RAIRO
MODÉLISATION MATHÉMATIQUE
ET ANALYSE NUMÉRIQUE

JEAN-PAUL CHEHAB

A nonlinear adaptative multiresolution method in finite differences with incremental unknowns

RAIRO – Modélisation mathématique et analyse numérique, tome 29, n° 4 (1995), p. 451-475.

<http://www.numdam.org/item?id=M2AN_1995__29_4_451_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE
(Vol. 29, n° 4, 1995, p. 451 à 475)

A NONLINEAR ADAPTATIVE MULTIRESOLUTION METHOD IN FINITE DIFFERENCES WITH INCREMENTAL UNKNOWNS (*)

by Jean-Paul CHEHAB (1)

Communicated by R. TEMAM

Abstract. — *In this article, we propose a new method well suited for the calculation of unstable solutions of nonlinear eigenvalues problem. This method is derived from the classical Marder-Weitzner scheme (MW) which can be seen as a nonlinear Richardson method. First we adapt to (MW) the usual extension of the classical Linear Richardson scheme (LR) which consists in computing the relaxation parameter in order to minimize the iterative residual in a suitable norm. This method is then generalized with the utilization of the Incremental Unknowns (I.U.) inducing the minimizing relaxation parameter in the embedded hierarchical subspaces. We obtain in this way both generalizations of the MW and the LR algorithms. The numerical illustrations we give allowing comparisons between the differents LR schemes (for linear problems) and some versions of the MW method (for nonlinear eigenvalue problems), point out the better speed of convergence of the new algorithms.*

1. INTRODUCTION

The Incremental Unknowns, introduced in [11], is a multiresolution method well adapted to the solution of nonlinear problems when finite differences are used. It is related to the nonlinear Galerkin method [8] [9] and it can be seen as the analog of the Hierarchical Basis Finite Elements Method for finite differences. Using several levels of discretization, the I.U. method generates different structures or scales in distinct points of a grid. The difference of magnitude of the several unknowns leads us to treat them differently, according to the grid level associated, in a given scheme. This idea introduced by the Nonlinear Galerkin Method was applied in [2] for the solution of nonlinear eigenvalue problems and gave efficient generalizations of The Marder-Weitzner Method (MW). These new methods were based on substituing the scalar relaxation step of (MW) by a matricial multirelaxation

(*) Manuscript received January, 10, 1994, revised May 25, 1994.

(1) Université de Paris-Sud, Laboratoire d'Analyse Numérique, Bâtiment 425, 91405 Orsay Cedex, France.

step : each grid level unknown was relaxed by an appropriate parameter. The drawback of the scheme proposed in [2] was in the fact that one must choose, at the begining of a program, the relaxation matrix and keep it fixed, while it would be more efficient to change it along the iterations (the MW scheme is a local method).

This article is a continuation of [2]. Our aim here is to transform the MW scheme in order to calculate automatically, along the iterations, the relaxation matrix adaptating it to the current approximation. For that purpose, we point out a very simple analogy between MW and the classical (linear) Richardson method (LR) comparing the propagation error equations. We adapt to MW the usual extension of LR consisting of replacing the relaxation parameter by a current relaxation parameter which minimizes at each step the error, then the residual, in a suitable norm. MW is the modified by the corresponding adaptative calulus of the relaxation parameter.

This paper is organized as follows : first, and after a brief presentation of the basic methods, we establish an analogy between MW and LR *via* their respective propagation error equations. We then show that the MW method is nothing but a nonlinear Richardson method. After that, considering the usual extension of LR (LRA), we define, proceeding again by analogy, a new generalization of MW where the scalar relaxation parameter is, as in LRA, calculated along the iterations such as to minimize the current residual. In section 4 we consider a multilevel discretization using the I.U. method and we extend the « minimizing » relaxation step to each grid level. The relaxation step is then completed with a diagonal matrix in the I.U. basis. This extension gives generalizations of both LRA and MW. Finally, in section 5, we present some numerical results. They concern two type of problems :

- a linear problem where, comparing LRA and its multilevel extension we point out a better speed of convergence ;
- a nonlinear eigenvalue problem where, comparing our new MW scheme with the classical one and those introduced in [2], we observe a much better speed of convergence for a comparable CPU computing time.

2. THE MARDER-WEITZNER AND THE RICHARDSON METHODS

2.1. The Richardson scheme

Let us consider the linear problem :

$$\begin{cases} \text{Find } X \in IR^n \text{ such that } M \,.\, X = b \\ \text{where } M \text{ is an } n \times n \text{ positive definite matrix .} \end{cases} \tag{1}$$

One of the simplest methods to solve (1) is the Richardson scheme which is defined as follows :

$$\begin{cases} \text{Let } X^0 \text{ be an initial guess} \\ \text{For } k = 0, \ldots \\ X^{k+1} = X^k + \alpha (b - M \,.\, X^k) \\ \text{where } \alpha \text{ is a nonnegative real parameter .} \end{cases} \tag{2}$$

Let us examine the propagation error equation of (1). We set $\varepsilon^k = X^k - X$, the error at the k^{th} step. We have

$$\varepsilon^{k+1} = (I - \alpha M)\, \varepsilon^k \tag{3}$$

where I is the $n \times n$ identity matrix. It is well known that a necessary and sufficient condition of convergence of the scheme (1) is

$$\alpha < \frac{2}{\rho (M)}$$

and the optimal relaxation parameter is given by

$$\alpha_{\text{opt}} = \frac{2}{\rho (M) + \mu (M)}$$

where $\rho (M)$ (resp $\mu (M)$) is the larger (resp. the smaller) eigenvalue of the matrix M.

2.2. The Marder-Weitzner scheme

Let us consider now the following nonlinear problem :

$$\begin{cases} \text{Find } X \in IR^n \text{ such that } X = T(X) \\ \text{where } \quad T : IR^n \to IR^n \text{ is a nonlinear mapping .} \end{cases} \tag{4}$$

We recall that the Marder-Weitzner Method is a fixed point method consisting of a three steps scheme generalizing the Picard Iterates (PI) ; it is well suited for the calculation of unstable solutions (unlike PI). In particular, we can apply MW to the calculation of a solution after a bifurcation (see [7] and [10]). We define it as follows.

Let X^0 be the initial guess, assumed to be close enough to X, a local solution of (4). The sequence X^k is defined by :

$$\begin{cases} X^{k+1/3} = T(X^k) \\ X^{k+2/3} = T(X^{k+1/3}) \\ X^{k+} \;\; = X^k + \alpha (2\, X^{k+1/3} - X^k - X^{k+2/3}) , \end{cases} \tag{5}$$

where α is again a nonnegative real parameter.

Now, as for the Richardson method, we analyse the propagation error equation of MW. Denoting by Ψ the jacobian matrix of T at X we have :

$$\varepsilon^{k+1} = (I - \alpha (I - \Psi)^2)\, \varepsilon^k + o(\varepsilon^k)\,. \tag{6}$$

Hence, setting $M = (I - \Psi)^2$, we see that the linear (dominant) part of (6) is nothing but (3). Furthermore we recover here the same conditions on the relaxation parameters. They are given by the following Theorem due to M. Sermange [10].

THEOREM 1 : *Assume that T is Fréchet differentiable at the solutions of (4). We denote by Ψ the Fréchet differential. Assume that the eigenvalues of Ψ are reals and different of 1 at X a solution X of (4).*

Then for $0 < \alpha < \alpha_c$, there exists a neighborood V_α of X such that if $X^0 \in V_\alpha$ and X^k is defined by (5),

$$X^{k+i/3} \to X \quad \textit{when} \quad k \to \infty \quad i = 0, 1, 2\,.$$

Moreover if we write $a = \sup_{y \in SP(\Psi)} |1 - y|$ and $b = \inf_{y \in SP(\Psi)} |1 - y|$, we have the following values of the critical and the optimal relaxation parameter :

$$\alpha_c = \frac{2}{a^2} \quad \textit{and} \quad \alpha_{\text{opt}} = \frac{2}{a^2 + b^2}\,.$$

The analogy of the two methods presented above shows that MW is a **nonlinear Richardson Method.** It is then natural to try to adapt to MW the usual extensions of LR.

3. THE MINIMIZING RESIDUAL RELAXATION PARAMETER

3.1. The Richardson method case

A classical extension of (1) consists of replacing a fixed relaxation parameter α by a variable parameter which depends of the current iterate and which minimizes the residual in a suitable norm. The scheme is :

$$X^{k+1} = X^k + \alpha_k (b - M \,.\, X^k)\,.$$

Setting $r^k = b - M \,.\, X^k$, the residual at the k^{th} step and $\| \,.\, \|$ the euclidian norm, we easily find that

$$r^{k+1} = r^k - \alpha_k M \,.\, r^k$$

and that the α_k which minimizes $\|r_{k+1}\|$ is given by

$$\alpha_k = \frac{(M \cdot r^k, r^k)}{(M \cdot r^k, M \cdot r^k)}$$

where $(\cdot, \cdot)$ is the euclidian scalar product.

Now, we shall adapt this extension to the MW scheme.

3.2. The Marder-Weitzner method case

We consider again the problem (4) and we introduce the following notations :

$$\begin{aligned} M &= (I - \Psi)^2 \\ N &= I - \Psi \\ r^k &= X^k - T(X^k) \end{aligned}$$

where X^k is defined by the MW scheme (5) and we denote again by X a local solution of (4).

First, we relate the error to the residual. By the definition of ε^k, we have

$$\begin{aligned} r^k &= X + \varepsilon^k - T(X + \varepsilon^k) \\ &= N \cdot \varepsilon^k + o(\varepsilon^k) \end{aligned}$$

because $X = T(X)$. But

$$r^k = X^k - X^{k+1/3} .$$

Then multiplying on the left by N each side of (6), we find

$$r^{k+1} = (I - \alpha_k \cdot M)\, r^k + o(r^k)$$

and the parameter α_k which minimizes $\|r^{k+1}\|$ is given by

$$\alpha_k = \frac{(M \cdot r^k, r^k)}{(M \cdot r^k, M \cdot r^k)} .$$

Remark 1 : It is not precisely a minimization of the residual : we minimize only the linearized part of the iteration operator of *MW*.

For the determination of α_k we need to known, at each iteration of MW, the vectors $r^k = X^k - X^{k+1/3}$, which can be calculated along the iterations of MW and $M \cdot r^k$ which can not be deduced of the sequence $X^{k+i/3}$. For this reason we introduce a supplementary Picard iterate. We set

$$V = T(X^{k+2/3}) .$$

The vector we want to deduce is

$$M \cdot r^k = (I - \Psi)^2 r^k = (I - 2\,\Psi + \Psi^2) r^k .$$

Hence, we need to evaluate Ψr^k and $\Psi^2 r^k$.

Evaluation of $\Psi^2 r^k$

We have

$$V = T(X + \Psi^2 \varepsilon^k + o(\varepsilon^k))$$

and then

$$V = X + \Psi^3 \varepsilon^k + o(\varepsilon^k) .$$

Thus

$$V - X^{k+2/3} = \Psi^3 \varepsilon^k - \Psi^2 \varepsilon^k + o(\varepsilon^k) .$$

Consequently

$$V - X^{k+2/3} = - \Psi^2 (I - \Psi)\, \varepsilon^k + o(\varepsilon^k) ,$$

that is to say

$$V - X^{k+2/3} = - \Psi^2 r^k + o(r^k).$$

Hence

$$\Psi^2 r^k = X^{k+2/3} - V + o(r^k) .$$

Now we conclude with the

Evaluation of Ψr^k

We have

$$X^{k+1/3} - X^{k+2/3} = T(X^k) - T(X^{k+1/3}) .$$

Thus

$$X^{k+1/3} - X^{k+2/3} = T(X + \varepsilon^k) - T(X + \Psi \varepsilon^k + o(\varepsilon^k)) .$$

Hence,

$$X^{k+1/3} - X^{k+2/3} = (\Psi - \Psi^2)\, \varepsilon^k + o(\varepsilon^k) = \Psi r^k + o(r^k) .$$

Then

$$\Psi r^k = X^{k+1/3} - X^{k+2/3} + o(r^k) ,$$

and finally,

$$\begin{aligned} M \cdot r^k &= X^{k+2/3} - V + 2(X^{k+2/3} - X^{k+1/3}) + X^k - X^{k+1/3} + o(r^k) \\ &\simeq 3(X^{k+2/3} - X^{k+1/3}) + X^k - V . \end{aligned}$$

Setting $W = 3(X^{k+2/3} - X^{k+1/3}) + X^k - V$, we find that

$$\alpha_k = \frac{(W, r^k)}{(W, W)}.$$

We note that the matrix $M = (I - \Psi)^2$ is never calculated and that the matrix-vector products are estimated only using combinations of successive Picard iterates.

This relaxation technics leads to a new MW scheme, called A1, which one can define as follows :

The A1 method

Let X^0 be the initial guess supposed sufficiently close to X, a local solution of (4). For $k = 0, \ldots$

$$\left\{\begin{array}{l}
\textbf{step 1:} \text{ solve} \\
X^{k+1/3} = T(X^k) \\
\textbf{step 2:} \text{ solve} \\
X^{k+2/3} = T(X^{k+1/3}) \\
\text{Computation of the supplementary Picard iterate} \\
\textbf{solve} \\
V = T(X^{k+2/3}) \\
\text{Compute the vectors :} \\
r^k = X^k - X^{k+1/3} \\
W = 3(X^{k+2/3} - X^{k+1/3}) + X^k - V \\
\text{Compute the relaxation parameter} \\
\alpha_k = \dfrac{(W, r^k)}{(W, W)} \\
\text{Relax the Picard iterates to compute } X^{k+1} \\
X^{k+1} = X^k + \alpha_k(2X^{k+1/3} - X^k - X^{k+2/3}).
\end{array}\right. \tag{7}$$

4. THE MULTILEVEL RELAXATION

In the previous sections we have presented a method which determines automatically (and independentely of the program's datas) the relaxation

parameter in the MW scheme. This method does not take into account the discretization technics used. As we said in the introduction, when we discretize the problem with the Incremental Unknowns, the unknown solution vector to compute has a multilevel structure : its components are not of the same order of magnitude according to the grid level they are associated with (see [2]). In this section, we construct an adapted scheme derived from the generalization of the MW scheme presented above. First of all let us recall briefly the definition and the main properties of the Incremental Unknowns.

4.1. The incremental unknowns. Definition and properties

The construction of the Incremental Unknowns is composed of two steps. For the sake of simplicity we consider first two levels of discretization for the Incremental Unknowns.

4.1.1. *Hierarchization*

The first step consists in a hierarchization of the components as a function of the grid level they are associated with. Like the Multigrid Method one considers a regular meshing of an open set Ω associated to the spatial mesh size h. At this point we distinguish the coarse grid G_H associated to the mesh $H = 2 \cdot h$, and the fine grid G_h associated to the mesh h. The hierarchization consists in arranging in a vector (which represents for instance the approximation of a function at the grid points) first the components lying in G_H and after that those lying on $G_h \backslash G_H$ with the standard lexicographic order in each family of components.

$$\cdot\ o \times o \times o \times o \times o\ \cdot$$

Figure . — Dimension 1, $\Omega =]0, 1[$. $\times$: G_H points, o : $G_h \backslash G_H$ points.

$$\begin{array}{ccccccccc}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & o & o & o & o & o & o & o & \cdot \\
\cdot & o & \times & o & \times & o & \times & o & \cdot \\
\cdot & o & o & o & o & o & o & o & \cdot \\
\cdot & o & \times & o & \times & o & \times & o & \cdot \\
\cdot & o & o & o & o & o & o & o & \\
\cdot & o & \times & o & \times & o & \times & o & \cdot \\
\cdot & o & o & o & o & o & o & o & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}$$

Figure 2. — Dimension 2, $\Omega =]0, 1[^2$. $\times$: G_H points, o : $G_h \backslash G_H$ points.

A hierarchized vector U is written $U = (U_c, U_f)^t$ with $U_c = Y \in G_H$ and $U_f \in G_h \backslash G_H$.

4.1.2. *Change of variable*

The second step of the construction of the Incremental Unknowns consists in a change of variable which operates only in $G_h \backslash G_H$. We can express it in the form :

$$Z_f = U_f - RY \tag{8}$$

where $R : G_H \to G_h \backslash G_H$ is a second order interpolation operator and then, according to Taylor's formula, the unknowns of $G_h \backslash G_H$ are of order $O(h^2)$. The numbers Z are the *incremental unknowns*. We can, of course, repeat recursively the process described above, using l levels of discretization defining then l Z-levels.

We introduce now the following notations :

We shall say that a grid has a $C_{k,l}$ configuration if it is obtained with l dyadic refinements of a grid composed of k points in each direction of the domain. The fine grid is thus composed of $2^l(k+1)-1$ points in each direction. Then denoting by S the transfer matrix, we have

$$\begin{pmatrix} Y \\ U_{f_1} \\ U_{f_2} \\ \vdots \\ U_{f_l} \end{pmatrix} = S \begin{pmatrix} Y \\ Z_1 \\ Z_2 \\ \vdots \\ Z_l \end{pmatrix}$$

with obvious notations.

4.2. The multilevel Richardson methods

We assume that the fine grid is decomposable into a $C_{k,l}$ grid and that we have discretized the problem with $l+1$ nested grids. Let us consider the following bloc decomposition of the matrix M in function of the hierarchical decomposition of the approximating space $\mathbf{V}$:

$$M = \begin{pmatrix} M_{0,0} & M_{0,1} & \cdots & M_{0,l} \\ M_{1,0} & M_{1,1} & \cdots & M_{1,l} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ M_{l,0} & \cdot & M_{l,l-1} & M_{l,l} \end{pmatrix}. \tag{9}$$

We consider the splitting of $\mathbf{V} = V_0 \overset{l}{\underset{j=1}{\oplus}} W_j$ where V_0 is the subspace associated to the coarsest grid and the W_j are the subspaces associated to the successive complementary grids. Hence, the subspace corresponding to j hierarchised grids, $j < l$, is $V_j = V_0 \overset{l}{\underset{m=1}{\oplus}} W_m$. The question now is how to approximate the minimizing residual relaxation parameter, introduced in section 3 and defined on the finest grid, on the subspace V_j. We recall that this parameter is defined on the fine grid by

$$\alpha_k = \frac{(M \,.\, r^k,\, r^k)}{(M \,.\, r^k,\, M \,.\, r^k)}$$

where r^k and $M \,.\, r^k$ are written in the $\mathbf{V}$ basis as

$$r^k = (r_0^k,\, r_1^k,\, r_2^k,\, \ldots,\, r_l^k)^t \quad \text{and} \quad M \,.\, r^k =$$

$$= \left(\sum_{j=0}^{l} M_{0j}\, r_j^k,\ \sum_{j=0}^{l} M_{1j}\, r_j^k,\ \sum_{j=0}^{l} M_{2j}\, r_j^k,\ \ldots,\ \sum_{j=0}^{l} M_{lj}\, r_j^k \right)^t .$$

A natural way to adapt this parameter to the subspace V_j is to consider in the previous formula the projection of the vectors on this subspace. We denote by $P_j(r^k)$ (resp. $P_j(M \,.\, r^k)$) the projection of r^k (resp. of $M \,.\, r^k$) on V_j, and we define the relaxation parameter on V_j by

$$\alpha_j^k = \left| \frac{(P_j(M \,.\, r^k),\, P_j(r^k))}{(P_j(M \,.\, r^k),\, P_j(M \,.\, r^k))} \right| , \tag{10}$$

where $(\,.\,,\,.\,)$ is still the euclidian scalar product on $\mathbf{V}$. We have taken the modulus of the expression in order to obtain a positive relaxation parameter.

We conclude by pointing out that $Z_j \in W_j \subset V_j$ and then the relaxation on W_j can be realized with α_j^k. Moreover it is clear that the (multi)relaxation method described above is applicable on all kinds of Richardson methods and then in particular on both the linear (L.R.) and nonlinear (MW) ones. In the following we set for convenience $P_j(r^k) = r_j^k$ and $P_j(M \,.\, r^k) = (M \,.\, r^k)_j$.

4.2.1. *A multivel linear Richardson method*

We consider the linear problem :

$$\begin{cases} \text{Find } X \in IR^n \text{ such that } A \,.\, X = b \\ \text{where } A \text{ is an } n \times n \text{ positive definite matrix .} \end{cases} \tag{11}$$

We assume that the grid is a $C_{k,l}$ one and we hierarchize (11). The unknown solution vector of the problem has the structure

$$\begin{pmatrix} Y \\ U_{f_1} \\ U_{f_2} \\ \vdots \\ U_{f_l} \end{pmatrix}.$$

Introducing the Incremental Unknowns by the transfer matrix S we obtain a multistructural solution vector related to the previous one by

$$\begin{pmatrix} Y \\ U_{f_1} \\ U_{f_2} \\ \vdots \\ U_{f_l} \end{pmatrix} = S \begin{pmatrix} Y \\ Z_1 \\ Z_2 \\ \vdots \\ Z_l \end{pmatrix}$$

with obvious notations.

The linear problem to solve, which is equivalent to (11), is

$$\text{Find } \hat{X} \in IR^n \text{ such that } A \,.\, S\hat{X} = b \,. \tag{12}$$

Multiplying on the left every term of (12) by ${}^tS$, we obtain the equivalent problem

$$\text{Find } \hat{X} \in IR^n \text{ such that } M\hat{X} = {}^tSb = \hat{b} \,, \tag{13}$$

where $M = {}^tSAS$.

Using the formula (10) we can define the following generalization of the linear Richardson Method with the minimizing residual relaxation parameter.

Algorithm MLR

$$\left\{ \begin{array}{l} \text{Let } \hat{X}^0 \text{ be an initial guess. For } k = 0, \ldots \\ \text{Compute } W_k = M \,.\, \hat{X}^k \\ \text{Compute } r^k = \hat{b} - W_k \\ \text{For } j = 0, \ldots, l \\ \alpha_j^k = \left| \dfrac{((M \,.\, r^k)_j, \, r_j^k)}{((M \,.\, r^k)_j, \, (M \,.\, r^k)_l)} \right| \\ Z_j^{k+1} = Z_j^k - \alpha_j^k r_j^k \,, \end{array} \right. \tag{14}$$

where we have set for convenience $Z_0 = Y$.

Remark 2 : We have multiplied on the left each term of (12) by ${}^tS$ because when the matrix A is symmetric and positive definite, e.g. when it represents the discretization of an elliptic self adjoint operator, one obtains again a symmetric definite posite matrix ${}^tSAS$. This is usefull in the elliptic case : indeed, $M = {}^tSAS$ has a condition number much smaller than A and, it is well known, both the gradient and the Richardson methods have their speed of convergence increased when the condition number of the matrix is decreased.

4.2.2. *A Multilevel nonlinear Richardson Method. A new generalization of the MW scheme*

Considering the MW method as a nonlinear Richardson method and using the same technic of multirelaxation, we can define the following MW type scheme.

Let us consider the discrete nonlinear eigenvalue problem

$$\text{Find } X \in IR^n \text{ such that } AX = \gamma F(X), \tag{15}$$

where A is the discretization matrix of $-\Delta$, written in the hierarchical basis, $F : IR^n \to IR^n$ is a C^2 function such that the hypothesis on $T(\,.\,) = \gamma A^{-1} F(\,.\,)$ are those of Theorem 1, and γ is a nonnegative real parameter. We assume that the mesh is decomposable into a $C_{k,l}$ grid.

Now as for the linear problem, we introduce the incremental unknowns with the variable change S and we let $X = S\,.\,\hat{X}$. Hence (15) is equivalent to

$$\text{Find } \hat{X} \in IR^n \text{ such that } AS\hat{X} = \gamma F(S\hat{X}). \tag{16}$$

Here $\hat{X}$ is the $(l+1)$-level vector $\hat{X} = (Y, Z_1, Z_2, \ldots, Z_l)^t$ built on $l+1$ nested meshes.

Multiplying on the left every term of (16) by ${}^tS$ so that the linear operator is positive definite, we obtain :

$$\text{Find } \hat{X} \in IR^n \text{ such that } {}^tSAS\hat{X} = \gamma\, {}^tSF(S\hat{X}). \tag{17}$$

We set $\hat{A} = {}^tSAS$.

Now we can define the algorithm MWIUa (MWIU adaptative) for which the relaxation matrix is determined as in the MLR method.

Let $\hat{X}^0$ be the initial guess supposed sufficiently close to $\hat{X}$, a local solution of (17).

The MWIUa Algorithm For $k = 0, \ldots$

$$
\left\{
\begin{array}{l}
\textbf{step 1 : } solve \\
\hat{A}\hat{X}^{k+1/3} = \gamma^t SF\,(S\hat{X}^k) \\
\textbf{step 2 : } solve \\
\hat{A}\hat{X}^{k+2/3} = \gamma^t SF\,(S\hat{X}^{k+1/3}) \\
\text{Computation of the supplementary Picard iterate} \\
\textbf{solve} \\
\hat{A}V = \gamma^t SF\,(S\hat{X}^{k+2/3}) \\
\text{Compute the vectors:} \\
r^k = \hat{X}^k - \hat{X}^{k+1/3} \text{ and} \\
W = 3\,(\hat{X}^{k+2/3} - \hat{X}^{k+1/3}) + \hat{X}^k - V \\
\text{Multirelaxation} \\
\text{For } j = 0, \ldots, l \\
\alpha_j^k = \left| \dfrac{((W)_j,\ r_j^k)}{((W)_j,\ (W)_j)} \right| \\
Z_j^{k+1} = Z_j^k + \alpha_j^k (2\,Z_j^{k+1/3} - Z_j^{k+2/3} - Z_j^k)\,.
\end{array}
\right. \tag{18}
$$

Remark 3 : The CPU Computing time per iteration for MWIUa is the same as for A1IU (the A1 scheme written with IU). Modulo l supplementary divisions, the determination of the relaxation parameter(s) involves the same operation.

5. NUMERICAL RESULTS

In this section, we illustrate the efficiency of the adaptative multirelaxation. We give some numerical results allowing comparisons between the several versions of both the linear and the nonlinear Richardson methods. For the linear problems, we compare the adaptative Richardson methods in the classical case and in the multilevel case. For the nonlinear problems, we compare the classical MW algorithm and its generalizations introduced in [2], using I.U., with the new adaptative nonlinear Richardson algorithms built in the previous section.

5.1. The linear case : solution of the Dirichlet problem

We consider the classical Dirichlet problem

$$
\left\{
\begin{array}{rl}
-\Delta u = f & \text{in } \Omega =]0, 1[^2 \\
u = 0 & \text{on } \partial\Omega\,,
\end{array}
\right. \tag{19}
$$

which is discretized by the usual five points scheme in finite differences on a $C_{k,l}$ grid. We arrange the unknowns in the hierarchical order and we introduce the Incremental Unknowns *via* the transfer matrix S. The discrete problem to solve is then

$$AS \,.\, \hat{X} = F \tag{20}$$

where A is the discretization matrix of $-\Delta$ written in the hierarchical basis. After the symmetrization of (20) by multiplication on the left of each term by ${}^tS$, we obtain the symmetric system

$${}^tSAS \,.\, \hat{X} = {}^tSF = \hat{F} \,. \tag{21}$$

M. Chen and R. Temam have shown in [4] that the condition number of $\hat{A}$ is $C_1 \,.\, (l+1)^2 = C_1 \,.\, \left(\mathrm{Log}_2 \left(\frac{1}{h}\right)\right)^2$ which is much smaller than $K(A) = \dfrac{C_2}{h^2}$ where C_1 and C_2 are positive numbers independant of the mesh size. This property points out an obvious advantage of solving the discrete problem under the form (21) rather than in the usual nodal basis, in particular when one uses a conjugate gradient method (see [3]). It is well known that the speed of convergence of this method is related to $\sqrt{K(A)}$. One recover this improved speed of convergence with the LRA scheme. We have indeed (see [5])

$$\frac{(r^{k+1}, A^{-1} r^{k+1})}{(r^k, A^{-1} r^k)} \leqslant \left(\frac{K(A)-1}{K(A)+1}\right)^2 .$$

For this reason we did not compare the Richardson methods involving Incrementals Unknowns with the corresponding schemes in the nodal basis where the convergence is very slow.

The numerical results we present here correspond to $F = 0$: there is no loss of generality in taking a null source term.

We have chosen as initial data $\hat{X}^0 = \sin\,(16 \,.\, x \,.\, y \,.\, (1-x) \,.\, (1-y))$.

In the figures (3) to (5), one can compare the evolution along the iterations of the relative error $\left(\dfrac{\|X^k - X\|}{\|X^0\|}\right)$ and the residual for the LRA and the MLR schemes. As one can see, the number of iterations is reduced by about 30 % for the MLR method as compared to the LRA scheme. This proportion seems to be independent of the fine grid mesh size. The new linear Richardson Method can not be considered as a powerful elliptic solver but it gives a natural illustration of the efficiency of the multirelaxation technics proposed in section 4. We think that it is a useful step for the extension of the technics in the nonlinear case.

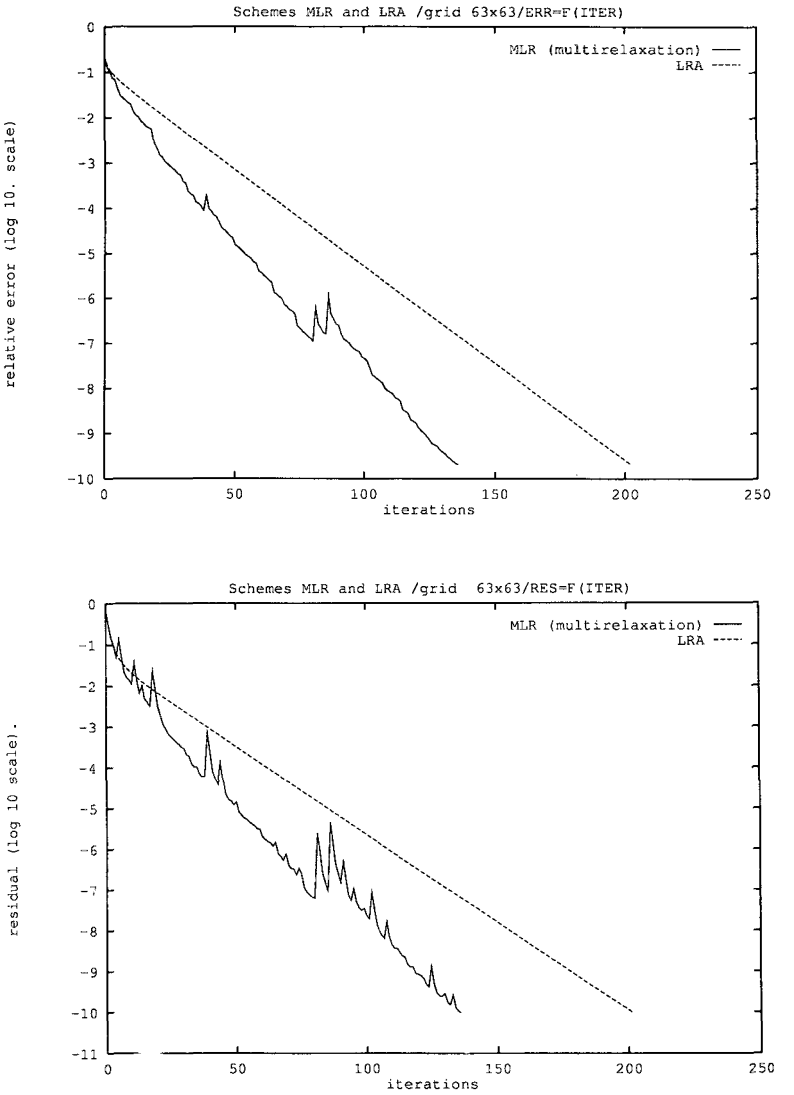


Figure 3. — Comparison between the LRA and MLR Methods. The relative error and the residue are plotted against iterations. The grid is $C_{1.5}$.

5.2. The nonlinear case : solution of a nonlinear eigenvalue problem

5.2.1. *The model problem*

We consider the following problem : we want to calculate some unstable

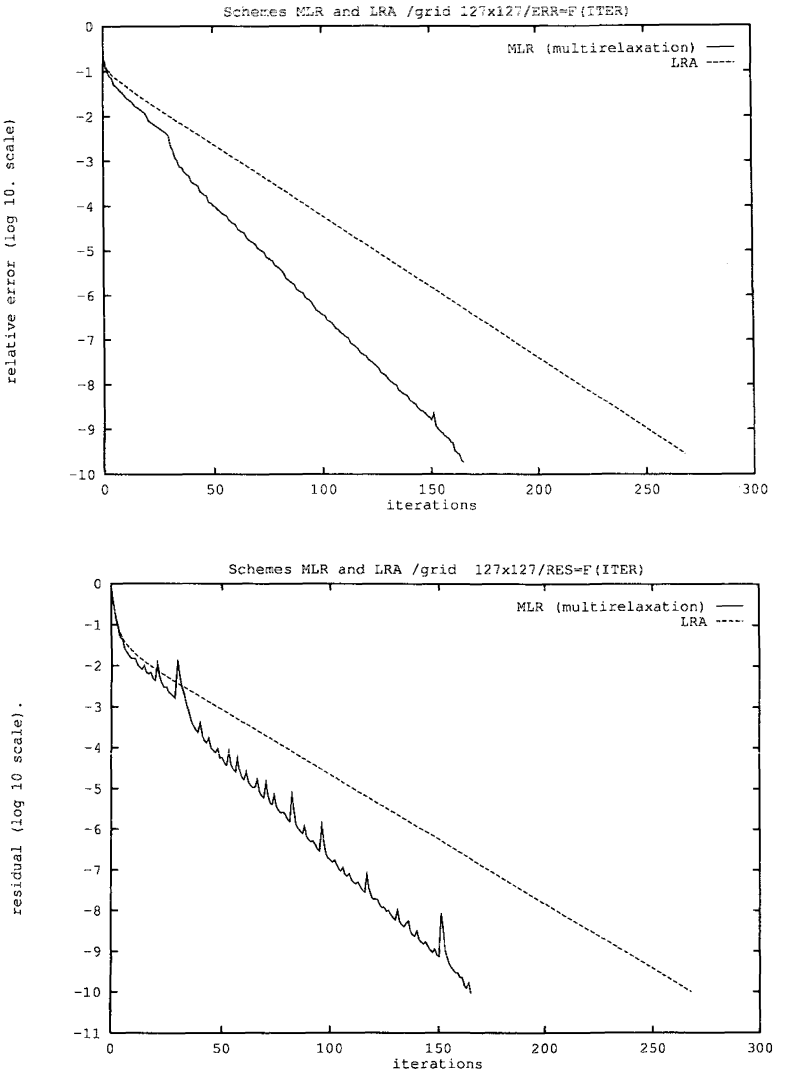


Figure 4. — Comparison between the LRA and MLR Methods. The relative error and the residue are plotted against iterations. The grid is $C_{1.6}$.

solutions of

$$\begin{cases} -\Delta u = \gamma u - \nu |u|^{\varepsilon} u & \text{in} \quad \Omega =]0, 1[^2 \\ u = 0 & \text{on} \quad \partial\Omega , \end{cases} \tag{22}$$

with γ and $\nu > 0$ and $0 < \varepsilon \leqslant 2$.

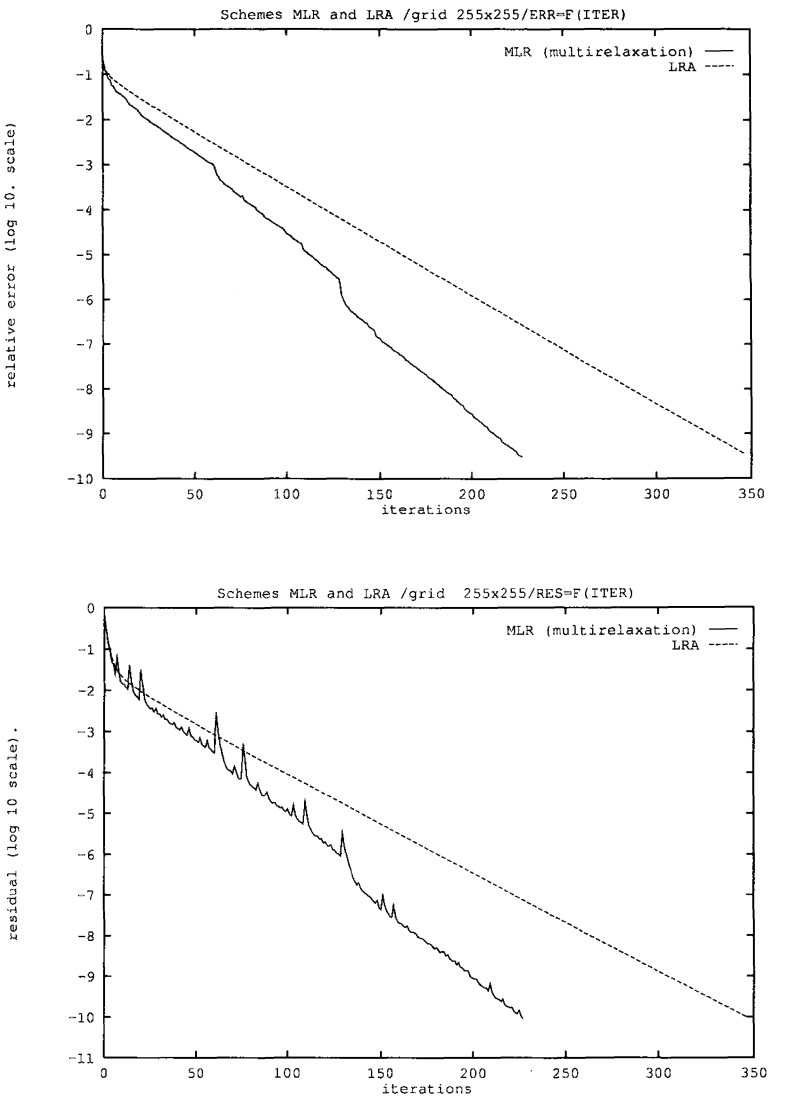


Figure 5. — Comparison between the LRA and MLR Methods. The relative error and the residue are plotted against iterations. The grid is $C_{1.7}$.

It is well known that such a problem exhibits bifurcations everytime the parameter γ crosses an eigenvalue of $-\Delta$. We easily verify that the hypothesis of theorem 1 are satisfied, and that, consequently, the MW scheme is well suited to compute unstable solutions of (22).

5.2.2. *Properties of the solutions and choices of the initial datas*

Let $\Lambda_{p,q} = \pi^2(p^2 + q^2)$, $p \cdot q \neq 0$, be an eigenvalue of $-\Delta$ and let $\Phi_{p,q} = \sin(p\pi x)\sin(q\pi y)$ be the corresponding eigenfunction. We recall the following results (see [1] and [6] for the dimension one) :

— When $\gamma < \Lambda_{1,1}$, the trivial solution, $u \equiv 0$, is the onlv one and it is stable.

— When $\Lambda_{1,1} < \gamma \leq \Lambda_{1,2}$, the trivial solution becomes unstable and there exist two stable solutions, denoted by $K(1, 1)$ which are deformations of the eigenfunction $\Phi_{1,1}$ (i.e. the set of zeros and the extremas are at the same points).

— When $\Lambda_{p,q} < \gamma$, $p^2 + q^2 > 2$, all the solutions (including the trivial one) are unstables except the $K(1, 1)$ type ones. Let a and b be such that $a^2 + b^2 \leq p^2 + q^2$. These unstable solutions are deformation of basically two kinds of functions :

- The eigenfunctions of $-\Delta$, the $\Phi_{a,b}$. The corresponding unstable solutions are of $K(a, b)$ type. Then, to compute them, we take $U_0 = k \cdot \Phi_{a,b}$ as the initial guess.
- The functions $\Theta_{a,b} = \sin(a\pi x)\sin(b\pi y)\,Z(x, y)$, where $Z(x, y)$ vanishes on a segment parallel to the lines of equation $y = x$ and $y = -x$. These solutions are said to be of $\Delta(a, b)$ type. To compute them, we take $U_0 = \Theta_{a,b}$ as the initial guess.

5.2.3. *The MW schemes used*

We first discretize the model problem by finite differences and using the I. U., we obtain the discrete non linear problem :

$$\hat{A}\hat{X} = {}^tSF(S\hat{X}) = {}^tS(\gamma S\hat{X} - \nu \left|S\hat{X}\right|^{\varepsilon} S\hat{X}). \tag{23}$$

Now, we introduce the following MW schemes which we shall compare with those constructed in the previous sections.

First we recall the definition of the MWIU method introduced in [2] and which may be seen as the multilevel version of the classical MW algorithm : the relaxation parameters are fixed once for all at the begining of the program.

The MWIU method

Let $\hat{U}^0$ be the initial guess supposed sufficiently close to $\hat{X}$, a local solution of (23). For $k = 0, \ldots$

$$\left\{\begin{array}{l}\textbf{step 1 : solve}\\ \hat{A}\hat{X}^{k+1/3} = \gamma^{l}\, SF\,(S\hat{X}^{k})\\ \textbf{step 2 : solve}\\ \hat{A}\hat{X}^{k+2/3} = \gamma^{l}\, SF\,(S\hat{X}^{k+1/3})\\ \textbf{step 3 : multirelaxation}\\ \hat{X}^{k+1} = \hat{X}^{k} + \Gamma(2\,\hat{X}^{k+1/3} - \hat{X}^{k+2/3} - \hat{X}^{k})\\ \text{Or equivalentely}\\ \text{For } j = 0, \ldots, l\\ Z_j^{k+1} = Z_j^{k} + \alpha_j(2\,Z_j^{k+1/3} - Z_j^{k+2/3} - Z_j^{k}),\end{array}\right. \tag{24}$$

where $\Gamma = \text{DIAG}\ (\alpha_0, \alpha_1, \ldots, \alpha_l)$ is the diagonal multirelaxation matrix. In [2] we have proposed a method for the construction of this matrix pointing out, with the help of numerical observations, that the α_j must be chosen as an increasing sequence in j with the first grid levels parameter closes to the relaxation parameter used in the MW for the same problem and the same datas of course.

The stoping test

The several MW schemes considered here are iterative methods. It is then necessary to define a numerical criteria which indicates that the current iterate is accurate enough. We then introduce two residuals :

- The iterative residual at the $(k+1)^{th}$ step :

$$\omega^{k+1} = \frac{\|X^{k+1} - X^{k}\|}{\|X^{k}\|}\,.$$

It measures the relative progression of the process at the $(k+1)^{th}$ step. Looking at the relaxation step of MW, we have

$$X^{k+1} - X^{k} = \alpha\,(2\,.\,X^{k+1/3} - X^{k+2/3} - X^{k})\,,$$

and then $\alpha\,(2\,.\,X^{k+1/3} - X^{k+2/3} - X^{k})$ is the correction added to X^{k} to compute X^{k+1}. Taking the relative value of this correction, it is reasonable to consider that the algorithm has converged when ω^{k+1} is small. Moreover it indicates that the solution computed is nontrivial. To be sure that this is a solution of the problem, we use also

- The classical residual defined by

$$r^{k} = \|A\,.\,X^{k} - F\,(X^{k})\|\,.$$

In general, the first residue suffices to measure the accuracy of the process. We then choose a small real parameter η and the iterations will stop when $\omega^{k+1} \leqslant \eta$.

Calculation of a K(2,2) solution

Values of the relaxation parameters

For the classical MW scheme $\alpha = 0.117$.
The matrix Γ is

$$\alpha_0 = 0.117 \qquad \alpha_1 = 0.7099 \qquad \alpha_2 = 0.99899$$
$$\alpha_3 = 1.08907 \qquad \alpha_4 = 1.1505 \qquad \alpha_5 = 1.2505$$
$$\alpha_6 = 1.28907 .$$

Figures (6) and (7) represent the evolution of the residuals along the iterations and the CPU time. As one can remark, the incremental schemes, say MWIU, A1IU and MWIU have a much better speed of convergence. We note also that A1IU is more efficient than A1 althrough these algorithms are the same but written in a different way. This is probably due to the number of condition of $M = (I - \Psi)^2$ which may be smaller when it is preconditionned with the I.U. The accuracy is for all methods 5.10^{-8}.

For a better analysis of the relative efficiency of the methods, we have compared in figures (8) and (9) only the incremental algorithms, say MWIU, MWIUa and A1IU. The accuracy is for all methods 5.10^{-9}. We note that the multirelaxation methods (MWIU and MWIUa) have a much better speed of convergence with a regular reduction ratio per iteration (2 for MWIU and 4 for MWIUa). The convergence is obtained by MWIUa in less CPU time than by MWIU. This gain is not very important but it is signifiant because the relaxation is « automatically » provided in MWIUa.

We recover comparable results in dimension one for the calculation of bifurcated solutions of the Chafee-Infante equation.

6. CONCLUSION

Thanks to the several analogies between the classical Richardson Method and the original MW scheme, we have built two efficient families of generalizations of MW involving the incremental unknowns :

- When the relaxation parameter(s) is (are) given at the begining of the program, the nonlinear multirelaxation associated to the I.U. yields a much better speed of convergence and an important gain of CPU time for about the same bassin of attraction (see also [2] and compare MWIU with MW).
- When the relaxation parameter(s) is (are) deduced of the iterates, the numerical results point out again a better speed of convergence obtained by the incremental schemes (compare A1 with A1IU, A1IU with MWIUa).

According to the numerical results, the more efficient scheme built is the MWIUa. It is associated to a minimizing residual process and to a multilevel relaxation. This new algorithm is a new powerful tool for the calculation of unstable solutions. It can be also used for the determination of bifurcation branches (with no turning point).

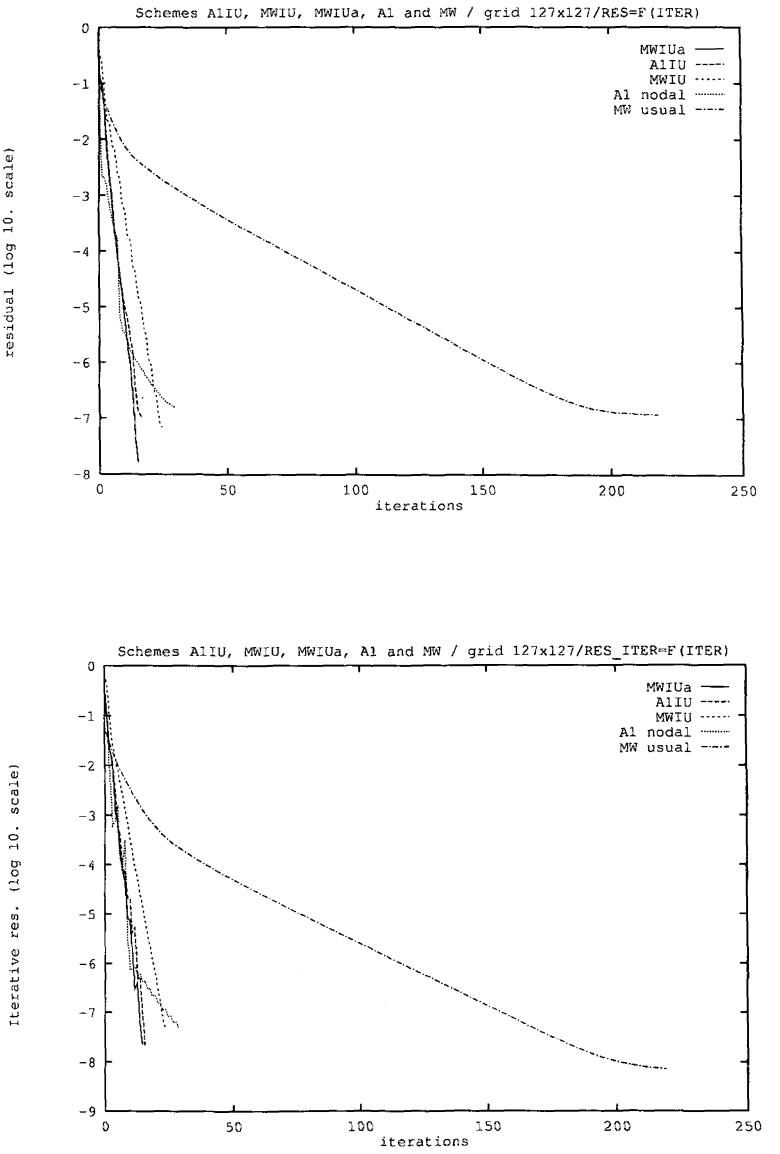
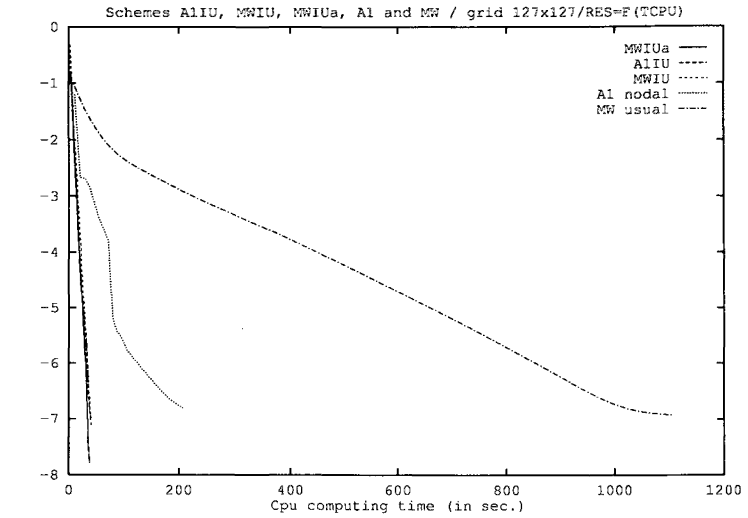


Figure 6. — Comparison of the methods MWIU, MWIUa, A1IU, A1 nodal and MW. The evolution of the residue and the iterative residue are plotted against the iterations. The grid is $C_{1.6}$; $\gamma = \nu = 120$, $\varepsilon = 2$.



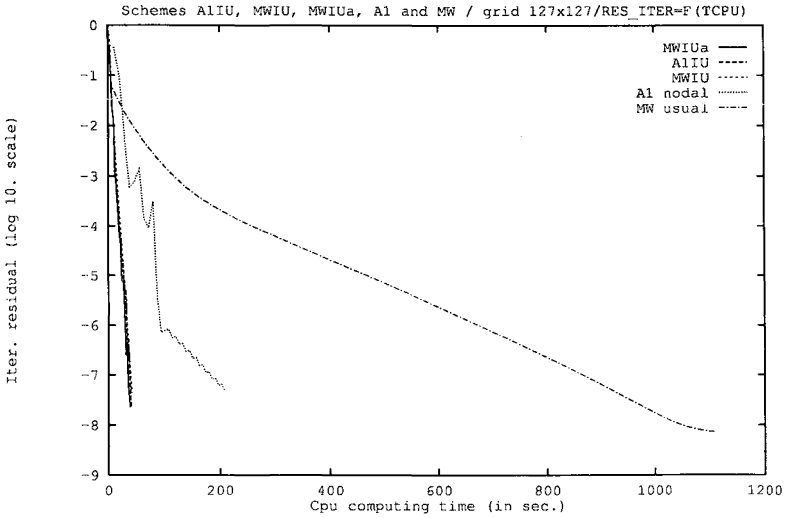


Figure 7. — Comparison of the methods MWIU, MWIUa, A1IU, A1 nodal and MW. The evolution of the residue and the iterative residue are plotted against the CPU time (in seconds). The grid is $C_{1.6}$; $\gamma = \nu = 120$, $\varepsilon = 2$.

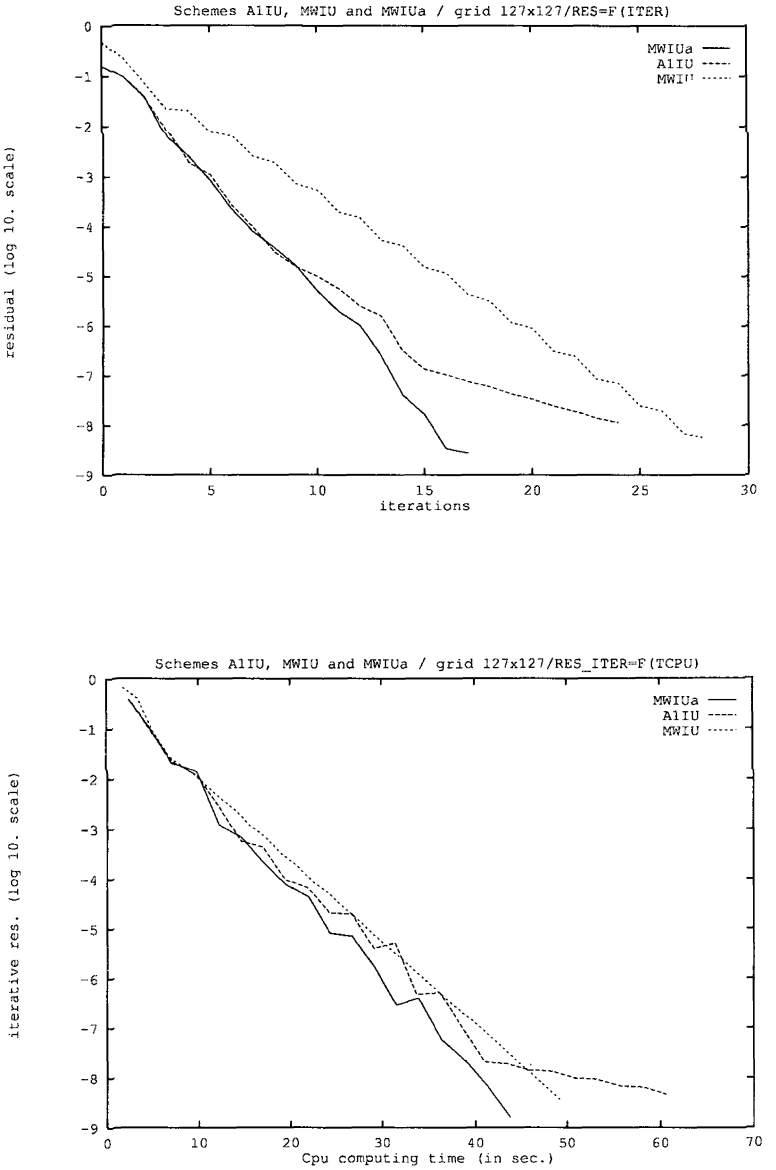


Figure 8. — Comparison of the methods MWIU, MWIUa and A1IU. The evolution of the residue and the iterative residue are plotted against the iterations. The grid is $C_{1.6}$; $\gamma = \nu = 120$, $\varepsilon = 2$.

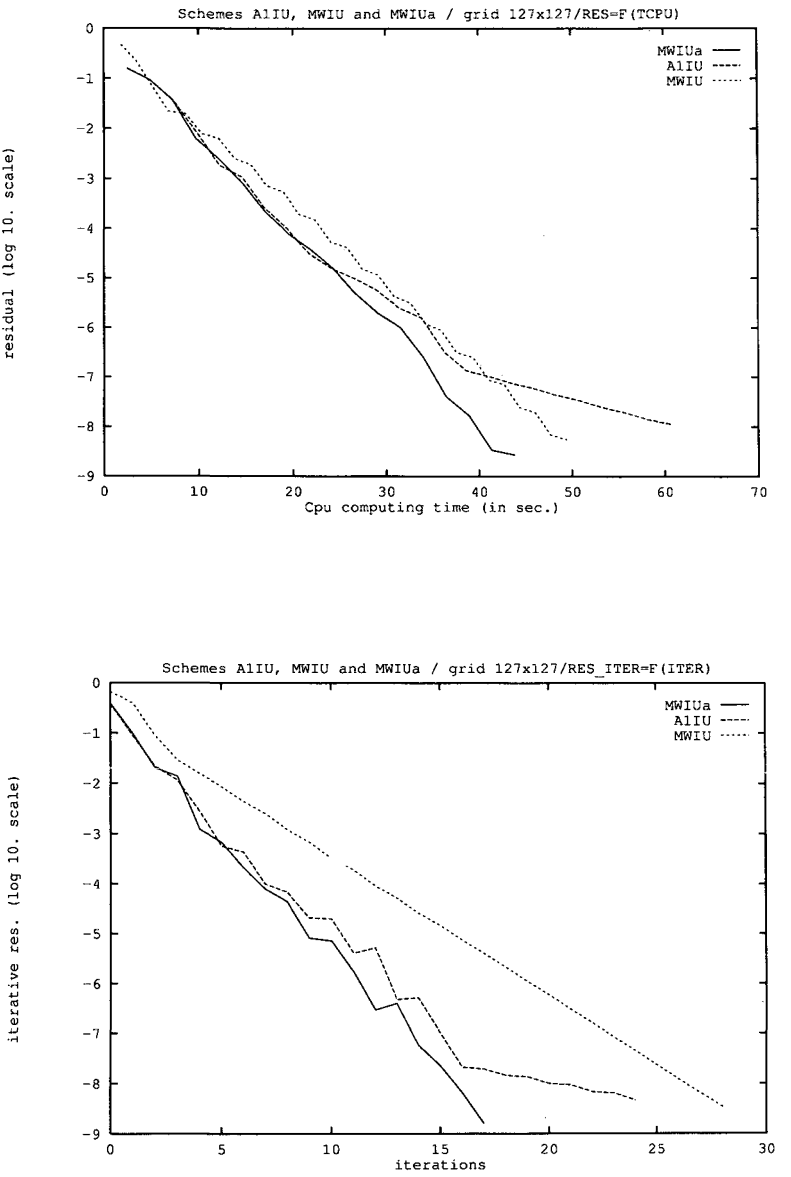


Figure 9. — Comparison of the methods MWIU, MWIUa, A1IU. The evolution of the residue and the iterative residue are plotted against the CPU time. The grid is $C_{1.6}$; $\gamma = \nu = 120$, $\varepsilon = 2$.

REFERENCES

[1] C. BOLLEY, 1978, Multiple Solutions of a Bifurcation Problem, in *Bifurcation and Nonlinear Eigenvalue Problems,* ed. C. Bardos, Proceedings Univ. Paris XIII Villetaneuse, Springer Verlag, n° 782, 42-53.

[2] J.-P. CHEHAB, R. TEMAM, Incremental Unknowns for Solving Nonlinear Eigenvalue Problems. New Multiresolution Methods, *Numerical Methods for PDE's,* **11,** 199-228 (1995).

[3] M. CHEN, R. TEMAM, 1991, Incremental Unknowns for Solving Partial Differential Equations, *Numerische Matematik,* **59,** 255-271.

[4] M. CHEN, R. TEMAM, 1993, Incremental Unknowns in Finite Differences : Condition Number of the Matrix, *SIAM J. of Matrix Analysis and Applications (SIMAX),* **14,** n° 2, 432-455.

[5] G. H. GOLUB, G. A. MEURANT, 1983, *Résolution numérique des grands systèmes linéaires,* Ecole d'été d'Analyse Numérique CEA-EDF-INRIA, Eyrolles.

[6] D. HENRY, 1981, *Geometric Theory of Semilinear Parabolic Equations,* Springer Verlag, n° 840.

[7] H. MARDER, B. WEITZNER, 1970, A Bifurcation Problem in E-layer Equilibria, *Plasma Physics,* **12,** 435-445.

[8] M. MARION, R. TEMAM, 1989, Nonlinear Galerkin Methods, *SIAM Journal of Numerical Analysis,* **26,** 1139-1157.

[9] M. MARION, R. TEMAM, 1990, Nonlinear Galerkin Methods ; The Finite elements case, *Numerische Matematik,* **57,** 205-226.

[10] M. SERMANGE, 1979, Une méthode numérique en bifurcation. Application à un problème à frontière libre de la physique des plasmas, *Applied Mathematics and Optimization,* 127-151.

[11] R. TEMAM, 1990, Inertial Manifolds and Multigrid Methods, *SIAM J. Math. Anal.,* **21,** 154-178.