Stopping criteria, forward and backward errors for perturbed asynchronous linear fixed point methods in finite precision

Jean-Claude Miellou, Pierre Spiteri, Didier El Baz

To cite this version:


HAL Id: hal-01153719
https://hal.archives-ouvertes.fr/hal-01153719
Submitted on 20 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Stopping criteria, forward and backward errors for perturbed asynchronous linear fixed point methods in finite precision

J.C. Miellou, P. Spiteri, D. El Baz

October 5, 2004

Abstract

This paper deals with perturbed linear fixed point methods in the presence of round off errors. Successive approximations, as well as the more general asynchronous iterations are treated. Forward and backward error estimates are presented, and these are used to propose theoretical stopping criteria for these methods. In the case of asynchronous iterations, macro iterations are used as a tool in order to obtain estimates.

Keywords. approximate contraction, parallel algorithms, asynchronous iterations, fixed point methods, successive approximation methods, round off errors, stopping criterion, forward errors, backward errors.

AMS Subject Classification. 65F10, 65G05, 65Y05, 68Q22, 68Q10.

1 Introduction

The termination of parallel asynchronous iterations is a complex and crucial problem. Thus, it is of primary importance to study stopping criteria for such iterations. In this paper, we consider in particular several theoretical
topics related to stopping criteria for parallel asynchronous iterations in the case of perturbation by floating point errors of affine mappings. This issue is particularly complex due to the nondeterminism and lack of synchronization. In [24] we have studied perturbations of parallel asynchronous fixed point methods by round off errors. In this paper, we give some extensions related to forward and backward errors (see [6] and [16]). We will consider first the case of successive approximation methods. Then, we will consider the more general case of parallel asynchronous iterations. In each case, we develop a theoretical analysis on how to choose a stopping criterion if one desires the resulting approximation to satisfy certain forward or backward error estimates. We point out that our contribution is theoretical and that we do not present here any application to actual computational situations. The reader is referred to [10] and to [11] for different studies concerning the termination of asynchronous iterations in nonperturbed and perturbed contexts, respectively.

There are many situations where the solution of a lot of auxiliary linear problems with the same matrix occur in a numerical simulation. This is the case in particular in numerical fluid dynamics simulation whereby the Navier-Stokes equation is solved by using auxiliary linear operators such as the discretized Laplacian (see [25] and [15]). Then, it can be possible and useful to perform some specific preprocessing with respect to the linear operator. The preprocessing will have the purpose to approximate the Perron-Frobenius eigenvalues of the associated Jacobi or Gauss-Seidel mapping issued from multi-overlapping domain decomposition. We note that such a preprocessing can also occur in a variant of the Schwarz alternating method whereby we have to solve a linear fixed point problem via asynchronous iterations in order to obtain the value of the solution on the boundary of the overlapping domain (see [20]). Another promising perspective is probably to derive extensions of the results presented in this paper to the case of general multisplitting techniques (see [3]).

In the sequel, we will use the following weighted norm variant of the convexity theorem of Riesz established in [24].

**Theorem 1** Let us consider a matrix $B \in \mathcal{L}(\mathbb{R}^n)$, with entries $b_{ij}$. Let us denote by $\rho(|B|)$ the spectral radius of the matrix $|B|$ with entries $|b_{ij}|$. Then, for all real numbers $\epsilon > 0, \epsilon$ small enough, there exist two positive vectors $e(\epsilon)$ and $e^*(\epsilon)$ and a scalar $\lambda_\epsilon \geq 0$, such that

\[
\begin{cases}
|B| e(\epsilon) \leq \lambda_\epsilon e(\epsilon), \\
|B^T| e^*(\epsilon) = |B| e^*(\epsilon) \leq \lambda_\epsilon e^*(\epsilon),
\end{cases}
\]
where \( \lambda \in [\rho(|B|), \rho(|B|) + \epsilon] \). Moreover for each \( p \in [1, \infty] \), we have

\[
\| B \|_{e,p} \leq \lambda,
\]

where \( \| \cdot \|_{e,p} \) is a matrix norm derived from the following Perron-Frobenius weighted \( p \)-norm

\[
\| x \|_{e,p} = \left[ \sum_{i=1}^{m} e_i(\epsilon) e_i^*(\epsilon) \frac{|x_i|^p}{e_i(\epsilon)} \right]^{\frac{1}{p}}.
\] (1.1)

**Remark 1** In the case of parallel asynchronous iterations, we will use the following weighted norm (see [12])

\[
\| x \|_{e,\infty} = \max_{1 \leq i \leq m} \left( \frac{|x_i|}{e_i(\epsilon)} \right),
\] (1.2)

which corresponds to the case where \( p = \infty \) and we have also \( \| B \|_{e,\infty} \leq \lambda \) (see [13]).

**Remark 2** The situation where \( \epsilon \) is equal to zero corresponds to the case where the matrix \( B \) is irreducible. Then \( e \) and \( e^* \), respectively, are eigenvectors of matrices \( |B| \) and \( |B|^T \), respectively, associated with the respective spectral radius.

The paper is organized as follows. In section 2 we analyze successive approximation methods for affine mappings and one of the main results deals with the adaptation of the approximate contraction concept to round off perturbations. Some results concerning usual stopping criteria are also given. Furthermore, an expression of the backward error and an associated stopping criterion for the successive approximation method are proposed. Moreover, a sharp bound on the condition number of the matrix arising in the affine problem, with respect to the Perron-Frobenius weighted \( p \)-norms (1.1) is given and aspects of backward errors are presented, for the successive approximation method. In section 3, we consider the case of general affine fixed point methods on a product space, i.e. of asynchronous iterative methods. The corresponding study needs a special case of the previous topological context of Perron-Frobenius weighted \( p \)-norms (1.1) to the maximum weighted norm (1.2). In the general case of asynchronous iterations, we derive stopping criteria based on bounds of the error by using the concept of macro-iteration developed in [18]. The previous results can be also extended to the case of periodic asynchronous iterations with multiple initializations in the specific case of affine mappings (see [21]).

3
2 Perturbation by round off errors of affine mappings

2.1 Notations and complementary background

Let $m$ be an integer. Let us consider two matrices $A, B \in \mathcal{L}(\mathbb{R}^m)$ such that

$$A = I - B.$$ 

Assume that the spectral radius of the matrix $|B|$ satisfies

$$\rho(|B|) < 1.$$  \hspace{1cm} (2.1)

Let $c \in \mathbb{R}^m$ be a vector; consider the following linear system

$$Au^* = c,$$  \hspace{1cm} (2.2)

and the associated fixed point mapping

$$u \in \mathbb{R}^m \rightarrow v = \mathcal{T}(u) = Bu + c.$$  \hspace{1cm} (2.3)

Clearly, $u^*$ is the unique solution of (2.2) if and only if $u^*$ is the unique fixed point of the mapping $\mathcal{T}$. In order to approximate the fixed point $u^*$, let us consider parallel asynchronous iterations $\{u^n\}_{n \in \mathbb{N}}$ defined recursively as follows (see [2]), for all $n \geq 0$ and $i \in \{1, \ldots, m\}$,

$$u_i^{n+1} = \begin{cases} 
\mathcal{T}_i(\ldots, u_j^{(n)}, \ldots), & \text{if } i \in J(n), \\
 u_i^n, & \text{if } i \not\in J(n), 
\end{cases}$$  \hspace{1cm} (2.4)

where $u^0$ is the initial guess and

$$J = \{J(n)\}_{n \in \mathbb{N}}$$ is a sequence of non void subsets of $\{1, 2, \ldots, m\}$,  \hspace{1cm} (2.5)

$$S = \{s_1(n), s_2(n), \ldots, s_m(n)\}_{n \in \mathbb{N}}$$ is a sequence of elements of $\mathbb{N}^m$.  \hspace{1cm} (2.6)

Moreover $J$ and $S$, satisfy

$$\forall i \in \{1, 2, \ldots, m\}, \text{Card}\{n \in \mathbb{N} \mid i \in J(n)\} = +\infty,$$  \hspace{1cm} (2.7)

$$\forall j \in \{1, 2, \ldots, m\}, \forall n \in \mathbb{N}, s_j(n) \leq n,$$  \hspace{1cm} (2.8)

$$\forall j \in \{1, 2, \ldots, m\}, \lim_{n \to \infty} s_j(n) = +\infty.$$  \hspace{1cm} (2.9)
In the framework of floating point computations and round off error perturbations, we must replace in (2.4) the exact mapping $\bar{T}$ by an approximate mapping denoted by $T$. In order to deal with this perturbation, let us introduce now the following notations (see [24]). We define the real number

$$\tau = \mu(t + 1)\chi,$$

(2.10)

where $t$ is the maximum number of nonzero elements in a row of the matrix $B$, $\mu$ is a positive constant equal to $\mu = 1.0101...$ (see [14] on page 63 and [16] on page 75), and $\chi$ is such that for any floating point number $y$ we have

$$fI(y) = y(1 + \delta), |\delta| \leq \chi.$$

In fact, if $\beta$ denotes the base and $s$ the precision, then $\chi$ is defined by $\chi = \frac{1}{2}\beta^{1-s}$ in the case of rounding and $\chi = \beta^{1-s}$ in the case of chopping (see [14] and [16]).

Let us also consider a strengthened form of the assumption (2.1) by substituting $B$ by $\frac{1}{1+\tau}B$, in which the spectral radius of the matrix $|B|$ satisfies

$$\rho(|B|) < \frac{1}{1+\tau}.$$  (2.11)

According to (2.11) and Theorem 1, $\epsilon$ can be chosen sufficiently small so that

$$\alpha = (1 + \tau)\lambda_{\epsilon} < 1.$$  (2.12)

Let us associate to $u \in R^m$ the vectorial norm $|u|$ with components $|u_i|, i = 1, \ldots, m$. By using (2.10), we can write (see [24])

$$|Tu - \bar{T}u| \leq \tau(|B||u| + |c|).$$  (2.13)

We present now an important property related to approximate contraction; see [19] and also [24] where the proof is given.

**Lemma 1** Assume that $\epsilon$ is sufficiently small so that (2.12) is valid. Then, the perturbed fixed point mapping satisfies

$$|u^* - Tv| \leq (1 + \tau)|B||v - u^*| + \tau(I - |B|)^{-1}|c|, \forall v \in R^m,$$

and for each $p \in [1, \infty]$ we have

$$\|u^* - Tv\|_{p,\epsilon} \leq (1 + \tau)\lambda_{\epsilon}\|u^* - v\|_{p,\epsilon} + \frac{\tau}{1 - \lambda_{\epsilon}}\|c\|_{p,\epsilon}, \forall v \in R^m.$$
Thus, for each $p \in [1, \infty]$, the mapping $T$ satisfies the approximate contraction property
\[ \| u^* - T v \|_{e,p} \leq \alpha \| u^* - v \|_{e,p} + \theta^*_p, \forall v \in R^m, \] (2.14)
where $\theta^*_p$ satisfies
\[ \theta^*_p = \frac{\tau}{1 - \lambda_e} \| c \|_{e,p}. \] (2.15)

2.2 The particular case of successive approximations

Let us concentrate now on the classical successive approximation method associated with the mapping $T$, which of course is included in the broad class of asynchronous iterative methods. In this case we have

\[ J(n) = \{1, \ldots, m\}, \forall n \in N, \]
and
\[ s_j(n) = 0, \forall n \in N, \forall j \in \{1, \ldots, m\}. \]

Using (2.14) and results in [19] we develop now a stopping criterion.

Proposition 1 Assume that (2.12) holds for $\epsilon$ sufficiently small. After $n$ iterations of the successive approximation method associated with the perturbed fixed point mapping $T$, we have the following estimate of the error for each $p \in [1, \infty]$
\[ \| u^* - u^n \|_{e,p} \leq \frac{1}{1 - \alpha}(\alpha^n \| u^1 - u^0 \|_{e,p} + \frac{\tau}{1 - \lambda_e} \| c \|_{e,p}). \] (2.16)

Proof. It follows from Lemma 1 that (2.14) holds. Then, we have
\[ \| u^* - u^n \|_{e,p} \leq \alpha^n \| u^* - u^0 \|_{e,p} + \frac{1 - \alpha^n}{1 - \alpha} \theta^*_p. \]

Furthermore,
\[ \| u^1 - u^0 \|_{e,p} \geq \| u^* - u^0 \|_{e,p} - \| u^* - u^1 \|_{e,p}, \]
and since
\[ \| u^* - u^1 \|_{e,p} \leq \alpha \| u^* - u^0 \|_{e,p} + \theta^*_p \]
we have
\[ \| u^* - u^0 \|_{e,p} \leq \frac{\| u^1 - u^0 \|_{e,p} + \theta^*_p}{1 - \alpha}. \]
Thus,
\[
\|u^* - u^n\|_{e,p} \leq \frac{\alpha^n}{1-\alpha} \|u^1 - u^0\|_{e,p} + \frac{\theta_p}{1-\alpha},
\]
and by using (2.15) the proof is complete. ■

The a priori stopping criterion (see [19]) can be stated as follows: for any \( p \in [1, \infty] \) and \( \eta > 0 \), let us evaluate the real number \( \bar{n}_\eta \) defined by
\[
\frac{\alpha^n \|u^1 - u^0\|_{e,p}}{1-\alpha} = \eta, \quad \text{i.e.} \quad \alpha^n = \eta \frac{1-\alpha}{\|u^1 - u^0\|_{e,p}}.
\]
(2.17)

Therefore,
\[
\bar{n}_\eta = \frac{1}{\log(\alpha)} \left( \log(\eta) + \log(1-\alpha) - \log(\|u^1 - u^0\|_{e,p}) \right).
\]

Then, let us define the integer \( n_\eta \) by
\[
n_\eta = \lceil \bar{n}_\eta \rceil = \min \{ n \in N; n \geq \bar{n}_\eta \};
\]
(2.18)

by using (2.16), we can assert that
\[
\|u^* - u^n\|_{e,p} \leq \eta + \frac{\tau}{(1-\alpha)(1-\lambda_e)} \|e\|_{e,p}, \forall n \geq n_\eta.
\]
(2.19)

**Remark 3** i) If the range \( R(T) \) of \( T \) is a finite set (see [19], on page 84), then, whatever the cardinality of \( R(T) \), there exists an integer \( n^* \) such that for each \( p \in [1, \infty] \)
\[
\|u^* - u^n\|_{e,p} \leq \frac{\tau}{(1-\alpha)(1-\lambda_e)} \|e\|_{e,p}, \forall n \geq n^*.
\]

However, it seems difficult to express a value of \( n^* \) with respect to the data of the problem.

ii) Consider the estimate (2.16) written at the first iteration; then
\[
\|u^* - u^1\|_{e,p} \leq \frac{1}{1-\alpha} \left( \alpha_0 \|u^1 - u^0\|_{e,p} + \frac{\tau}{1-\lambda_e} \|e\|_{e,p} \right).
\]

Analogously, at the \( n \)-th iteration we can write
\[
\|u^* - u^n\|_{e,p} \leq \frac{1}{1-\alpha} \left( \alpha_n \|u^n - u^{n-1}\|_{e,p} + \frac{\tau}{1-\lambda_e} \|e\|_{e,p} \right).
\]
(2.20)
Then, the corresponding stopping criterion reads as follows: for any $p \in [1, \infty]$ and the given real number $\eta \geq 0$, the successive approximation method is stopped when
\[
\frac{\alpha}{1 - \alpha} \|u^n - u^{n-1}\|_{e,p} \leq \eta, \quad \text{i.e.} \quad \|u^n - u^{n-1}\|_{e,p} \leq \eta \frac{1 - \alpha}{\alpha}.
\] (2.21)

By using (2.20), we can assert that, if the condition (2.21) is satisfied, then
\[
\|u^* - u^n\|_{e,p} \leq \eta + \frac{\tau}{1 - \alpha} \|c\|_{e,p}.
\] (2.22)

Practically, it can be noted that, if the term $\frac{\tau}{1 - \alpha} \|c\|_{e,p}$ is large, due to the fact that $\lambda_e$ is very close to 1 and $\|c\|_{e,p}$ is large, then, there is no need to look for good accuracy. For example, if we consider the case of ill conditioned problems, such as the discretized Laplacian problem or the discretized convection diffusion problem, the spectral radius of the iteration matrix can be close to one. In this case $\lambda_e$ will be also very close to one.

### 2.3 An estimate of the forward error of the perturbed successive approximation method

In the case where the range of $T$ is finite then, according to [19], whatever the cardinality of $R(T)$, there exists a finite integer $n^0$ such that for all $n \geq n^0$, $u^n \in B_{E}(u^*; \frac{\tau}{1 - \alpha})$, where the set $B_{E}$ is a closed ball of center $u^*$ and radius $\frac{\tau}{1 - \alpha}$. $R(T)$ is of course a finite set in the framework of finite precision arithmetics associated with the use of floating point numbers. Then we are able to give expressions of forward errors according to the work of N. J. Higham (see definitions of chapter 1, section 1-10 in [16]). Concerning the forward error, let (2.1) to (2.13) hold. For any $p \in [1, \infty]$ the condition number of the matrix $A = (I - B)$, with respect to the Perron-Frobenius norm is defined as
\[
\kappa_{e,p}(I - B) = \|I - B\|_{e,p} \|(I - B)^{-1}\|_{e,p}.
\]

If $\epsilon > 0$ is small enough so that $\lambda_e < 1$, then,
\[
\|I - B\|_{e,p} \leq 1 + \lambda_e, \quad \|(I - B)^{-1}\|_{e,p} \leq \frac{1}{1 - \lambda_e},
\]
and the condition number $\kappa_{e,p}(I - B)$ satisfies
\[
\kappa_{e,p}(I - B) \leq \hat{\kappa}(\lambda_e) = \frac{1 + \lambda_e}{1 - \lambda_e},
\] (2.23)
Proposition 2 Assume that (2.12) holds for \( c \) sufficiently small. Then, the forward error at the \( n - \) th iteration of the successive approximation method associated with the perturbed fixed point mapping \( T \) satisfies, for each \( p \in [1, \infty] \), the following estimate,

\[
\frac{\|u^* - u^n\|_{e,p}}{\|u^*\|_{e,p}} \leq \hat{\kappa}(\lambda_c) \left( \frac{\tau}{1 - \alpha} + \frac{1 - \lambda_c}{\lambda_c} \frac{\alpha^n \|u^1 - u^0\|_{e,p}}{\|e\|_{e,p}} \right). 
\] (2.24)

**Proof.** \( u^* \) being the fixed point of \( \tilde{T} \), it follows from (2.3) that

\[
|u^*| = |Bu^* + c| \geq |c| - |Bu^*| \geq |c| - |B||u^*|,
\]

which implies that

\[
(I + |B||u^*|) \geq |c|.
\]

Then, using the Perron-Frobenius norm, we obtain for each \( p \in [1, \infty] \),

\[
(1 + \lambda_c) \|u^*\|_{e,p} \geq \|c\|_{e,p},
\]

which implies

\[
\frac{1}{\|u^*\|_{e,p}} \leq \frac{1 + \lambda_c}{\|c\|_{e,p}}.
\] (2.25)

From Lemma 1, it follows that the approximate contraction property (2.14) holds. Multiplying member by member each side of (2.16) and (2.25), we can express the forward error as follows

\[
\frac{\|u^* - u^n\|_{e,p}}{\|u^*\|_{e,p}} \leq \frac{1 + \lambda_c}{1 - \alpha} \frac{\|c\|_{e,p}}{\|u^*\|_{e,p}} \left( \alpha^n \|u^1 - u^0\|_{e,p} + \frac{\tau}{1 - \lambda_c} \frac{\alpha^n \|u^1 - u^0\|_{e,p}}{\|c\|_{e,p}} \right).
\]

We can also write

\[
\frac{\|u^* - u^n\|_{e,p}}{\|u^*\|_{e,p}} \leq \hat{\kappa}(\lambda_c) \left( \frac{\tau}{1 - \alpha} + \frac{1 - \lambda_c}{\lambda_c} \frac{\alpha^n \|u^1 - u^0\|_{e,p}}{\|c\|_{e,p}} \right),
\]

which implies (2.24) by taking into account (2.12), and the proof is complete.

\[\blacksquare\]

As a consequence, let us modify the stopping criterion defined by (2.17) in the following way: for \( p \in [1, \infty] \) and for any real number \( \eta > 0 \), let us estimate the real number \( \tilde{n}_\eta \) defined by

\[
\frac{1}{\|e\|_{e,p}} \frac{\alpha^n \|u^1 - u^0\|_{e,p}}{\|c\|_{e,p}} = \eta, \; i.e. \; \alpha^n \eta = (1 - \frac{\tau}{1 - \lambda_c} \frac{\alpha^n \|u^1 - u^0\|_{e,p}}{\|c\|_{e,p}}) \eta.
\]
Then \( \bar{n}_\eta^* \) is given by

\[
\bar{n}_\eta^* = \frac{1}{\log(\alpha)} \left( \log(\eta) + \log(\gamma) + \log(\|c\|_{e,p}) - \log(\|u^1 - u^0\|_{e,p}) \right)
\]  

(2.26)

where \( \gamma = 1 - \frac{\tau \lambda_e}{1 - \lambda_e} \). Assuming that \( \frac{\tau \lambda_e}{1 - \lambda_e} \) is small enough, then

\[
\bar{n}_\eta^* \approx \frac{1}{\log(\alpha)} \left( \log(\eta) + \log(\|c\|_{e,p}) - \frac{\tau \lambda_e}{1 - \lambda_e} - \log(\|u^1 - u^0\|_{e,p}) \right);
\]

by using (2.26) and according to (2.18), let us define the integer \( n^*_\eta \) by \( n^*_\eta = \lceil \bar{n}_\eta^* \rceil \); thus, by using (2.24), we obtain

\[
\frac{\|u^*-u^n\|_{e,p}}{\|u^*\|_{e,p}} \leq \tilde{\kappa}(\lambda_e) \left( \frac{\tau}{1 - \alpha} + \frac{\alpha}{1 - \frac{\tau \lambda_e}{1 - \lambda_e}} \|c\|_{e,p} \right), \forall n \geq n^*_\eta.
\]  

(2.27)

**Remark 4** In a way similar to point ii) of Remark 3, the estimate (2.24) can be given here also at the first iteration. Then, analogously, we obtain at the \( n \)-th iteration

\[
\frac{\|u^*-u^n\|_{e,p}}{\|u^*\|_{e,p}} \leq \tilde{\kappa}(\lambda_e) \left( \frac{\tau}{1 - \alpha} + \frac{1}{1 - \frac{\tau \lambda_e}{1 - \lambda_e}} \frac{\alpha}{\|c\|_{e,p}} \right).
\]

The corresponding stopping criterion, based on the forward error, can be written as follows: for \( p \in [1, \infty] \) and \( \eta > 0 \) given, the successive approximation method is stopped when

\[
\frac{1}{1 - \frac{\tau \lambda_e}{1 - \lambda_e}} \frac{\alpha}{\|c\|_{e,p}} \leq \eta,
\]

which implies

\[
\|u^n - u^{n-1}\|_{e,p} \leq (1 - \frac{\tau \lambda_e}{1 - \lambda_e}) \frac{\|c\|_{e,p} \eta}{\alpha},
\]

and the corresponding forward error satisfies

\[
\frac{\|u^*-u^n\|_{e,p}}{\|u^*\|_{e,p}} \leq \tilde{\kappa}(\lambda_e) \left( \frac{\tau}{1 - \alpha} + \eta \right).
\]

10
2.4 Example

Consider an $m \times m$ matrix $\hat{A}$ and a vector $\hat{c} \in \mathbb{R}^m$. Let us denote by $\hat{a}_{ij}$ the entries of the matrix $\hat{A}$ and by $\hat{D}$ the diagonal part of $\hat{A}$. Let us assume that

$$\hat{A} \text{ is an H-matrix,}$$

that is to say, according to [4], that the comparison matrix $\mathcal{M}(A) = |\hat{D}| - |\hat{L}| - |\hat{U}|$ is an M-matrix, where $-\hat{L}$ and $-\hat{U}$ represent the lower and the upper parts of $\hat{A}$.

We consider now the situation related to the Jacobi method; in this case, our previous results can be used. Let us consider the matrix $B_J = I - \hat{D}^{-1}\hat{A}$; then, the assumption (2.28) is equivalent to

$$\rho(|B_J|) < 1, \text{ where } |B_J| = |I - \hat{D}^{-1}\hat{A}|.$$

Let us consider now the problem

$$\hat{A}u = \hat{c},$$

which is of course equivalent to the problem

$$u = B_J u + \hat{D}^{-1} \hat{c} = B_J u + c, \text{ where } c = \hat{D}^{-1} \hat{c}.$$  

For this particular situation, let us denote $\lambda_c$ by $\lambda_c^J$, where $\lambda_c^J \in [\rho(|B_J|), \rho(|B_J|) + \epsilon]$. Let us also denote by

$$\hat{\kappa}(\lambda_c^J) = \frac{1 + \lambda_c^J}{1 - \lambda_c^J},$$

the condition number associated with the Jacobi matrix and consider the real number

$$\alpha = \lambda_c^J (1 + \tau),$$

where $\tau$ is given by (2.10). Taking into account the round off perturbation error of the fixed point mapping associated with the Jacobi iteration, by using the result of Proposition 2, we obtain for any $p \in [1, \infty]$

$$\frac{\|u^* - u^n\|_{\epsilon,p}}{\|u^*\|_{\epsilon,p}} \leq \hat{\kappa}(\lambda_c^J) \left( \frac{\tau}{1 - \alpha} + \frac{1}{1 - \frac{\tau\lambda_c^J}{1 - \lambda_c^J}} \frac{\alpha^n\|u^{n-1} - u^0\|_{\epsilon,p}}{\|c\|_{\epsilon,p}} \right).$$

Moreover, (2.20) gives

$$\|u^* - u^n\|_{\epsilon,p} \leq \frac{1}{1 - \alpha} \left( \alpha\|u^n - u^{n-1}\|_{\epsilon,p} + \frac{\tau}{1 - \lambda_c^J} \|c\|_{\epsilon,p} \right),$$

11
and for any $p \in [1, \infty]$, thanks to (2.21) we obtain

$$\alpha \frac{\|u^{n} - u^{n-1}\|_{e,p}}{1 - \alpha} \leq \eta \text{ i.e. } \|u^{n} - u^{n-1}\|_{e,p} \leq \frac{\eta}{\alpha}.$$ 

So, when the previous inequality holds, according to (2.22), we obtain, for any $p \in [1, \infty]$, the estimate

$$\|u^{*} - u^{n}\|_{e,p} \leq \eta + \frac{\tau}{(1 - \alpha)(1 - \lambda^{j})} \|c\|_{e,p}.$$ 

By using the same argument, we can obtain an analogous estimate for the Gauss-Seidel method and more generally for any compatible splitting of an H-matrix $A$.

### 2.5 An estimate of the backward error for the perturbed successive approximation method

Let us assume that

$$\hat{\kappa}(\lambda_{e}) \frac{\tau}{1 - \alpha} < 1. \quad (2.30)$$

For each $p \in [1, \infty]$, the backward error is defined as follows (see [16])

$$\frac{\| (I - B)u^{n} - c \|_{e,p}}{\| I - B \|_{e,p} \| u^{n} \|_{e,p}} = \frac{\| (I - B)(u^{n} - u^{*}) \|_{e,p}}{\| I - B \|_{e,p} \| u^{n} \|_{e,p}}. \quad (2.31)$$

For each $p \in [1, \infty]$, we have the over-estimation

$$\frac{\| (I - B)u^{n} - c \|_{e,p}}{\| I - B \|_{e,p} \| u^{n} \|_{e,p}} \leq \frac{\| u^{*} - u^{n} \|_{e,p}}{\| u^{n} \|_{e,p}};$$

and because $\| u^{*} \|_{e,p} - \| u^{*} - u^{n} \|_{e,p} \leq \| u^{n} \|_{e,p}$, we obtain

$$\frac{\| (I - B)u^{n} - c \|_{e,p}}{\| I - B \|_{e,p} \| u^{n} \|_{e,p}} \leq \frac{\| u^{*} - u^{n} \|_{e,p}}{\| u^{n} \|_{e,p} - \| u^{*} - u^{n} \|_{e,p}},$$

which can be written finally as follows

$$\frac{\| (I - B)u^{n} - c \|_{e,p}}{\| I - B \|_{e,p} \| u^{n} \|_{e,p}} \leq \frac{\| u^{*} - u^{n} \|_{e,p}}{\| u^{n} \|_{e,p} \left( \frac{1}{1 - \| u^{*} - u^{n} \|_{e,p}} \right)}.$$ 

Clearly, since the left hand side of the last inequality is nonnegative, we must have

$$\| u^{*} - u^{n} \|_{e,p} \leq \| u^{*} \|_{e,p}. \quad (2.32)$$
Now, using (2.30), let us choose a real positive number $\eta$, such that

$$\hat{k}(\lambda_c)\left(\frac{\tau}{1-\alpha} + \eta\right) < 1.$$ 

Therefore for all $n \geq n^*_\eta$

$$\frac{(I - B)u^n - c}{\|I - B\|_{e,p}\|u^n\|_{e,p}} \leq \hat{k}(\lambda_c)\left(\frac{\tau}{1-\alpha} + \eta\right) \frac{1}{1 - \hat{k}(\lambda_c)\left(\frac{\tau}{1-\alpha} + \eta\right)}.$$  

3 The general case of asynchronous iterations

Let us now consider the general case of perturbed asynchronous iterations; then, we have to replace the mapping $T$ by its perturbation $T$ in (2.4). In this case, the asynchronous iteration produces a sequence of iterates, denoted by $\{u^n\}_{n \in \mathbb{N}}$, initialized by $u^0$, and defined as follows for all $n \geq 0$ and $i \in \{1, \ldots, m\}$

$$u_i^{n+1} = \begin{cases} T_i(...) \cdot u_j^{t(n)}(n), & \text{if } i \in J(n), \\ u_i^n, & \text{if } i \not\in J(n), \end{cases} \quad (3.1)$$

where the sequences $J$ and $S$ are defined respectively by (2.5) and (2.6) and the conditions (2.7) to (2.9) are satisfied. We point out that in the case of parallel asynchronous iterative methods, whereby computations are performed in parallel without any order nor synchronization, it is particularly important and challenging to derive criteria so that certain bounds on the error hold. These aspects are developed in the remaining of this section. In this context, the notion of macro iteration plays a major part. Let us consider the numbers

$$s(n) = \min_{r \geq n} \min_{1 \leq j \leq m} (s_j(r)). \quad (3.2)$$

Let us now define the sequence $\{n_k\}_{k \in \mathbb{N}}$ as follows

$$\begin{cases} n_0 = 0, \\ n_{k+1} = \min\{n|\bigcup_{n_k \leq s(t) \leq n} J(t) = \{1, \ldots, m\}\}. \quad (3.3)\end{cases}$$

In several works concerning the convergence of asynchronous iterations, the behavior of the successive iterates is based on the relations

$$u^n \in E^k, \forall n \geq n_k, \quad (3.4)$$
where \( \{E^k\}_{k \in \mathbb{N}} \) is a sequence of nested Cartesian product sets (see [5] and [19]), satisfying
\[
E^{k+1} \subset E^k, \forall k \in \mathbb{N}.
\] (3.5)

In the context of asynchronous iterations, we have to use the Perron-Frobenius weighted maximum norm (corresponding to the case \( p = \infty \), see [12]). In such a framework, when considering convergence properties of iterations (2.4) associated with the fixed point mapping \( \tilde{T} \), (see [12] and [18]) the nested sets are defined by
\[
E^k = \{ u \in \mathbb{R}^m \mid \| u^* - u \|_{e, \infty} \leq \alpha^k \| u^* - u^0 \|_{e, \infty} \}.
\]
In the case of asynchronous iterations (3.1) associated with the perturbed mapping \( T \), the nested sets are given as follows (see [19] and [24])
\[
E^k = \{ u \in \mathbb{R}^m \mid \| u^* - u \|_{e, \infty} \leq \alpha^k \| u^* - u^0 \|_{e, \infty} + \frac{1 - \alpha^k}{1 - \alpha} \theta^*_\infty \},
\] (3.6)
where \( \theta^*_\infty = \frac{\tau}{1 - \lambda_c} \| c \|_{e, \infty} \). Then, in the case of perturbed asynchronous iterations, we have the following result

**Proposition 3** Assume that (2.12) holds for \( \epsilon \) sufficiently small. Then, the absolute error of the perturbed asynchronous iterations (3.1) satisfies for all \( n \geq n_k \) and \( k \in \mathbb{N} \)
\[
\| u^* - u^n \|_{e, \infty} \leq \frac{1}{1 - \alpha} \left( \alpha^k \| u^1 - u^0 \|_{e, \infty} + \frac{\tau}{1 - \lambda_c} \| c \|_{e, \infty} \right),
\] (3.7)
and for all \( n \geq n_k \) and \( k \in \mathbb{N} \), the corresponding forward error satisfies the following estimate
\[
\frac{\| u^* - u^n \|_{e, \infty}}{\| u^* \|_{e, \infty}} \leq \hat{\epsilon}(\lambda_c) \left( \frac{\tau}{1 - \alpha} + \frac{1 - \alpha^k}{1 - \alpha} \frac{\alpha^k \| u^1 - u^0 \|_{e, \infty}}{\| c \|_{e, \infty}} \right).
\] (3.8)

**Proof.** It follows from Lemma 1 that (2.14) holds; then for all \( u \in \mathbb{R}^m \), we have
\[
\| u - Tu \|_{e, \infty} = \| u - u^* + u^* - Tu \|_{e, \infty} \geq \| u - u^* \|_{e, \infty} - \| u^* - Tu \|_{e, \infty}.
\]
Thus,
\[
\| u - Tu \|_{e, \infty} \geq (1 - \alpha) \| u - u^* \|_{e, \infty} - \theta^*_\infty.
\]
which implies
\[
\| u - u^* \|_{e, \infty} \leq \frac{1}{1 - \alpha} (\| u - Tu \|_{e, \infty} + \theta^*_{\infty}), \forall u \in E^k.
\]

Then by using (3.6)
\[
\| u - u^* \|_{e, \infty} \leq \frac{\alpha^k}{1 - \alpha} \left(\| u^0 - Tu^0 \|_{e, \infty} + \theta^*_{\infty} + \frac{1}{1 - \alpha} \theta^*_{\infty}\right), \forall u \in E^k.
\]

Therefore
\[
\| u - u^* \|_{e, \infty} \leq \frac{1}{1 - \alpha} \left(\alpha^k \| u^0 - Tu^0 \|_{e, \infty} + \theta^*_{\infty}\right), \forall u \in E^k.
\]

This shows that (3.7) is true. Furthermore, since \( u^* = (I - B)^{-1} c \) and \( \|c\|_{e, \infty} = \| (I - B)(I - B)^{-1} c \|_{e, \infty} \leq \| (I - B) \|_{e, \infty} (I - B)^{-1} \|c\|_{e, \infty} \), we obtain
\[
\frac{1}{\|u^*\|_{e, \infty}} = \frac{1}{\| (I - B)^{-1} c \|_{e, \infty}} \leq \frac{\| I - B \|_{e, \infty}}{\| c \|_{e, \infty}} \leq \frac{1 + \lambda_e}{\|c\|_{e, \infty}}.
\]

Consequently, by multiplying member by member each side of the previous inequality and (3.7), we obtain for each \( k \) and all \( n \geq n_k \)
\[
\frac{\| u^n - u^* \|_{e, \infty}}{\| u^* \|_{e, \infty}} \leq \frac{1 + \lambda_e}{1 - \alpha} \left(\frac{\tau}{1 - \lambda_e} + \frac{\alpha^k \| u^0 - Tu^0 \|_{e, \infty}}{\| c \|_{e, \infty}}\right).
\]

Since \( \hat{\lambda}(\lambda_e) = \frac{1 + \lambda_e}{1 - \lambda_e} \), \( \alpha = (1 + \tau) \lambda_e \) and
\[
\frac{1 - \lambda_e}{1 - \alpha} = \frac{1 - \lambda_e}{1 - \lambda_e - \tau \lambda_e} = \frac{1}{1 - \frac{\tau \lambda_e}{1 - \lambda_e}},
\]

we finally obtain for each \( k \) and all \( n \geq n_k \)
\[
\frac{\| u^n - u^* \|_{e, \infty}}{\| u^* \|_{e, \infty}} \leq \hat{\lambda}(\lambda_e) \left(\frac{\tau}{1 - \alpha} + \frac{1}{1 - \frac{\tau \lambda_e}{1 - \lambda_e}} \right),
\]
and the proof is achieved. 

Consequently, by using the bound on the error (3.7), we can provide a stopping criterion for general asynchronous iterations as follows: for a given real number \( \eta > 0 \), let us estimate the real number \( \tilde{k}_\eta \) defined by
\[
\frac{\alpha\tilde{k}_\eta \| u^1 - u^0 \|_{e, \infty}}{1 - \alpha} = \eta, \text{ i.e. } \alpha\tilde{k}_\eta = \eta \frac{1 - \alpha}{\| u^1 - u^0 \|_{e, \infty}}.
\]  

(3.9)
Then
\[ \tilde{k}_\eta = \frac{1}{\log(\alpha)} \left( \log(\eta) + \log(1 - \alpha) - \log(\|u^1 - u^0\|_{e, \infty}) \right). \]  
(3.10)

By using (3.10) and according to (2.18), let us define the integer \( k_\eta \) by \( k_\eta = \lceil \tilde{k}_\eta \rceil \); by using (3.7), we can assert that
\[ \|u^* - u^n\|_{e, \infty} \leq \eta + \frac{\tau}{(1 - \alpha)(1 - \lambda_e)} \|e\|_{e, \infty}, \forall n \geq n_{k_\eta}. \]  
(3.11)

With respect to the forward error, let us modify the previous criterion of the general asynchronous iterations defined by (3.9) to (3.11) in the following way: for a real number \( \eta > 0 \), let us evaluate the real number \( \tilde{k}^*_\eta \) defined by
\[ \frac{1}{1 - \frac{\tau \lambda_e}{1 - \lambda_e}} \frac{\alpha^{\tilde{k}^*_\eta} \|u^1 - u^0\|_{e, \infty}}{\|e\|_{e, \infty}} = \eta, \] i.e. \( \alpha^{\tilde{k}^*_\eta} = (1 - \frac{\tau \lambda_e}{1 - \lambda_e}) \frac{\|e\|_{e, \infty}}{\|u^1 - u^0\|_{e, \infty}} \eta; \)
so that \( \tilde{k}^*_\eta \) is given by
\[ \tilde{k}^*_\eta = \frac{1}{\log(\alpha)} \left( \log(\eta) + \log(\|e\|_{e, \infty}) - \log(\|u^1 - u^0\|_{e, \infty}) \right). \]  
(3.12)

where \( \gamma = 1 - \frac{\tau \lambda_e}{1 - \lambda_e} \). Assuming that \( \frac{\tau \lambda_e}{1 - \lambda_e} \) is small enough, then
\[ \tilde{k}^*_\eta \approx \frac{1}{\log(\alpha)} \left( \log(\eta) + \log(\|e\|_{e, \infty}) - \frac{\tau \lambda_e}{1 - \lambda_e} - \log(\|u^1 - u^0\|_{e, \infty}) \right). \]

Using (3.12) and according to (2.18), let us define the integer \( k^*_\eta \) by \( k^*_\eta = \lceil \tilde{k}^*_\eta \rceil \); so that, using (3.8), we obtain
\[ \frac{\|u^* - u^n\|_{e, \infty}}{\|u^*\|_{e, \infty}} \leq \tilde{k}^*_\eta \left( \eta + \frac{\tau}{1 - \alpha} \right), \forall n \geq n_{k^*_\eta}. \]  
(3.13)

**Remark 5** The above stopping criterion for asynchronous iterations, is based on the use of the macro iteration index \( k \). Nevertheless, we point out that the current value of \( k \) is not easy to obtain practically; note that procedures such as the one proposed in [10] can be implemented in order to compute the current value of \( k \) concurrently with the value of the iterate.

Concerning the backward error, we have
\[ \frac{\|(I - B)u^n - c\|_{e, \infty}}{\|I - B\|_{e, \infty} \|u^n\|_{e, \infty}} = \frac{\|(I - B)(u^n - u^*)\|_{e, \infty}}{\|I - B\|_{e, \infty} \|u^n\|_{e, \infty}}. \]
Therefore
\[
\frac{\| (I - B) u^n - c \|_{e, \infty}}{\| I - B \|_{e, \infty} \| u^n \|_{e, \infty}} \leq \frac{\| u^* - u^n \|_{e, \infty}}{\| u^n \|_{e, \infty}}.
\]
Because \( \| u^* \|_{e, \infty} - \| u^* - u^n \|_{e, \infty} \leq \| u^n \|_{e, \infty} \), we obtain
\[
\frac{\| (I - B) u^n - c \|_{e, \infty}}{\| I - B \|_{e, \infty} \| u^n \|_{e, \infty}} \leq \frac{\| u^* - u^n \|_{e, \infty}}{\| u^n \|_{e, \infty} - \| u^* - u^n \|_{e, \infty}},
\]
which can be written finally as follows
\[
\frac{\| (I - B) u^n - c \|_{e, \infty}}{\| I - B \|_{e, \infty} \| u^n \|_{e, \infty}} \leq \frac{\| u^* - u^n \|_{e, \infty}}{\frac{1}{\| u^n \|_{e, \infty}} - \frac{1}{\| u^* - u^n \|_{e, \infty}}}.\]

Thus, if the condition (3.13) is satisfied and if furthermore
\[
k(\lambda_e) \left( \frac{\tau}{1 - \alpha} + \eta \right) < 1,
\]
then the corresponding backward error is bounded by
\[
\frac{\| (I - B) u^n - c \|_{e, \infty}}{\| I - B \|_{e, \infty} \| u^n \|_{e, \infty}} \leq k(\lambda_e) \left( \frac{\tau}{1 - \alpha} + \eta \right) \frac{1}{1 - k(\lambda_e) \left( \frac{\tau}{1 - \alpha} + \eta \right)}, \forall n \geq n_k^e.
\]

**Remark 6** The previous estimate of the backward error is an extension of the classical one (see [16]) to the parallel context.

**Remark 7** The results presented in this section can be extended to the case of periodic asynchronous iterations with multiple initializations (see [9], [7] and [22]) that can also be regarded as successive approximation methods on a product space (see [21]).

**Acknowledgements** The authors wish to thank the referees and the editor for their helpful comments and remarks.

**References**


