{ee^\star,p} = \rho(B),\ \forall p \in [1,\infty],$$

where $\rho(B)$ is the spectral radius of matrix B.

5. Approximate contraction for linear parallel asynchronous iterations. We consider now the case where $B \in L(\mathbb{R}^m)$ is not necessarily nonnegative and $c \in \mathbb{R}^m$. Consider the affine mapping

$$\bar{T}v = Bv + c. \tag{5.1}$$

Let T be the perturbation of the mapping $\bar{T}$; in practice T is the floating point realisation of the mapping $\bar{T}$ on a computer.

LEMMA 5.1. *Let $|B| = (|b_{ij}|)$ and assume that*

$$\rho(|B|) < 1, \tag{5.2}$$

and

$$q(Tv - \bar{T}v) \leq \tau(|B|q(v) + q(c)),\ \forall v \in \mathbb{R}^m, \tag{5.3}$$

where $q(.)$ is the vectorial norm defined by

$$q(v) = (|v_1|, .., |v_i|, .., |v_m|), \tag{5.4}$$

and τ is a positive real number. Then the following inequality holds

$$q(Tv - u^\star) \leq (1+\tau)\ |B|q(v-u^\star) + \tau \left(\sum_{k=0}^{\infty} |B|^k\right) q(c),\ \forall v \in \mathbb{R}^m. \tag{5.5}$$

Proof. First of all,

$$q(\bar{T}v - u^\star) = q(B(v-u^\star)), \forall v \in \mathbb{R}^m.$$

Since the components of the vectorial norm are the absolute values of the components of the vector $B(v-u^\star)$, the previous equation leads to

$$q(\bar{T}v-u^\star)\leq |B|q(v-u^\star),\forall v\in\mathbb{R}^m. \tag{5.6}$$

Furthermore, from the triangle inequality we obtain

$$q(Tv-u^\star)\leq q(Tv-\bar{T}v)+q(\bar{T}v-u^\star),\forall v\in\mathbb{R}^m. \tag{5.7}$$

It follows from (5.3), (5.6) and (5.7) that we have

$$q(Tv-u^\star)\leq \tau(|B|q(v)+q(c))+|B|q(v-u^\star),\forall v\in\mathbb{R}^m, \tag{5.8}$$

and we obtain

$$q(Tv-u^\star)\leq \tau(|B|q(v-u^\star)+|B|q(u^\star)+q(c))+|B|q(v-u^\star),\forall v\in\mathbb{R}^m. \tag{5.9}$$

Thus, we have

$$q(Tv-u^\star)\leq (1+\tau)|B|q(v-u^\star)+\tau(|B|q(u^\star)+q(c)),\forall v\in\mathbb{R}^m. \tag{5.10}$$

Since $u^\star=\bar{T}u^\star=Bu^\star+c$, we have

$$q(u^\star)\leq |B|q(u^\star)+q(c).$$

Thus,

$$(I-|B|)q(u^\star)\leq q(c).$$

By assumption (5.2), $(I-|B|)$ is an M-matrix (see [15] and [17]) and we obtain

$$q(u^\star)\leq (I-|B|)^{-1}q(c).$$

Thus, inequality (5.10) becomes

$$q(Tv-u^\star)\leq (1+\tau)|B|q(v-u^\star)+\tau(|B|(I-|B|)^{-1}+I)q(c),\ \forall v\in\mathbb{R}^m. \tag{5.11}$$

From assumption (5.2) and the fact that $(I-|B|)^{-1}$ is nonnegative, it follows that

$$|B|(I-|B|)^{-1}=\sum_{k=1}^{\infty}|B|^k.$$

Thus, inequality (5.11) can be written as

$$q(Tv-u^\star)\leq (1+\tau)\,|B|\,q(v-u^\star)+\tau\left(\sum_{k=0}^{\infty}|B|^k\right)\,q(c),\ \forall v\in\mathbb{R}^m, \tag{5.12}$$

and the lemma is true. □

REMARK 7. *If we consider the vectorial norm of $\bar{T}v$, then we obtain*

$$q(\bar{T}v)=q(Bv+c)\leq q(Bv)+q(c),\ \forall v\in\mathbb{R}^m.$$

It follows from the particular choice of the vectorial norm (5.4) that we obtain the following inequality

$$q(Bv+c)\leq |B|q(v)+q(c),\ \forall v\in\mathbb{R}^m.$$

So (5.3) is the vectorial analogue of assumption (2.12). In the sequel, we will show that assumption (5.3) is satisfied and is sufficient in order to study roundoff errors in the case of parallel asynchronous linear iterations.

Let us assume that the matrix B satisfies assumption (5.2). Then there exists

$$\lambda \in [\rho(|B|), \rho(|B|) + \epsilon] \subset [0, 1[, \tag{5.13}$$

and similar to (4.1) there also exists a strictly positive vector e, such that

$$|B|e \leq \lambda e, \tag{5.14}$$

thus,

$$]|B|[_{e,\infty} =]|\,|B|\,|[_{e,\infty} \leq \lambda, \tag{5.15}$$

where $\mathbb{R}^m$ is endowed with the weighted maximum norm (4.3), and (5.13) to (5.15) follow from Proposition 4.3 in which we take $p = \infty$ and B is replaced by $|B|$. We can deduce from Lemma 5.1 the following result.

PROPOSITION 5.2. *Assume that the space $\mathbb{R}^m$ is normed by the weighted maximum norm (4.3). Let B be a matrix such that assumption (5.2) is satisfied. Assume also that (5.3) is satisfied. Moreover suppose that*

$$\tau < \frac{1-\lambda}{\lambda}. \tag{5.16}$$

Then, the associated fixed point mapping T is a-contracting, with respect to the weighted maximum norm (4.3) with contraction constant $l = (1+\tau)\lambda$ and approximation constant

$$\theta_\star = \left(\frac{\tau}{1-\lambda}\right) ||c||_{e,\infty}.$$

Proof. Let y, x and z be three strictly positive vectors of dimension m, and d a real positive number such that

$$y \leq d|B|x + z.$$

Then, for a given monotone scalar norm $||\ .\ ||$, we can obtain

$$||y|| \leq ||d|B|x + z|| \leq d]|B|[||x|| + ||z||,$$

where $]|B|[$ is the subordinate matrix norm of B associated with the scalar norm $||\ .\ ||$. In particular, we have

$$||y||_{e,\infty} \leq ||d|B|x + z||_{e,\infty} \leq d]|B|[_{e,\infty}||x||_{e,\infty} + ||z||_{e,\infty}.$$

Applying the previous inequalities to (5.5) we obtain

$$||Tv - u^\star||_{e,\infty} \leq (1+\tau)\ ]|B|[_{e,\infty}||v - u^\star||_{e,\infty} + \tau \left(\sum_{k=0}^{\infty}]|B|[^k_{e,\infty}\right) ||c||_{e,\infty}, \forall v \in \mathbb{R}^m,$$

and it follows from the considered assumptions that

$$||Tv - u^\star||_{e,\infty} \leq (1+\tau)\ \lambda\ ||v - u^\star||_{e,\infty} + \tau \left(\sum_{k=0}^{\infty} \lambda^k\right) ||c||_{e,\infty}, \forall v \in \mathbb{R}^m,$$

ETNA
Kent State University
etna@mcs.kent.edu

and we obtain the following inequality

$$\|Tv - u^{\star}\|_{e,\infty} \leq (1+\tau)\,\lambda\,\|v - u^{\star}\|_{e,\infty} + \frac{\tau}{1-\lambda}\,\|c\|_{e,\infty}, \forall v \in \mathbb{R}^m,$$

and the linear fixed point mapping is a-contracting. Thus, the proof is complete. □

According to the results recalled in section 3, we can easily derive the following result

COROLLARY 5.3. *If assumptions (5.2) and (5.13) to (5.15) hold, the parallel asynchronous iteration (3.1) associated with T, produces a sequence of iterates $\{u^n\}$ such that $a(\{u^n\}) \subset B_E\left(u^{\star}; \frac{\theta_\star}{1-\lambda}\right)$.*

Proof. We use the result of Theorem 3.1 and Corollary 3.2, where according to (5.16), assumption (2.15) is satisfied. □

COROLLARY 5.4. *Consider the parallel asynchronous linear fixed point method (3.1). Under the assumptions of Theorem 3.1 and Corollary 3.2, the limit of subsequences of $\{u^n\}$ produced by the parallel asynchronous iteration belongs to the ball $B_E\left(u^{\star}; \frac{\theta_\star}{1-\lambda}\right)$, where, according to (2.15), $(1+\tau)\lambda < 1$ and $\theta_\star = \left(\frac{\tau}{1-\lambda}\right)\|c\|_{e,\infty}$.*

REMARK 8. *Inequality (5.5) clearly shows an approximate contraction property with respect to the vectorial norm $q(.)$ defined by (5.4). The result of Proposition 5.2 shows that if (5.5) is satisfied, then the property of approximate contraction is satisfied for the weighted maximum norm, which extends a result of J.C. Miellou valid only in the case of classical contraction (see [12]).*

REMARK 9. *According to the results of section 4 (see in particular Corollary 4.4), the number λ will be chosen in the sequel as follows: if $|B|$ is an irreducible matrix, then $\lambda \equiv l = \rho(|B|)$ and if $|B|$ is a reducible matrix then $\lambda = \rho(|B|) + \epsilon$.*

REMARK 10. *We consider now the successive approximation method. If the assumptions of Proposition 5.2 hold, then according to the result of Lemma 4.2, we can obtain in a similar way the analogue of the previous a-contraction property for the linear fixed point mapping, in the space $\mathbb{R}^m$ normed by the p-norm (4.2) for all $p \in [1, \infty[$,*

$$\|Tv - u^{\star}\|_{ee^\star,p} \leq (1+\tau)\,\lambda\,\|v - u^{\star}\|_{ee^\star,p} + \frac{\tau}{1-\lambda}\,\|c\|_{ee^\star,p}.$$

For all $p \in [1, \infty[$, the approximation constant related to the perturbation of the fixed point mapping is given by

$$\theta_\star = \left(\frac{\tau}{1-\lambda}\right)\,\|c\|_{ee^\star,p},\ \forall p \in [1, \infty[.$$

If we consider now the successive approximation method (2.8) described in section 2, then under the assumptions of Theorem 2.5, the limit of subsequences of $\{u^n\}$ produced by the iteration (2.8) belongs to the ball $B_E\left(u^{\star}; \frac{\theta_\star}{1-l}\right)$.

REMARK 11. *According to the result of Proposition 5.2 and Remarks 6 and 10, we have obtained for the affine mapping T, the property of a-contraction in $\mathbb{R}^m$ normed by $\|\,.\,\|_{ee^\star,p}, \forall p \in [1, \infty]$.*

6. Application to roundoff errors in the case of parallel asynchronous linear iterations. It is well known that arithmetic operations are affected by roundoff errors when calculations are performed on a computer. According to the books of N.J. Higham (see [9]) and G.H. Golub and C.F. Van Loan (see [10]), it is possible to define the fl operator which satisfy $fl(x) = x(1+\epsilon), |\epsilon| \leq \chi$, where χ is defined by $\chi = \frac{1}{2}\hat{b}^{1-s}$, in the case of rounding and $\chi = \hat{b}^{1-s}$, in the case of chopping, where $\hat{b}$ denotes the base and s the precision. Let g

ETNA
Kent State University
etna@mcs.kent.edu

and h be any two floating point numbers and let *op* denote any of the four basic arithmetic operations. Then in the model of floating point arithmetic defined by N.J. Higham (see [9]) and G.H. Golub and C.F. Van Loan (see [10]), it is assumed that the computed version of $g\ op\ h$ is given by $fl(g\ op\ h)$. It follows that

$$fl(g\ op\ h) = (g\ op\ h)(1+\epsilon), |\epsilon| \leq \chi.$$

Thus

$$\frac{|fl(g\ op\ h) - (g\ op\ h)|}{|g\ op\ h|} \leq \chi, \text{ if } g\ op\ h \neq 0,$$

which shows that the relative error associated with individual arithmetic operations is small.

Consider now the parallel asynchronous linear iteration (3.1) associated with the exact mapping $\bar{T}$ defined by (5.1). Note that the components of vector u^n are obtained by making m dot products of vectors from $\mathbb{R}^{m+1}$. Let y and x be two vectors of $\mathbb{R}^{m+1}$; it was established by G.H. Golub and C.F. Van Loan (see [10]) and also by N.J. Higham (see [9]) that there exists a positive number μ = 1.0101 such that

$$|fl(x^t y) - x^t y| \leq \mu\ (m+1)\ \chi\ q(x)^t q(y),$$

where $q(.)$ is the vectorial norm defined by (5.4).

By applying the previous result to the case considered in this paper, we finally obtain the following inequality

$$|(\bar{T}u)_i - (Tu)_i| = \left| \sum_{j=1}^{m} b_{ij}u_j + c_i - fl\left(\sum_{j=1}^{m} b_{ij}u_j + c_i\right)\right| \leq \mu(m+1)\chi\left(\sum_{j=1}^{m} |b_{ij}||u_j| + |c_i|\right)$$

which leads to

$$q(\bar{T}u - Tu) \leq \mu\ (m+1)\ \chi\ (|B|\ q(u) + q(c)), \tag{6.1}$$

and assumption (5.3) is well verified, with $\tau = \mu(m+1)\chi$; thus, if the approximation constant $l = (1 + \mu(m+1)\chi)\lambda$ is strictly less than one, the result of Corollary 5.3 holds.

REMARK 12. *The previous estimate depends on m, the dimension of the system to be solved. If the system is a large scale system, then the number $\tau = \mu(m+1)\chi$ can be considerably large. In the case of a sparse matrix, it is possible to improve the above estimation by replacing the term $(m+1)$ by the maximum number of nonzero elements in a row of the matrix B, denoted by t, and we obtain*

$$q(\bar{T}u - Tu) \leq\ \mu\ (t+1)\ \chi\ (|B|\ q(u) + q(c)).$$

We refer to [14] for a study of some stopping criteria, forward and backward errors with respect to roundoff errors of fixed point methods including
- the specific case of the successive approximation method considered in the topological context of a large familly of p-norms (4.2),
- the general situation of asynchronous iterations by using weighted maximum norm (4.3).

7. Examples. We present in this section two simple examples which illustrate the present study.

ETNA
Kent State University
etna@mcs.kent.edu

7.1. Example 1. Consider a strictly positive vector $e \in \mathbb{R}^m$ with components $e_i > 0, \forall i = 1, .., m$. Note that with the convention $e_{m+1} = e_1$, we have

$$\prod_{i=1}^{m} \frac{e_i}{e_{i+1}} = 1; \tag{7.1}$$

consider also two matrices B_1 and B_2 defined by

$$B_1 = \begin{vmatrix} 0 & b^1_{1,2} & 0 & 0 & \dots & 0 \\ 0 & 0 & b^1_{2,3} & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 0 & 0 & b^1_{m-1,m} \\ b^1_{m,1} & 0 & \dots & 0 & 0 & 0 \end{vmatrix},$$

$$B_2 = \begin{vmatrix} 0 & 0 & 0 & 0 & \dots & b^2_{1,m} \\ b^2_{2,1} & 0 & 0 & 0 & \dots & 0 \\ 0 & b^2_{3,2} & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 & 0 \\ 0 & 0 & \dots & 0 & b^2_{m,m-1} & 0 \end{vmatrix},$$

the entries of which are defined by

$$b^1_{i,i+1} = \frac{e_i}{e_{i+1}}, b^1_{m,1} = \frac{e_m}{e_1}, b^2_{i,i-1} = \frac{e_i}{e_{i-1}} \text{ and } b^2_{1,m} = \frac{e_1}{e_m}.$$

Note that it can be easily verified that

$$B_1 e = e \text{ and } B_2 e = e. \tag{7.2}$$

Thus, according to (7.1), the eigenvalues of the matrices B_1 and B_2 satisfy

$$(\lambda_i)^m - 1 = 0\,, \forall\, i = 1, 2.$$

Then, according to (7.2), their spectral radii are equal to one.

Consider now the matrix B, which is a linear combination of the matrices B_1 and B_2, such that

$$B = \rho_1 B_1 + \rho_2 B_2, \rho_1 > 0 \text{ and } \rho_2 > 0,$$

then the matrix B is a non-negative irreducible matrix. Assume also that $1 > \rho_1 >> \rho_2$; thus $\rho_1 B_1$, $\rho_2 B_2$, respectively, correspond to the strong and weak weights, respectively, arising in the matrix B; assume also that

$$\rho = \rho_1 + \rho_2 < 1. \tag{7.3}$$

Thus, according to (7.2), we obtain easily

$$Be = (\rho_1 + \rho_2)e.$$

Consider now the solution of the following fixed point problem

$$u^* = Bu^* + c, \tag{7.4}$$

where c is a vector of $\mathbb{R}^m$.

REMARK 13. *Note that, if we consider the numerical solution of a 1-dimensional convection-diffusion problem with periodic boundary conditions, which are defined by*

$$\begin{cases} -\epsilon \frac{d}{dx}(a(x)\frac{du}{dx}) + b(x)\frac{du}{dx} + q(x)u = f, \text{ everywhere in } \Omega =]0,1[\,, \\ u(0) = u(1),\ \frac{du(0)}{dx} = \frac{du(1)}{dx}, \end{cases}$$

with $q(x) > \bar{q} > 0$, $\epsilon > 0$, then we can obtain a discretization matrix with the same shape as the matrix B. If furthermore the convection is dominant, then we can obtain a situation where strong and weak weights occur in the solution of a fixed point problem of the kind (7.4).

In order to solve the equation (7.4) consider a sequential chaotic algorithm (see [12]) where the strong and the weak weights are taken into account alternatively; more precisely, starting from u^0, ν iterations are carried out in order to solve the fixed point equation

$$u^{n+1} = \rho_1 B_1 u^n + \bar{c}, \tag{7.5}$$

where ν is a given integer and $\bar{c} = \rho_2 B_2 u^0 + c$. So after ν chaotic iterations of algorithm (7.5), we obtain easily the following estimation

$$||u^\nu - u^*||_{e,\infty} \le \left(\rho_1^\nu + \rho_2 \frac{1-\rho_1^\nu}{1-\rho_1}\right) ||u^0 - u^*||_{e,\infty}.$$

Note that the previous inequality, corresponds to an approximate contraction property. Then, starting from u^ν, only one classical successive approximation iteration is then performed for the global fixed point iteration, and we obtain for this chaotic algorithm

$$||u^{\nu+1} - u^*||_{e,\infty} \le (\rho_1 + \rho_2) \left(\rho_1^\nu + \rho_2 \frac{1-\rho_1^\nu}{1-\rho_1}\right) ||u^0 - u^*||_{e,\infty}. \tag{7.6}$$

Then, starting from $u^{\nu+1}$, we repeat the chaotic algorithm (7.5) ν times followed by only one global iteration.

If we compare now, the previous iteration scheme to the classical successive approximation scheme for a similar computational cost (i.e. for $(\frac{\nu}{2}+1)$ complete successive approximation iterations) the estimation (7.6) must be compared to

$$||u^{\frac{\nu}{2}} - u^*||_{e,\infty} \le (\rho_1 + \rho_2)^{\frac{\nu}{2}+1} ||u^0 - u^*||_{e,\infty}. \tag{7.7}$$

Then, the chaotic algorithm will perform better, if the following inequality is satisfied

$$\left(\rho_1^\nu + \rho_2 \frac{1-\rho_1^\nu}{1-\rho_1}\right) < (\rho_1 + \rho_2)^{\frac{\nu}{2}}. \tag{7.8}$$

For convenient values of ν, it can be verified that the previous inequality is satisfied. Moreover it can be noted that

$$\lim_{\nu\to\infty} (\rho_1 + \rho_2)^{\frac{\nu}{2}} = 0 \text{ and } \lim_{\nu\to\infty} \left(\rho_1^\nu + \rho_2 \frac{1-\rho_1^\nu}{1-\rho_1}\right) = \frac{\rho_2}{1-\rho_1}. \tag{7.9}$$

So the curves $\nu \to v1(\nu) = \left(\rho_1^\nu + \rho_2 \frac{1-\rho_1^\nu}{1-\rho_1}\right)$ and $\nu \to v2(\nu) = (\rho_1 + \rho_2)^{\frac{\nu}{2}}$ meet for only one value of ν. Indeed it can be verified that the curve $\nu \to v2(\nu)$ is strictly decreasing

TABLE 7.1
ν_{max} *as a function of* ρ_1

ρ_1	ν_{max}	ρ_1	ν_{max}
0.9	47	0.4	9
0.8	28	0.3	7
0.7	19	0.2	5
0.6	14	0.1	4
0.5	11	0.05	3

and the curve $\nu \to v1(\nu)$ is not increasing; moreover for very small values of ν we have $v1(\nu) < v2(\nu)$ and (7.9) is valid (for more details the reader is referred to figures 1 and 2 of Appendix). More precisely, Table 7.1 shows the maximum value of ν, denoted by ν_{max}, for which the inequality (7.8) is satisfied as a function of ρ_1, ρ_2 being fixed to $\rho_2 = 0.01$. The previous experimentation shows that the greater the value of ρ_1, the larger the maximal value of ν_{max}.

Then, according to the results of Proposition 5.2 and (6.1), when the above chaotic context is satisfied, the iterate vector is localized in a ball of center u^* and of radius δ_*, where

$$\delta_* = \frac{\mu(m+1)\chi}{(1-\rho_1-\rho_2)(1-(1+\mu(m+1)\chi)(\rho_1+\rho_2))}||c||_{e,\infty},$$

and

$$u^n \in B_E\left(u^*; \frac{\tau}{(1-\rho_1-\rho_2)(1-(1+\tau)(\rho_1+\rho_2))}||c||_{e,\infty}\right),$$

where $\tau = \mu(m+1)\chi$, and $\mu = 1.0101$ according to [9].

7.2. Example 2. Consider the 3x3 nonnegative matrix defined by

$$B = \begin{vmatrix} 0 & 0 & b \\ c & 0 & 0 \\ 0 & d & 0 \end{vmatrix},$$

where b , c , d are three positive numbers. Note that this case corresponds to a particular case of the matrix B_2 considered in the previous example. After simple calculations, we can compute the spectral radius of the matrix B to be

$$\rho(B) = \sqrt[3]{bcd}.$$

Note that the above matrix is nonnormal, because

$$BB^t - B^tB = Diag(b^2 - c^2, c^2 - d^2, d^2 - b^2); \tag{7.10}$$

so, in this case such a matrix can display a signifiant amount of spectral instability in finite precision computation even if $\rho(B) < 1$ (see [4]).

By choosing, for example, $b = 1$, $d = \epsilon$ and $c = \frac{1}{2\epsilon}$, we obtain

$$\rho(B) = \frac{1}{\sqrt[3]{2}} < 1;$$

furthermore

$$e_1 = \sqrt[3]{2}\, e_3,\ e_2 = \frac{e_3}{\epsilon\,\sqrt[3]{2}},$$

and

$$e_1^\star = \sqrt[3]{2}\, e_3^\star,\, e_2^\star = \frac{\epsilon\, e_3^\star}{\sqrt[3]{2}}.$$

Thus, $\|x\|_{e,\infty}$ and $\|x\|_{ee^\star,p}$ can be easily computed for every x and for every $p \in [1,\infty[$. Consequently, according to the previous results, the mapping $Tu = Bu + c$ is approximately contracting, and the sequence $\{u^n\}$, $n \in \mathbb{N}$ is such that

$$u^n \in B_E\left(u^\star; \frac{\tau}{(1-\lambda)(1-(1+\tau)\lambda)}\|c\|_{ee^\star,p}\right), \forall p \in [1,\infty], \forall n \in \mathbb{N},$$

where τ and μ are previously defined.

According to (7.10), for small values of ϵ, this case corresponds to a property of high nonnormality of the matrix B, for which all our results apply in the topological framework of the weighted norms used, provided that we stay outside of underflow or overflow situations.

Acknowledgments. The authors wish to thank the referee for his helpful comments and suggestions.

REFERENCES

[1] G. BAUDET, *Asynchronous iterative methods for multiprocessors*, J. Assoc. Comput. Mach., 25 (1978), no. 2, pp. 226–244.
[2] D. P. BERTSEKAS AND J. N. TSITSIKLIS, *Parallel and distributed computation: numerical methods*, Prentice Hall, Englewood Cliffs, N.J., 1987.
[3] F. CHAITIN-CHATELIN AND V. FRAYSSE, *Lectures on finite precision computations*, SIAM Publications, Philadelphia, PA., 1996.
[4] F. CHAITIN-CHATELIN AND S. GRATTON, *Convergence in finite precision of successive iteration methods under high nonnormality*, BIT, 36 (1996), no. 3, pp. 455–469.
[5] D. CHAZAN AND W. MIRANKER, *Chaotic relaxation*, Linear Algebra Appl., 2 (1969), pp. 199–222.
[6] M. N. EL TARAZI, *Some convergence results for asynchronous algorithms*, Numer. Math., 39 (1982), no. 3, pp. 325–340.
[7] A. FROMMER, H. SCHWANDT AND D. SZYLD, *Asynchronous weighted additive Schwarz methods*, Electron. Trans. Numer. Anal., 5 (1997), June, pp. 48-61.
[8] A. FROMMER AND D. SZYLD, *On asynchronous iterations*, Numerical Analysis 2000, Vol. III. Linear Algebra. J. Comput. Appl. Math., 123 (2000), no. 1-2, pp. 201–216.
[9] N. J. HIGHAM, *Accuracy and stability of numerical algorithms*, SIAM Publications, Philadelphia, PA., 1996.
[10] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, Second ed., The Johns Hopkins University Press, Baltimore, MD, 1989.
[11] M. A. KRASNOSEL'SKII, G. M. VAINIKKO, P. P. ZABREIKO, YA. B. RUTITSKII AND V. YA. STETSENKO, *Approximate solution of operator equations*, Wolters-Noordhoff Publishing, Groningen, 1972.
[12] J. C. MIELLOU, *Itérations chaotiques à retards*, RAIRO, R1 (1975), pp. 55–82.
[13] J. C. MIELLOU, P. CORTEY-DUMONT AND M. BOULBRACHÊNE , *Perturbation of fixed point iterative methods*, Advances in Parallel Computing, I (1990), pp. 81–122.
[14] J. C. MIELLOU, P. SPITERI AND D. EL BAZ, *Perturbation of fixed point methods by round off errors : stopping criteria, forward and backward errors*, Preprint of the University of Franche - Comté, 2002.
[15] J. ORTEGA AND W. C. RHEINBOLDT, *Iterative solution of non linear equations in several variables*, Academic Press, 1970.
[16] H. H. SCHAEFER, *Banach lattices and positive operators*, Springer Verlag, New York - Heidelberg - Berlin, 1974.
[17] R. S. VARGA, *Matrix iterative analysis*, Springer Verlag, 2000.
[18] J. H. WILKINSON, *Rounding errors in algebraic processes*, Prentice-hall, Englewood Cliffs, NJ., 1963.

Appendix

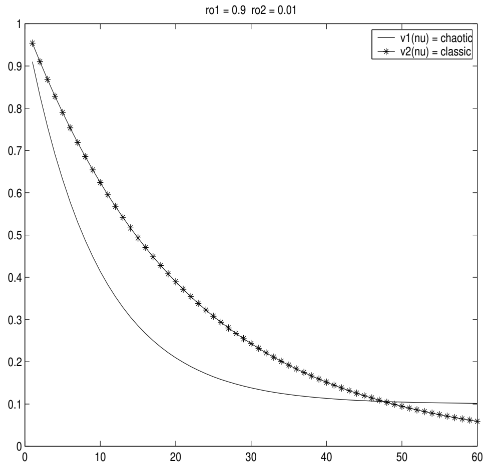

figure 1 - $\rho_1 = 0.9$

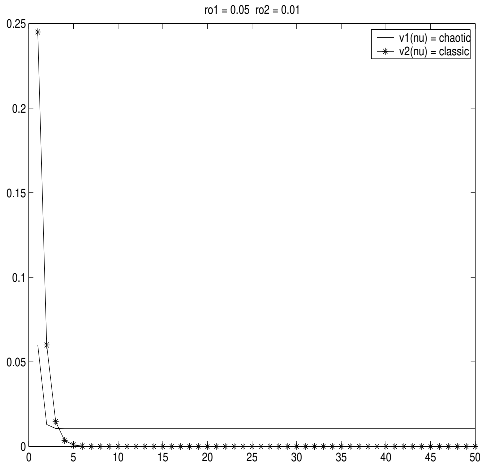

figure 2 - $\rho_1 = 0.05$