



Universitat d'Alacant
Universidad de Alicante

Esta tesis doctoral contiene un índice que enlaza a cada uno de los capítulos de la misma.

Existen asimismo botones de retorno al índice al principio y final de cada uno de los capítulos.

[Ir directamente al índice](#)

Para una correcta visualización del texto es necesaria la versión de [Adobe Acrobat Reader 7.0](#) o posteriores

Aquesta tesi doctoral conté un índex que enllaça a cadascun dels capítols. Existeixen així mateix botons de retorn a l'índex al principi i final de cadascun dels capítols .

[Anar directament a l'índex](#)

Per a una correcta visualització del text és necessària la versió d' [Adobe Acrobat Reader 7.0](#) o posteriors.

Universitat d'Alacant
Universidad de Alicante

**Departamento de Tecnología Informática y
Computación**

**Métodos Iterativos Paralelos para
la Resolución de Sistemas Lineales
basados en Multiparticiones**

JOSÉ PENADÉS MARTÍNEZ

Alicante, Noviembre de 1993.





Universitat d'Alacant
Universidad de Alicante

**A mis padres,
gracias por todo.**

**A Violeta,
por estar ahí.**

Índice

Prólogo	xi
1 Métodos iterativos secuenciales	1
1.1 Normas vectoriales y matriciales	2
1.2 Planteamiento y convergencia	6
1.3 Métodos iterativos clásicos	11
1.4 Métodos en dos etapas	19
2 Métodos iterativos paralelos	25
2.1 Introducción a los multiprocesadores	25
2.1.1 Tipos de arquitecturas paralelas	26
2.1.2 Medidas de paralelismo	35
2.2 Antecedentes y estado actual	37
2.2.1 Métodos iterativos basados en la técnica de multipartición	40
2.2.2 Métodos caóticos basados en la técnica de multipartición	49

3	Método de multipartición en dos etapas	61
3.1	Introducción	61
3.2	Modelo síncrono	66
3.3	Modelo asíncrono	88
3.4	Experimentos numéricos	99
3.4.1	Algoritmos	102
3.4.2	Resultados numéricos	105
3.5	Conclusiones	131
4	Método de multipartición relajado	133
4.1	Introducción	133
4.2	Modelo síncrono	136
4.3	Modelo asíncrono	152
4.4	Experimentos numéricos	156
4.4.1	Algoritmos	157
4.4.2	Resultados numéricos	160
4.5	Conclusiones	182
5	Método de multipartición basado en el método AOR	183
5.1	Introducción	183
5.2	Modelo síncrono	188
5.3	Modelo asíncrono	201
5.4	Experimentos numéricos	205
5.4.1	Algoritmos	206
5.4.2	Resultados numéricos	209
5.5	Conclusiones	233

Índice

ix

Líneas Futuras

235

Bibliografía

239

Prólogo

Una gran parte de los problemas físicos y técnicos, al ser formulados mediante un modelo matemático y tratados numéricamente, conducen a la resolución de un sistema algebraico de ecuaciones, usualmente de orden elevado. De acuerdo con el carácter físico del problema, dicho sistema algebraico será lineal o no. Por ejemplo, al discretizar ecuaciones en derivadas parciales por técnicas numéricas tan usuales hoy en día como las diferencias finitas o el método de elementos finitos, nos encontramos finalmente con un gran sistema de ecuaciones lineales a resolver.

Las ecuaciones en derivadas parciales surgen a la hora de resolver multitud de problemas. De los métodos de aproximación numérica que se emplean para la resolución de ecuaciones en derivadas parciales, los que se basan en diferencias finitas y elementos finitos son los más frecuentemente utilizados, proporcionando una solución tan satisfactoria como la que se obtiene mediante técnicas analíticas.

La actual posibilidad de disponer de computadores con una tecnología muy avanzada ha tenido un gran impacto en el mundo de la computación científica.

En concreto, son las arquitecturas paralelas las que nos van a permitir grandes velocidades de computación.

Son muchas las áreas de la ciencia y del conocimiento en las que son de gran utilidad estas arquitecturas. Así, podemos citar ejemplos tales como pronósticos meteorológicos a nivel mundial, búsqueda de bolsas de petróleo, predicción de terremotos, reconocimiento de imágenes, inteligencia artificial, . . . En la computación matricial son especialmente adecuadas estas arquitecturas, debido tanto al coste normalmente elevado de resolución de los problemas correspondientes a este campo, como a la dificultad a la hora de manejar estructuras de datos de gran tamaño o con una configuración especial (matrices con muchos elementos nulos).

La aparición de los *multiprocesadores* ha dado lugar no sólo a una adecuación de los algoritmos clásicos de la computación matricial para su implementación y ejecución en dichos multiprocesadores, sino que además sugiere y motiva la búsqueda de nuevos métodos que obtengan un buen rendimiento de máquina. Estas reflexiones, nos llevan al objetivo general que se pretende con esta Tesis: *Diseño y estudio de métodos iterativos paralelos para la resolución de grandes sistemas de ecuaciones lineales, basados en multiparticiones.*

Nos vamos a centrar en el planteamiento de métodos iterativos paralelos, prestando especial atención a la distribución de la carga entre los procesadores. Uno de los propósitos de los *modelos caóticos* es que dicha carga de trabajo entre los procesadores esté equilibrada mediante la realización, por parte de cada procesador, de ciertos cálculos locales, independientes del resto de procesadores.

Esta memoria está dividida en cinco capítulos, atendiendo a la siguiente estructura.

El Capítulo 1, titulado *Métodos iterativos secuenciales*, consta de cuatro secciones. En la primera sección se recuerdan aquellos conceptos relativos a *normas vectoriales y matriciales* que nos han sido de utilidad en la elaboración de esta memoria. En la Sección 1.2 se estudian de forma general los métodos iterativos secuenciales. Los métodos iterativos clásicos serán objeto de estudio en la Sección 1.3. Haremos mención especial, en la Sección 1.4, a los métodos iterativos en dos etapas.

El objetivo principal del Capítulo 2, *Métodos iterativos paralelos*, es revisar los resultados conocidos, hasta el momento, sobre estos métodos. Así, la Sección 2.1 nos introduce en el concepto de *multiprocesador* y en la Sección 2.2 se hace una revisión de los métodos iterativos que se utilizan para la resolución de un sistema de ecuaciones lineales sobre un multiprocesador. Nos detenemos en aquellos que se basan en la técnica de multipartición y dentro de éstos, en los caóticos.

La parte principal de nuestra investigación la forman los Capítulos 3, 4 y 5, titulados respectivamente:

- *Método de multipartición en dos etapas.*
- *Método de multipartición relajado.*
- *Método de multipartición basado en el método AOR (acceleration overrelaxation).*

Todos ellos tienen la misma estructura. En la primera sección, se introducen los resultados que han motivado el estudio del método en cuestión. La segunda y tercera sección abordan el problema concreto, planteando el nuevo método y estudiando la convergencia para el caso síncrono y asíncrono respectivamente. A

continuación, también en cada uno de estos capítulos, se presentan los resultados numéricos obtenidos al implementar dichos métodos en un multiprocesador con memoria compartida, en concreto en el Alliant FX/80. Los Capítulos 3, 4 y 5 finalizan con una sección dedicada a las conclusiones, en la que recordamos los resultados más importantes que hemos obtenido, tanto desde el punto de vista teórico como práctico.

Finalizamos esta memoria con dos apéndices. El primero está dedicado a las líneas futuras de trabajo y el segundo a las referencias bibliográficas utilizadas para la elaboración de esta memoria.

Capítulo 1

Métodos iterativos secuenciales

Consideremos el sistema de ecuaciones lineales

$$Ax = b, \quad (1.1)$$

donde A es una matriz invertible de tamaño $n \times n$ y b un vector columna de tamaño n . Existen, básicamente, dos familias de métodos para resolver estos sistemas de ecuaciones, los *métodos directos* y los *métodos iterativos*.

Puesto que la matriz A es invertible, la solución del sistema lineal (1.1) es única. Los métodos directos obtienen la solución exacta, salvo errores de redondeo. Estos métodos se basan en sucesivas transformaciones algebraicas de la matriz inicial A , y en ellos el número de operaciones requeridas para obtener la solución se puede predecir.

Nosotros, sin embargo, vamos a dedicar esta memoria al estudio de los *métodos iterativos*. En ellos, no se obtiene la solución exacta, sino una aproximación, pero la importancia de los errores de redondeo disminuye. Realizaremos un



estudio detallado de estos métodos desde el punto de vista secuencial en las Secciones 1.2, 1.3 y 1.4.

En el estudio de los métodos iterativos, juegan un papel muy importante los conceptos de norma vectorial, norma matricial y radio espectral. Hacemos una revisión de dichos conceptos en la siguiente sección.

1.1 Normas vectoriales y matriciales

Antes de revisar los conceptos relativos a normas, para que no exista ningún tipo de ambigüedad, introducimos la notación relativa a vectores y matrices definidos sobre el cuerpo de los números reales, que denotaremos por \mathbf{R} .

Para todo $\mathbf{x} \in \mathbf{R}^n$, denotaremos por $|\mathbf{x}|$ el vector cuyas componentes son los valores absolutos de las correspondientes componentes de \mathbf{x} .

Diremos que un vector \mathbf{x} es *positivo* (respectivamente, *no negativo*) y escribiremos $\mathbf{x} > \mathbf{0}$ (respectivamente, $\mathbf{x} \geq \mathbf{0}$), si tiene todas sus componentes estrictamente mayores que 0 (respectivamente, mayores o iguales que 0).

Diremos que $\mathbf{x} \leq \mathbf{y}$, si $\mathbf{y} - \mathbf{x} \geq \mathbf{0}$. Análogamente, diremos que $\mathbf{x} < \mathbf{y}$, si $\mathbf{y} - \mathbf{x} > \mathbf{0}$.

Una notación similar se ha utilizado para matrices. Denotaremos por \mathbf{K} el cuerpo \mathbf{R} o \mathbf{C} .

Definición 1. Sea V un espacio vectorial sobre el cuerpo \mathbf{K} . Diremos que la función $\|\cdot\| : V \rightarrow \mathbf{R}$ es una *norma vectorial*, si para todo $\mathbf{x}, \mathbf{y} \in V$, se verifican las siguientes propiedades.

1.1 Normas vectoriales y matriciales**3**

- (i) $\|x\| \geq 0$.
- (ii) $\|x\| = 0$, si y sólo si, $x = 0$.
- (iii) $\|cx\| = |c|\|x\|$, $\forall c \in \mathbb{K}$.
- (iv) $\|x + y\| \leq \|x\| + \|y\|$.

Demos algunos ejemplos conocidos de normas vectoriales sobre \mathbb{R}^n .

- Las l_p -normas, que para un vector $x = (x_1, x_2, \dots, x_n)$ se definen como

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \quad p \geq 1.$$

Nótese que la l_2 -norma es la conocida norma euclídea.

- La norma infinito, definida para un vector x de la siguiente forma

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|. \quad (1.2)$$

- Dado un vector positivo $z \in \mathbb{R}^n$, definimos la norma $\|\cdot\|_z$, como sigue

$$\|z\|_z = \inf\{\alpha > 0 \mid -\alpha z \leq z \leq \alpha z\}, \quad z \in \mathbb{R}^n. \quad (1.3)$$

Las normas que se han definido en las expresiones anteriores son *monótonas*, es decir si v y w son vectores de \mathbb{R}^n tales que $|v| \leq |w|$, entonces $\|v\| \leq \|w\|$.

Designamos por \mathcal{M}_n el espacio vectorial formado por todas las matrices cuadradas de tamaño $n \times n$ definidas sobre el cuerpo \mathbb{K} .

A continuación, definimos en este espacio vectorial, las normas matriciales.

Definición 2. Diremos que la función $\|\cdot\| : \mathcal{M}_n \rightarrow \mathbf{R}$ es una *norma matricial*, si para todo par de matrices A y B de \mathcal{M}_n , se verifican los siguientes axiomas

- (i) $\|A\| \geq 0$.
- (ii) $\|A\| = 0$, si y sólo si, $A = O$.
- (iii) $\|cA\| = |c|\|A\|$, $\forall c \in \mathbf{K}$.
- (iv) $\|A + B\| \leq \|A\| + \|B\|$.
- (v) $\|AB\| \leq \|A\|\|B\|$.

Hacemos notar que una norma matricial es una norma vectorial que satisface además el axioma (v), o propiedad *submultiplicativa*.

Para estudiar la convergencia de los métodos iterativos paralelos, que planteamos en los Capítulos 3, 4 y 5 se han utilizado, en ocasiones, normas matriciales que se deducen de las normas vectoriales, utilizando la siguiente definición.

Definición 3. Sea $\|\cdot\|$ una norma vectorial sobre \mathbf{K}^n . Se llama *norma matricial inducida* por dicha norma vectorial, a la siguiente función definida sobre \mathcal{M}_n .

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

La norma introducida en la Definición 3 puede ser calculada de las siguientes formas equivalentes

$$\|A\| = \max_{\|x\|=1} \|Ax\| = \max_{\|x\| \leq 1} \|Ax\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

1.1 Normas vectoriales y matriciales

5

Toda norma matricial inducida por una norma vectorial es *compatible*, es decir, verifica la siguiente desigualdad

$$\|Ax\| \leq \|A\|\|x\|, \quad x \in \mathbb{K}^n.$$

Además, si denotamos por I a la matriz identidad y $\|\cdot\|$ es una norma compatible, entonces $\|I\| = 1$.

A continuación, como ejemplos de normas matriciales inducidas por normas vectoriales, damos las que nos serán de utilidad en el desarrollo de esta memoria.

- La *norma matricial infinito*, inducida por la norma vectorial infinito

$$\|A\|_{\infty} = \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |a_{ij}| \right), \quad \text{si } A = [a_{ij}]_{1 \leq i, j \leq n}.$$

- La *norma matricial* $\|\cdot\|_x$ inducida por la norma vectorial (1.3). Nos será útil la siguiente propiedad de esta norma, probada por Rheinboldt y Vandergraft [68], en 1973: para toda matriz B , se cumple

$$\|B\|_x = \| |B|_x \|_x. \quad (1.4)$$

Si una norma matricial cumple la propiedad (1.4), se dice que es una *norma matricial monótona*.

Otro concepto de gran importancia para el estudio de la convergencia de los métodos iterativos, es el de *radio espectral* de una matriz.

Definición 4. Sea A una matriz cuadrada de tamaño $n \times n$, definimos el *radio espectral* de A , y lo denotamos por $\rho(A)$, como el máximo de los módulos de los valores propios de A , es decir

$$\rho(A) = \max\{|\lambda_i| : \lambda_i \text{ es valor propio de } A\}$$

Es conocido, y así se refleja en el teorema siguiente (ver por ejemplo, Varga [78] y Young [83]), que cualquier norma matricial es una cota superior del radio espectral.

Teorema 1. *Si $\|\cdot\|$ es una norma matricial y $A \in \mathcal{M}_n$, entonces se satisface $\rho(A) \leq \|A\|$.*

El teorema que se enuncia a continuación, asegura que para toda matriz se puede encontrar una norma matricial de forma que dicha norma esté tan cerca del radio espectral como queramos (ver Varga [78] y Young [83]).

Teorema 2. *Sea $A \in \mathcal{M}_n$. Para todo $\epsilon > 0$, existe al menos una norma matricial $\|\cdot\|$ tal que $\rho(A) \leq \|A\| \leq \rho(A) + \epsilon$.*

Además (ver Horn y Johnson [48]), dada una norma matricial $\|\cdot\|_\alpha$ existe siempre una norma matricial inducida $\|\cdot\|_\beta$, y por tanto compatible, que verifica

$$\|A\|_\beta \leq \|A\|_\alpha.$$

Utilizando los conceptos revisados en esta sección, vamos a estudiar, en la Sección 1.2 las generalidades sobre los métodos iterativos secuenciales, para posteriormente, dedicar la Sección 1.3 a los métodos iterativos clásicos y la Sección 1.4 a los métodos iterativos en dos etapas.

1.2 Planteamiento y convergencia

Los métodos iterativos para la resolución de un sistema de ecuaciones lineales consisten en generar, a partir de un vector inicial $x^{(0)}$, una sucesión de vectores

1.2 Planteamiento y convergencia

7

$\{\mathbf{x}^{(k)}\}_{k \geq 1}$ que converja a la solución exacta $\xi = A^{-1}\mathbf{b}$ del mismo. Se dice que dicho método es *convergente* si para todo vector inicial $\mathbf{x}^{(0)}$, se cumple que $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \xi$.

El concepto de norma vectorial, dado en la Definición 1, se utiliza para el estudio de la convergencia de una sucesión de vectores, de la siguiente forma.

Definición 5. Sea V un espacio vectorial sobre el cuerpo K y sea $\|\cdot\|$ una norma vectorial sobre V . Decimos que la sucesión de vectores $\{\mathbf{x}^{(k)}\}_{k \geq 1}$ converge al vector $\mathbf{x} \in V$ con respecto a la norma $\|\cdot\|$, si y sólo si, la sucesión de números reales $\{\|\mathbf{x}^{(k)} - \mathbf{x}\|\}_{k \geq 1}$ tiende a 0 cuando k tiende a infinito.

En los espacios de dimensión finita, todas las normas son *equivalentes*, en el sentido en que si una sucesión $\{\mathbf{x}^{(k)}\}_{k \geq 1}$ converge al vector \mathbf{x} respecto a una norma, entonces converge al mismo vector con respecto a cualquier otra norma (ver Horn y Johnson [48]). Esto significa que la convergencia de una sucesión de vectores, puede ser estudiada con respecto a cualquier norma vectorial.

Para resolver el sistema (1.1) por procedimientos iterativos, consideraremos la matriz A expresada de la forma $A = M - N$, donde M y N son también matrices cuadradas de orden n .

Definición 6. Sea la matriz $A \in \mathcal{M}_n$, tal que

$$A = M - N, \quad M, N \in \mathcal{M}_n. \quad (1.5)$$

Si M es no singular, se dice que la expresión (1.5) es una *partición* de la matriz A .

A partir de la partición (1.5) de la matriz A , el sistema lineal $A\mathbf{x} = \mathbf{b}$ se puede escribir como

$$M\mathbf{x} = N\mathbf{x} + \mathbf{b}. \quad (1.6)$$

Para obtener la solución de (1.1), podemos iterar la ecuación (1.6) mediante el esquema

$$M\mathbf{x}^{(k)} = N\mathbf{x}^{(k-1)} + \mathbf{b}, \quad k = 1, 2, \dots, \quad (1.7)$$

o análogamente

$$\mathbf{x}^{(k)} = M^{-1}N\mathbf{x}^{(k-1)} + M^{-1}\mathbf{b}, \quad k = 1, 2, \dots, \quad (1.8)$$

y estudiar bajo qué condiciones de las matrices A , M y N , la sucesión de vectores $\{\mathbf{x}^{(k)}\}_{k \geq 1}$ generada por el esquema iterativo (1.8) converge a la solución exacta del sistema lineal (1.1). Se llama *matriz de iteración* a la matriz $H = M^{-1}N$.

En principio, se estudiaron de forma independiente diversas condiciones de convergencia para los métodos usuales conocidos con el nombre de Jacobi, Gauss-Seidel y SOR. Sin embargo, Varga [78] en 1962 introdujo el concepto de partición regular de una matriz, que en determinados casos engloba a los diversos métodos antes mencionados, e intentó dar las condiciones de convergencia sobre este tipo de particiones.

Definición 7. Se dice que la partición $A = M - N$ es *regular* si M^{-1} y N son matrices no negativas (es decir, si $M^{-1} \geq O, N \geq O$).

Posteriormente al concepto de partición regular se definió el siguiente tipo de partición.

Definición 8. La partición $A = M - N$ se dice que es *débilmente regular* si $M^{-1} \geq O$ y $M^{-1}N \geq O$.

Como se deduce de las definiciones de ambos tipos de particiones, toda partición regular es débilmente regular, ya que el producto de matrices no negativas es otra matriz no negativa.

1.2 Planteamiento y convergencia

9

En la versión original de la definición de partición débilmente regular (Definición 8), dada por Ortega y Rheinboldt [63], se incluía la hipótesis adicional $NM^{-1} \geq O$. Sin embargo, nosotros hemos elegido la definición más usada comúnmente en la literatura.

La convergencia del método (1.8) se puede estudiar analizando el error que se comete en cada iteración respecto de la solución exacta ξ . Así, si la ecuación (1.6) la escribimos de la forma

$$\xi = H\xi + M^{-1}b,$$

y si esta expresión se resta a la ecuación (1.8), se obtiene

$$e^{(k)} = He^{(k-1)}, \quad k = 1, 2, \dots, \quad (1.9)$$

donde

$$e^{(k)} = x^{(k)} - \xi,$$

se llama *vector error* en la k -ésima iteración. La ecuación (1.9) es equivalente a

$$e^{(k)} = H^k e^{(0)}, \quad k = 1, 2, \dots \quad (1.10)$$

Así pues, la convergencia de la sucesión definida por la ecuación (1.8) a ξ es equivalente a que la sucesión de vectores error $\{e^{(k)}\}_{k \geq 1}$, definida en la expresión (1.10), converja al vector cero. Es fácil demostrar (ver Young [83]) que esto ocurre, si y sólo si, las potencias de la matriz H convergen a la matriz nula, es decir si $\lim_{k \rightarrow \infty} H^k = O$.

El resultado teórico básico de convergencia del método iterativo (1.8) está determinado por el *radio espectral* de H y su demostración se basa en la forma canónica de Jordan de esta matriz, tal y como puede verse en Young [83].

Teorema 3. *Sea H una matriz cuadrada. Entonces $\{H^k\}_{k \geq 1}$ converge a la matriz nula, si y sólo si, $\rho(H) < 1$.*

En consecuencia, el método iterativo (1.8) converge a la solución del sistema lineal (1.1), para cualquier vector inicial $x^{(0)}$, si y sólo si, $\rho(M^{-1}N) < 1$.

Cuando la matriz A es *monótona*, es decir, cuando su inversa es no negativa ($A^{-1} \geq O$), existe un resultado de convergencia, que se basa en el siguiente lema (ver por ejemplo, Berman y Plemmons [7] y Varga [78]).

Lema 1. *Supongamos que la partición $A = M - N$ es regular o débilmente regular, entonces*

$$\rho(M^{-1}N) < 1 \iff A^{-1} \geq O.$$

El siguiente resultado de convergencia puede verse en Berman y Plemmons [7] y se refiere a particiones regulares o débilmente regulares.

Teorema 4. *Sea A una matriz monótona, y sea $A = M - N$ una partición regular o débilmente regular. Entonces el método (1.8) es convergente.*

El radio espectral de la matriz de iteración $H = M^{-1}N$ se utiliza para dar una medida de la rapidez de convergencia, ya que cuando menor sea el radio espectral, más rápida será la convergencia. Cuando un esquema iterativo posea una matriz de iteración con menor radio espectral, diremos que el esquema posee un *radio de convergencia superior*.

1.3 Métodos iterativos clásicos

Como hemos visto en la Sección 1.2, cada partición $A = M - N$, de la matriz del sistema lineal $Ax = b$, define un esquema iterativo como el que aparece en la expresión (1.8). Las distintas elecciones de las matrices M y N , dan como resultado los distintos métodos iterativos clásicos.

Supongamos que todos los elementos de la diagonal de $A = [a_{ij}]_{1 \leq i, j \leq n}$ son no nulos y que la matriz A se expresa de la siguiente forma

$$A = D - L - U, \quad (1.11)$$

donde $D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$, y L y U son respectivamente la parte estrictamente triangular inferior y superior de la matriz A .

Teniendo en cuenta la expresión (1.11), podemos escribir el sistema (1.1) como

$$Dx = (L + U)x + b. \quad (1.12)$$

Como los elementos diagonales a_{ii} , de A son no nulos, podemos establecer el siguiente método iterativo derivado de la expresión (1.12)

$$a_{ii}x_i^{(m+1)} = - \sum_{j=1, j \neq i}^n a_{ij}x_j^{(m)} + b_i, \quad 1 \leq i \leq n, \quad m = 0, 1, \dots, \quad (1.13)$$

donde $x_i^{(0)}$, $1 \leq i \leq n$, son las componentes de un vector inicial $x^{(0)}$.

Claramente, la expresión (1.13) puede ser reescrita como

$$x_i^{(m+1)} = - \sum_{j=1, j \neq i}^n \frac{a_{ij}}{a_{ii}} x_j^{(m)} + \frac{b_i}{a_{ii}}, \quad 1 \leq i \leq n, \quad m = 0, 1, \dots \quad (1.14)$$

La expresión matricial de (1.13) es

$$Dx^{(m+1)} = (L + U)x^{(m)} + b, \quad m = 0, 1, \dots,$$

y como D es una matriz no singular, la notación matricial de (1.14) es

$$\mathbf{x}^{(m+1)} = D^{-1}(L + U)\mathbf{x}^{(m)} + D^{-1}\mathbf{b}, \quad m = 0, 1, \dots$$

Este método iterativo se denomina *método de Jacobi*. Claramente, las siguientes elecciones de M y N en la partición $A = M - N$, dan lugar a dicho método

$$M = D \quad \text{y} \quad N = L + U. \quad (1.15)$$

La matriz de iteración asociada

$$B = M^{-1}N = D^{-1}(L + U),$$

se denomina usualmente *matriz de Jacobi* asociada a la matriz A .

Como se puede observar, en el método de Jacobi, para el cálculo de la i -ésima componente del vector $\mathbf{x}^{(m+1)}$, se utilizan los valores de las componentes $x_1^{(m)}, x_2^{(m)}, \dots, x_{i-1}^{(m)}, x_{i+1}^{(m)}, \dots, x_n^{(m)}$. Pero en el momento de calcular dicha componente i -ésima ya han sido calculadas las componentes $x_1^{(m)}, x_2^{(m)}, \dots, x_{i-1}^{(m)}$. Surge entonces, de modo natural, la pregunta siguiente: ¿Por qué no aprovechar la información que se acaba de obtener sobre las primeras $(i - 1)$ componentes del vector $\mathbf{x}^{(m+1)}$ en el cálculo de la componente i -ésima?

Una primera ventaja de actuar así sería la de no tener que almacenar dos vectores en el algoritmo de cálculo, pues las componentes del vector $\mathbf{x}^{(m+1)}$ se almacenarían en el lugar ocupado por las correspondientes del vector $\mathbf{x}^{(m)}$ (ya que éstas no serían necesarias en los sucesivos cálculos). Además parece probable que al utilizar datos más cercanos a la solución exacta en el cálculo de las componentes de $\mathbf{x}^{(m+1)}$, la velocidad de convergencia mejore.

El método iterativo de Gauss-Seidel consiste precisamente en actuar de dicha manera. Con ello el esquema de cálculo del método resulta ser:

$$x_i^{(m+1)} = - \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(m+1)} - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} x_j^{(m)} + \frac{b_i}{a_{ii}}, \quad 1 \leq i \leq n, \quad m = 0, 1, \dots$$

Matricialmente este método se puede escribir como sigue

$$(D - L)x^{(m+1)} = Ux^{(m)} + b, \quad m = 0, 1, \dots, \quad (1.16)$$

y como la matriz triangular inferior $D - L$ es no singular, (1.16) puede ser escrito equivalentemente como

$$x^{(m+1)} = (D - L)^{-1}Ux^{(m)} + (D - L)^{-1}b, \quad m = 0, 1, \dots$$

Análogamente al método de Jacobi, el método iterativo de Gauss-Seidel puede obtenerse considerando una partición $A = M - N$, en este caso

$$M = D - L \quad \text{y} \quad N = U.$$

Con el fin de acelerar la convergencia, a todo método iterativo de la forma (1.8) se le puede asociar un método extrapolado o relajado sustituyendo en cada etapa m , el vector $x^{(m+1)}$, por el vector extrapolado

$$\omega x^{(m+1)} + (1 - \omega)x^{(m)}, \quad \omega \neq 0, \quad \omega \in \mathbb{R}.$$

Esto sólo requiere un pequeño esfuerzo computacional adicional. Esta relajación corresponde al siguiente esquema iterativo

$$x^{(m+1)} = H(\omega)x^{(m)} + \omega M^{-1}b, \quad m = 0, 1, \dots,$$

con la matriz de iteración

$$H(\omega) = (1 - \omega)I + \omega M^{-1}N. \quad (1.17)$$

A este método iterativo se le llama ω -extrapolación del esquema (1.8) o también ω -relajación, siendo ω el factor de relajación. Lo denominaremos simplemente método extrapolado o relajado cuando no exista ambigüedad en la elección de ω .

Si en (1.17) utilizamos la partición de Jacobi, se obtiene el método de Jacobi relajado (JOR). El esquema iterativo es

$$\mathbf{x}^{(m+1)} = [(1 - \omega)I + \omega D^{-1}(L + U)] \mathbf{x}^{(m)} + \omega D^{-1}\mathbf{b}, \quad m = 0, 1, \dots,$$

o análogamente

$$x_i^{(m+1)} = -\omega \sum_{j=1, j \neq i}^n \frac{a_{ij}}{a_{ii}} x_j^{(m)} + (1 - \omega)x_i^{(m)} + \omega \frac{b_i}{a_{ii}}, \quad 1 \leq i \leq n, \quad m = 0, 1, \dots$$

El método de Gauss-Seidel puede ser sucesivamente ω -extrapolado, de la siguiente forma. Primero definimos el vector iterado auxiliar $\bar{\mathbf{x}}^{(m+1)}$ como

$$\bar{x}_i^{(m+1)} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(m+1)} - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} x_j^{(m)} + \frac{b_i}{a_{ii}}, \quad 1 \leq i \leq n, \quad m = 0, 1, \dots$$

La componente $x_i^{(m+1)}$ de este método iterativo se define como

$$x_i^{(m+1)} = \omega \bar{x}_i^{(m+1)} + (1 - \omega)x_i^{(m)}.$$

Combinando estas dos últimas expresiones obtenemos

$$x_i^{(m+1)} = -\omega \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(m+1)} - \omega \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} x_j^{(m)} + (1 - \omega)x_i^{(m)} + \omega \frac{b_i}{a_{ii}}. \quad (1.18)$$

La expresión matricial de este método es

$$(D - \omega L)\mathbf{x}^{(m+1)} = [(1 - \omega)D + \omega U] \mathbf{x}^{(m)} + \omega \mathbf{b}, \quad m = 0, 1, \dots,$$

y como la matriz $D - \omega L$ es no singular para cualquier elección de ω , entonces obtenemos

$$\mathbf{x}^{(m+1)} = (D - \omega L)^{-1} [(1 - \omega)D + \omega U] \mathbf{x}^{(m)} + \omega(D - \omega L)^{-1} \mathbf{b}, \quad m = 0, 1, \dots$$

El método iterativo que acabamos de contruir recibe el nombre de *método de sobrerrelajaciones sucesivas (SOR)*. La matriz de iteración se denota por

$$\mathcal{L}_\omega = (D - \omega L)^{-1} [(1 - \omega)D + \omega U].$$

Seleccionando $\omega = 1$, este método iterativo se reduce exactamente al método de Gauss-Seidel.

Las siguientes elecciones de M y N en (1.8) dan lugar respectivamente a los métodos extrapolados *JOR* y *SOR*.

$$\text{Método } JOR: \quad M = \omega^{-1} D \quad \text{y} \quad N = \omega^{-1} [(1 - \omega)D + \omega(L + U)]$$

$$\text{Método } SOR: \quad M = \omega^{-1}(D - \omega L) \quad \text{y} \quad N = \omega^{-1} [(1 - \omega)D + \omega U],$$

donde ω es el factor de relajación.

Hadjidimos [46] en 1978, estudia un nuevo método de resolución de sistemas de ecuaciones lineales. Este método constituye una generalización del método *SOR* en dos parámetros y su formulación es la siguiente

$$(D - \mu L) \mathbf{x}^{(m+1)} = [(1 - \omega)D + (\omega - \mu)L + \omega U] \mathbf{x}^{(m)} + \omega \mathbf{b}, \quad (1.19)$$

$$m = 0, 1, \dots,$$

donde los números reales μ y $\omega \neq 0$, son dos parámetros fijos.

Este método se denomina método de sobrerrelajación acelerada, y lo denotaremos indistintamente por método *AOR* o $M_{\mu, \omega}$ -método.

Claramente, para valores específicos de los parámetros μ y ω el $M_{\mu, \omega}$ -método se reduce a los métodos clásicos, definidos anteriormente. Así:

- El $M_{0,1}$ -método es el método de Jacobi.
- El $M_{1,1}$ -método es el método de Gauss-Seidel.
- El $M_{0,\omega}$ -método es el método *JOR*.
- El $M_{\omega,\omega}$ -método es el método *SOR*.

A partir de ahora, llamaremos a μ , parámetro de aceleración y a ω , parámetro de sobrerrelajación. Además denotaremos por $\mathcal{L}_{\mu,\omega}$ la matriz de iteración del método *AOR*, que viene dada por

$$\mathcal{L}_{\mu,\omega} = (D - \mu L)^{-1}[(1 - \omega)D + (\omega - \mu)L + \omega U].$$

A continuación estudiaremos la convergencia de estos métodos clásicos, para el caso de matrices diagonal dominantes e irreducibles. Dichos conceptos, que introducimos a continuación, pueden verse, por ejemplo, en Varga [78].

Definición 9. Para $n \geq 2$, una matriz A de tamaño $n \times n$ es *reducible* si existe una matriz de permutación P de tamaño $n \times n$ tal que

$$PAP^t = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix},$$

donde A_{11} es una submatriz de tamaño $r \times r$ y A_{22} es una submatriz de tamaño $(n-r) \times (n-r)$, con $1 \leq r < n$. Si no existe tal matriz de permutación, entonces A se llama *irreducible*.

Definición 10. Sea $A = [a_{ij}]_{1 \leq i,j \leq n}$ una matriz real de tamaño $n \times n$. Se dice que

(i) A es *diagonal dominante* si

$$|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{ij}|, \quad i = 1, 2, \dots, n. \quad (1.20)$$

(ii) A es *estrictamente diagonal dominante* si las n desigualdades (1.20) son estrictas.

(iii) A es *irreduciblemente diagonal dominante* si A es una matriz irreducible y diagonal dominante con al menos una desigualdad (1.20) estricta.

Si A es una matriz irreduciblemente diagonal dominante, entonces será también invertible con sus elementos diagonales no nulos. En tal caso, se puede demostrar el siguiente teorema (ver Hadjidimos [46]).

Teorema 5. *Sea A irreduciblemente diagonal dominante. Entonces el método AOR converge para $0 \leq \mu \leq 1$ y $0 < \omega \leq 1$.*

Puesto que el AOR generaliza a los otros métodos clásicos, a partir del Teorema 5 obtenemos el siguiente corolario.

Corolario 1. *Sea A irreduciblemente diagonal dominante. Entonces los métodos de Jacobi, Gauss-Seidel, JOR y SOR (estos dos últimos para $0 < \omega \leq 1$) convergen.*

Hadjidimos, en [46], también estudia la convergencia del método AOR, cuando la matriz $A = [a_{ij}]$ es una *L-matriz*, esto es, cuando los elementos de A satisfacen la siguiente relación

$$a_{ii} > 0, \quad i = 1, 2, \dots, n \quad \text{y} \quad a_{ij} \leq 0, \quad i \neq j, \quad i, j = 1, 2, \dots, n.$$

Teorema 6. *Si A es una L -matriz, entonces para todo μ y ω tal que $0 \leq \mu \leq \omega \leq 1$ ($\omega \neq 0$) el método AOR converge, si y sólo si, el método de Jacobi converge.*

En el Corolario 1 y en el Teorema 6 se estudia la convergencia de los métodos relajados, cuando $0 < \omega \leq 1$; sin embargo, para el método SOR se sabe que el valor óptimo de ω suele ser mayor o igual que 1. Hay que tener en cuenta que este método puede converger cuando el factor de relajación cumple $0 < \omega < 2$, pero nunca en otro caso, como expresa el siguiente resultado (ver Varga [78]).

Lema 2 (Lema de Kahan). *El método iterativo SOR no puede converger a menos que $0 < \omega < 2$.*

Muchos son los autores que han estudiado los métodos iterativos clásicos. En particular, existen condiciones de convergencia para los métodos clásicos cuando la matriz A es simétrica y definida positiva. Estos resultados pueden verse, por ejemplo, en Varga [78] o en Ortega [62].

En 1984, Albrecht y Klein [2] dan un teorema de convergencia para una matriz arbitraria A ; la condición necesaria y suficiente que exigen a la matriz A , con un ω apropiado, para que el método JOR converja es que todos sus valores propios pertenezcan exclusivamente a C^- o C^+ (parte real negativa o positiva respectivamente). Hacen notar además, que el método JOR converge especialmente si A es definida positiva o definida negativa.

Hadjidimos, en el artículo mencionado anteriormente (ver [46]), analiza la convergencia del método AGOR, para el caso en que A es consistentemente ordenada, es decir cuando $\det(\alpha L^{-1}L + \alpha^{-1}D^{-1}U - \beta D)$ es independiente de α para

$\alpha \neq 0$ y para todo β . Además muestra con ejemplos numéricos la superioridad del método *AOR*.

Recientemente, en 1991, Dancis en [15] considera la convergencia del sistema de ecuaciones lineales $Ax = b$, cuando la teoría usual del método *SOR* es aplicable tal y como puede estudiarse en Young [83].

No obstante, aunque la convergencia de un método iterativo esté asegurada, tanto si se trata de un método extrapolado (*JOR*, *SOR*) como de un método de sobrerrelajación acelerada (*AOR*), debemos intentar encontrar el valor óptimo del factor de relajación ω en el primer caso o de los dos parámetros ω y μ en el segundo caso. La determinación de estos parámetros no es sencilla y se basa en la mayoría de los casos en procedimientos heurísticos.

1.4 Métodos en dos etapas

Los métodos iterativos para resolver sistemas de ecuaciones lineales requieren de la resolución, en cada etapa, de un sistema de ecuaciones más simple. Cuando este sistema se aproxima a su vez por un método iterativo, el método global se llama *método iterativo en dos etapas*. A continuación vamos a describir más detalladamente dicho método.

Consideramos el sistema lineal $Ax = b$ y la partición $A = M - N$. Asociado a dicha partición tenemos el esquema iterativo

$$Mx^{(l+1)} = Nx^{(l)} + b, \quad l = 0, 1, 2, \dots, \quad (1.21)$$

donde $\mathbf{x}^{(0)}$ es un vector inicial arbitrario.

Los métodos en dos etapas, también llamados métodos "inner - outer", consisten en aproximar el sistema (1.21) de forma iterativa. Así, si consideramos la partición $M = F - G$ y suponemos que se realizan s iteraciones en el sistema (1.21), que llamaremos *internas*, entonces el esquema que resulta es el siguiente

$$\mathbf{x}^{(l+1)} = (F^{-1}G)^s \mathbf{x}^{(l)} + \sum_{j=0}^{s-1} (F^{-1}G)^j F^{-1}(N\mathbf{x}^{(l)} + \mathbf{b}), \quad l = 0, 1, 2, \dots \quad (1.22)$$

Podemos distinguir entre el método *estacionario* en dos etapas, en el que el número de iteraciones internas s permanece fijo en cada una de las etapas *externas*, y el método *no estacionario*, en el que el número de iteraciones internas $s(l)$ puede variar para cada etapa externa l . Hacemos notar que Lanzkron, Rose y Szyld en [53] llaman a los métodos no estacionarios en dos etapas, métodos *dinámicos* en dos etapas. En este caso el esquema iterativo es

$$\mathbf{x}^{(l+1)} = (F^{-1}G)^{s(l)} \mathbf{x}^{(l)} + \sum_{j=0}^{s(l)-1} (F^{-1}G)^j F^{-1}(N\mathbf{x}^{(l)} + \mathbf{b}), \quad l = 0, 1, 2, \dots \quad (1.23)$$

En un principio, los procesos iterativos en dos etapas se desarrollaron para la resolución de cierta clase de sistemas de ecuaciones lineales que surgían a partir de la discretización de problemas de valores frontera elípticos. Estos métodos se usan, en particular, para la resolución de aproximaciones a ecuaciones en diferencias parciales simultáneas (ver Smith [74]). Tales procedimientos son de interés puesto que pueden ser frecuentemente extendidos a la resolución de ecuaciones no lineales (Douglas [20], Dupont [21], D'Yakonov [82]). La convergencia de estos procesos se probó para problemas específicos y se estimó el número de operaciones necesarias para reducir la norma del error inicial en cada iteración.

Posteriormente, en 1973, los métodos en dos etapas serían estudiados por Nichols en [58], dando resultados de convergencia bajo condiciones bastante

generales. Concretamente, se demuestra la convergencia del método no estacionario bajo las hipótesis de que el número de iteraciones internas $s(l)$, para cada iteración externa l , verifique que $s(l) \geq S$, para un cierto S , y las iteraciones interna y externa sean ambas convergentes, es decir cumplan las condiciones razonables $\rho(M^{-1}N) < 1$ y $\rho(F^{-1}G) < 1$.

En este trabajo, también se demuestra la convergencia bajo condiciones más complejas para el método estacionario en dos etapas, es decir cuando $s(l) = s$. Las condiciones que Nichols plantea son las siguientes:

Los valores propios de $C = M^{-1}N$, $\lambda_i(C)$, y los de la matriz de iteración del proceso interno T_s , $\lambda_i(T_s)$ verifican

$$-1 < \lambda_i(C) < 1, \quad -1 < \lambda_i(T_s) < 1.$$

El valor de s es elegido de forma que $\rho(T_s) < \frac{1-\rho(C)}{1+\rho(C)}$. Y además se ha de cumplir una de las siguientes condiciones:

- (i) T_s y C son matrices reales simétricas.
- (ii) Existe una matriz P tal que PCP^{-1} y PT_sP^{-1} son ambas reales y simétricas.
- (iii) Existe una matriz P tal que PCP^{-1} es simétrica y T_s es simétrica y conmuta con P .

Golub y Overton en [40] y [41] estudian la convergencia de los métodos en dos etapas para el caso en el que la iteración externa se resuelve por el método de Richardson o Chebyshev. Para sistemas de ecuaciones no lineales, han sido ampliamente estudiados los métodos con iteración externa no lineal y con iteración interna lineal, además han sido aplicados a diferentes áreas de la

ciencia e ingeniería; por ejemplo, vease Bank y Rose [5], Dembo, Eisenstat y Steihaug [17].

Lanzkron, Rose y Szyld en [53] dan condiciones sobre las particiones externas e internas que aseguran la convergencia para cualquier número de iteraciones internas, del método en dos etapas, tanto estacionario como no estacionario. Las condiciones dadas hacen referencia a particiones regulares y débilmente regulares, y sus resultados sólo son aplicables a matrices monótonas, que como ya se ha dicho en la Sección 1.2 son matrices no singulares cuya inversa es no negativa. En concreto exigen que la partición externa sea regular y la partición interna débilmente regular. Los conceptos de partición regular y débilmente regular fueron introducidos en las Definiciones 7 y 8 respectivamente, de la Sección 1.2.

Teorema 7. *Sea $A = M - N$ una partición convergente y regular, y sea $M = F - G$ una partición convergente y débilmente regular. Entonces*

- (i) *El método en dos etapas estacionario definido en la expresión (1.22) converge para $s \geq 1$.*
- (ii) *El método en dos etapas no estacionario definido en la expresión (1.23) converge para $s(l) \geq 1$, $l = 0, 1, \dots$*

Bajo estas mismas condiciones demuestran que el radio espectral de la matriz de iteración global decrece cuando el número de iteraciones internas crece, es decir, dicho radio espectral es una función monótona decreciente de s , lo que se puede interpretar, obviamente, como que al aumentar las iteraciones internas mejora la correspondiente aproximación externa. Este resultado es intuitivo, pero sin embargo si las condiciones del Teorema 7 no se satisfacen, el resultado

puede no cumplirse, como puede apreciarse en diversos ejemplos que fueron objeto de estudio en [53].

Frommer y Szyld en [31] demuestran que si las particiones interna y externa son convergentes, el método no estacionario en dos etapas converge si

$$\lim_{l \rightarrow \infty} s(l) = \infty. \quad (1.24)$$

En particular, comprueban la convergencia para cualquier sucesión $\{s(l)\}_{l=1}^{\infty}$ satisfaciendo (1.24) y analizan la convergencia en el caso en que (1.24) no se cumple pero $s(l)$ es suficientemente grande. Además dan demostraciones alternativas de los resultados de convergencia vistos en [53], y usan estas demostraciones para extender los resultados de convergencia a particiones de H -matrices.

Definición 11. Sea $A = [a_{ij}]_{1 \leq i, j \leq n} \in \mathcal{M}_n$. Se define la *matriz comparación* de A y se denota por $\langle A \rangle$, como

$$\langle A \rangle = [b_{ij}] : b_{ij} = \begin{cases} |a_{ij}| & i = j \\ -|a_{ij}| & i \neq j \end{cases} \quad 1 \leq i, j \leq n.$$

Definición 12. Diremos que $A = [a_{ij}]_{1 \leq i, j \leq n} \in \mathcal{M}_n$ es un M -matriz si es no singular con $a_{ij} \leq 0$ para $i \neq j$ y $A^{-1} \geq O$. La llamaremos H -matriz, si $\langle A \rangle$ es una M -matriz.

Para introducir los resultados de convergencia para el método en dos etapas, demostrados por Fromer y Szyld [31], damos previamente las siguientes definiciones.

Definición 13. Diremos que la partición $A = M - N$ es H -partición si $\langle M \rangle - |N|$ es una M -matriz.

Definición 14. Diremos que la partición $A = M - N$ es H -compatible si $\langle A \rangle = \langle M \rangle - |N|$.

Teorema 8. Sea $A = M - N$ una H -partición, y sea $M = F - G$ una partición H -compatible. Entonces

- (i) El método en dos etapas estacionario definido en la expresión (1.22) converge para $s \geq 1$.
- (ii) El método en dos etapas no estacionario definido en la expresión (1.29) converge para $s(l) \geq 1$, $l = 0, 1, \dots$

Es conocido, (ver Frommer y Szyld [31]) que si la partición $A = M - N$ es una H -partición, entonces A y M son H -matrices, por tanto el Teorema 8 extiende los resultados de convergencia, de los métodos en dos etapas, a una clase de matrices no necesariamente monótonas, a las H -matrices.

Este tipo de matrices forman una clase bastante importante, según se refleja en el siguiente teorema que puede verse en Frommer y Mayer [29].

Teorema 9. Supongamos que la matriz A cumple una de las siguientes condiciones.

- (i) A es una M -matriz.
- (ii) A es estricta o irreduciblemente diagonal dominante.
- (iii) $\langle A \rangle$ es simétrica y definida positiva.

Entonces, A es una H -matriz.



Universitat d'Alacant
Universidad de Alicante

Capítulo 2

Métodos iterativos paralelos

2.1 Introducción a los multiprocesadores

Es de destacar la demanda de computadores con gran potencia de cálculo que hoy en día se está exigiendo, tanto a nivel científico como industrial. Como cabe esperar, esta demanda actual irá incrementándose en un futuro, sobre todo en aplicaciones de simulación y de tiempo real.

La velocidad en la computación tiene unos límites y restricciones lógicos y tecnológicos a los que poco a poco se va llegando. Las soluciones que se plantean ante esta demanda de computación se pueden concretar en tres tipos, soluciones de tipo software, soluciones de tipo hardware y paralelismo.

Las soluciones de tipo software consisten en realizar mejoras en los algoritmos para que estos se ejecuten requiriendo el menor tiempo posible. Pero los algoritmos poseen una complejidad y necesidad de recursos intrínsecas que

limitan este tipo de soluciones.

Las soluciones de tipo hardware consisten en mejoras en la tecnología de computadores, mejoras que deriven, por ejemplo, en una mayor rapidez en la ejecución de instrucciones o en la realización de operaciones de entrada y salida, conseguidas mediante la construcción de dispositivos electrónicos más rápidos.

Otra solución, que además es compatible con las anteriores, es el paralelismo, el cual consiste en replicar unidades de tratamiento de información con el objetivo de repartir tareas entre las mismas, espacial o temporalmente, y realizar la ejecución del programa en cuestión, en un tiempo inferior. A las arquitecturas de este tipo de computadores se las denomina *arquitecturas paralelas*. Dedicaremos la Subsección 2.1.1 al estudio de este tipo de arquitecturas, prestándole mayor atención a las arquitecturas tipo *MIMD* o *multiprocesadores*, ya que es en este tipo de arquitectura, en la que se realizarán los experimentos numéricos de esta memoria.

2.1.1 Tipos de arquitecturas paralelas

Dentro de las arquitecturas paralelas, cabe distinguir entre el *paralelismo temporal*, que es el que se da cuando se realizan un conjunto de operaciones de forma solapada en el tiempo, y el *paralelismo espacial*, que se consigue replicando ciertas unidades funcionales. En cuanto a los distintos tipos de arquitecturas paralelas destacan las arquitecturas segmentadas o pipeline, arquitecturas *SIMD*, procesadores sistólicos y arquitecturas *MIMD*.

Analizamos brevemente estos tipos.

- **Arquitecturas segmentadas o pipeline**

La idea básica de estas arquitecturas es la división de una unidad de proceso en partes o segmentos, dotándolos de registros capaces de almacenar los resultados intermedios. De esta forma si se tienen que procesar distintos conjuntos de datos, es posible hacerlo solapadamente en el tiempo, consiguiéndose así un paralelismo temporal.

El funcionamiento de los distintos segmentos es síncrono, el ciclo de reloj para la unidad de pipeline vendrá determinado por el segmento más lento.

Este tipo de arquitectura es útil cuando hay que procesar conjuntos homogéneos e independientes de datos, como pueden ser vectores. De aquí que en los computadores vectoriales se utilicen unidades segmentadas.

Aunque este tipo de arquitectura ha alcanzado su techo de perfeccionamiento, cabe destacar su compatibilidad con otro tipo de arquitecturas paralelas, mejorando así las prestaciones.

- **Arquitecturas tipo *SIMD***

Las siglas corresponden a *Single Instruction Multiple Data*. Este tipo de arquitectura corresponde a la idea de una única instrucción ejecutándose sobre distintos datos. Los computadores con este tipo de arquitectura suelen estar formados por varios procesadores pero con una única unidad de control. Todos los procesadores ejecutan la misma instrucción, bajo las órdenes de la unidad de control, pero cada uno de ellos sobre sus propios datos. Esto proporciona obviamente un funcionamiento síncrono del sistema. Cabe destacar la importancia de este tipo de arquitecturas para el

diseño de computadores con un fin muy específico, sin embargo no parece muy adecuada como solución general.

- **Arquitecturas sistólicas**

Este tipo de arquitectura permite disponer de un gran número de procesadores que pueden funcionar en paralelo con un coste relativamente bajo. Con un funcionamiento en principio síncrono puede considerarse como el límite al que tiende una arquitectura *SIMD* cuando los elementos de proceso tienden a la simplicidad máxima y la memoria local a desaparecer. Este tipo de arquitectura presenta gran utilidad en problemas cuya velocidad de ejecución viene limitada por el cálculo y no por las entradas-salidas. Aunque se ha implementado en pocas máquinas, tiene gran aceptación como coprocesador.

- **Arquitecturas tipo *MIMD***

La siglas corresponden a *Multiple Instructions Multiple Data* y describen la idea de un computador en el que distintas instrucciones se ejecutan sobre distintos conjuntos de datos. En este tipo de máquina cada elemento de proceso ejecuta su propio juego de instrucciones sobre los datos que eventualmente contiene. Existen por tanto, varios elementos de proceso y cada uno con su propia unidad de control, con lo que el funcionamiento del sistema es asíncrono.

A este tipo de máquinas se le denomina comúnmente *multiprocesador*. Según su configuración existen dos estructuras básicas:

- **Multiprocesador con memoria distribuida**

En este tipo de multiprocesador cada elemento de proceso tiene su propia memoria, inaccesible a los demás procesadores, donde almacena sus propios datos, no existiendo una memoria global común. La comunicación entre los distintos procesadores debe realizarse mediante *paso de mensajes*. Este paso de mensajes conlleva un retraso en la ejecución.

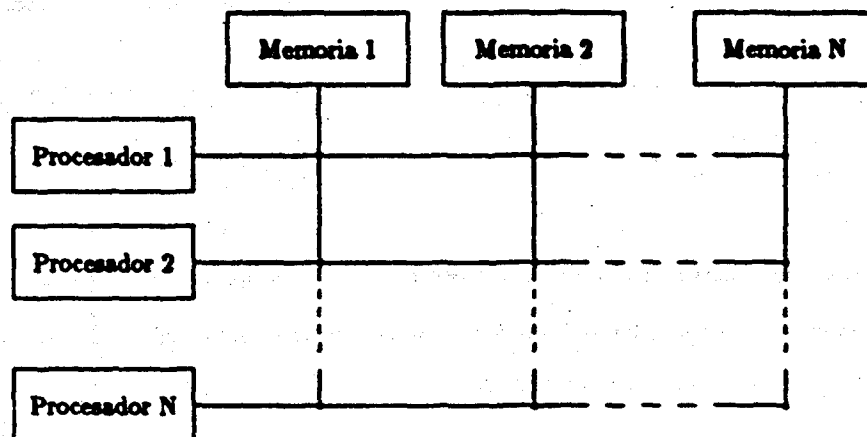
- **Multiprocesador con memoria compartida**

En estos computadores paralelos, todos los elementos de proceso tienen acceso a una memoria común, si bien independientemente, cada uno de ellos puede poseer una pequeña memoria local para código y resultados intermedios. La comunicación entre los distintos procesadores se realiza a través de la memoria común. La principal ventaja de este sistema es que permite comunicaciones muy rápidas entre los procesadores. Sin embargo pueden producirse conflictos en el acceso a datos. En algunas ocasiones un procesador no deberá poder acceder a un dato hasta que este no haya sido actualizado por otro, lo que debe preverse. También se producirán *tiempos de contención* cuando un procesador deba esperar a que otro haya leído o actualizado un dato. Este tipo de problema se incrementa al aumentar el número de procesadores.

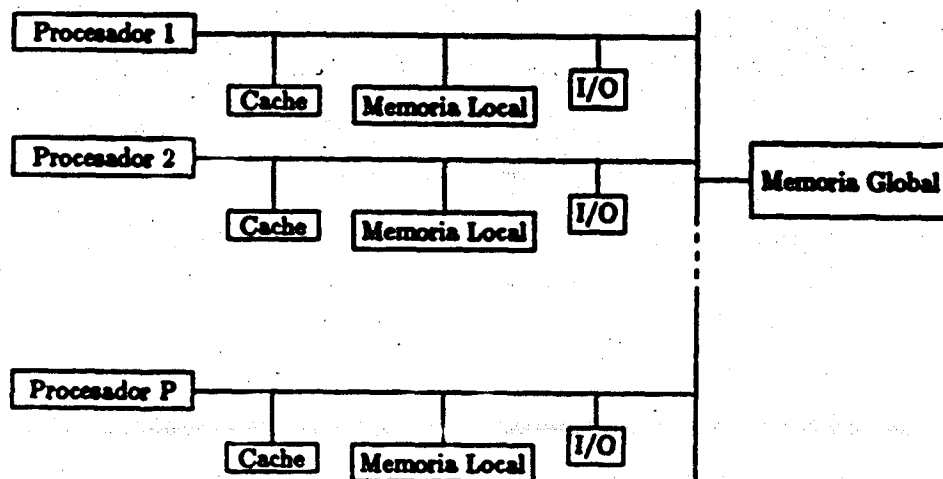
Un aspecto importante en un multiprocesador, ya sea de memoria compartida o distribuida, es la *topología de la red de interconexión*, es decir, la forma de conectarse los distintos procesadores para establecer la comunicación entre ellos. Una red de interconexión debería lograr conectar el mayor número de nodos posibles, con pocos enlaces por nodo, un diámetro pequeño y pocos con-

Conflictos de acceso en la red. Entre las redes más importantes se encuentran las que describimos a continuación.

Red crossbar: posee elementos de conmutación mediante los cuales se conectan todos los nodos entre sí.



Topología de bus: Todos los elementos están conectados a un bus rápido.



Topología de anillo: Cada elemento está conectado a los vecinos más próximos por un enlace unidireccional o bidireccional. Un elemento puede leer de un vecino mientras escribe en el otro vecino.

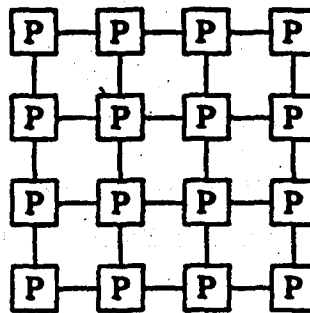


Anillo unidireccional

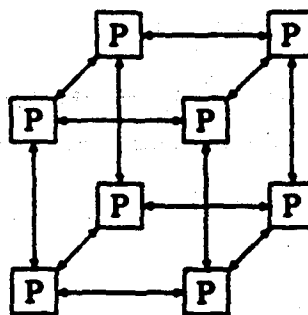


Anillo bidireccional

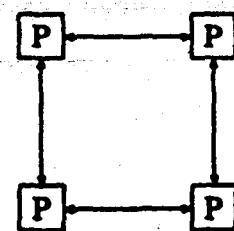
Topología de malla: Cada nodo está conectado a sus cuatro vecinos más próximos. Se dice de malla abierta, cuando además los enlaces de la periferia están anulados.



Topología de hipercubo: Este sistema de interconexión puede usarse para conectar $N = 2^n$ procesadores. Cada procesador correspondería a un vértice de un hipercubo n -dimensional y estaría conectado a otro si en el hipercubo hubiera una arista entre los correspondientes vértices. Con un número no demasiado elevado de conexiones se consigue un diámetro reducido, aunque en cualquier caso la dimensión n del hipercubo no puede ser muy grande. La ventaja fundamental del hipercubo es que permite conectar un gran número de procesadores (2^n) con un diámetro pequeño n y con pocas conexiones. Otra ventaja del hipercubo es que contiene otras estructuras, en particular al anillo.

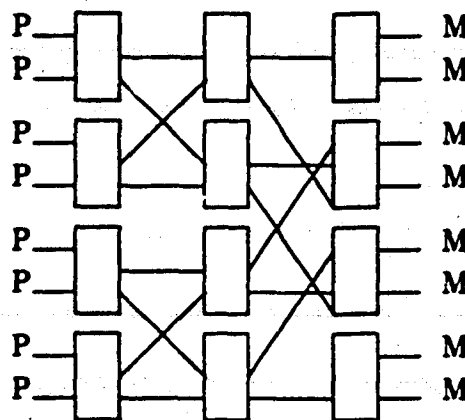


Hipercubo 3-dimensional



Hipercubo 2-dimensional

Red Multifase: Si se desea conectar $N = 2^n$ procesadores, puede hacerse en n fases, cada una de ellas con N conexiones. Este esquema puede utilizarse tanto para sistemas con memoria distribuida, como para sistemas con memoria compartida, según pongamos a la derecha procesadores o memorias.



Los experimentos numéricos que se expondrán en esta memoria, han sido implementados en un multiprocesador de memoria compartida Alliant FX/80. Resumimos, a continuación, sus características más importantes. Este multiprocesador posee dos tipos de procesadores, los procesadores interactivos o IP's y los elementos computacionales o CE's.

Las funciones principales de los IP's son la gestión y control de la periferia, gestión de tareas para el sistema operativo y gestión de los procesos de usuario

que no requieren cálculos (ediciones, compilaciones, ...). Poseen un juego de instrucciones no muy sofisticadas desde el punto de vista matemático (compatible con el motorola MC 68020). El Alliant FX/80 posee dos procesadores de este tipo.

Por otra parte, la función principal de los CE's es la ejecución de todo tipo de procesos usuarios, principalmente los que requieren cálculo matemático muy intenso. Poseen un juego de instrucciones básico (compatible MC 68020), instrucciones para operaciones en coma flotante, instrucciones vectoriales e instrucciones para la gestión de concurrencia (sincronización y comunicación). El Alliant FX/80 posee 8 procesadores de este tipo.

La red de interconexión es de tipo bus. El sistema de memoria soporta hasta 256 Mb de memoria central; además, este sistema dispone de dos tipos de memoria cache:

- IPC (Interactive Processor Cache). 1 banco de 32 Kb.
- CPC (Computational Processor Cache). 2 bancos de 256 Kb cada uno.

Como ya hemos comentado, los resultados numéricos que desarrollamos en las Secciones 3.4, 4.4 y 5.4, se han realizado sobre este multiprocesador, aprovechando las posibilidades de vectorización y paralelismo que ofrece el compilador asociado al lenguaje FX/FORTRAN.

Para obtener los distintos modos de ejecución, que nos interesan, hemos utilizado diferentes directivas de compilación, que detallamos a continuación.

- **Compilación escalar o secuencial:** Las operaciones se realizan secuencialmente. El procesador FX/FORTRAN realiza tantas instrucciones como le permiten sus "pipelines".

- **Compilación vectorial:** Las operaciones se realizan en grupos de 32 elementos (o menos, si no están disponibles los 32) por instrucciones vectoriales del hardware.
- **Compilación concurrente (escalar concurrente):** Las operaciones se realizan concurrentemente sobre un número predeterminado de elementos computacionales o CE's.
- **Compilación vectorial-concurrente:** Las operaciones se realizan en grupos de como mucho 32 elementos, concurrentemente sobre un número predefinido de CE's.

2.1.2 Medidas de paralelismo

En esta sección vamos a definir aquellos parámetros que nos van a ser de utilidad para evaluar las prestaciones de los algoritmos en paralelo. Estos son el incremento de velocidad o speed-up y la eficiencia. Para ello, vamos a suponer que disponemos de r procesadores.

Definición 1. Se llama *incremento de velocidad* (o *speed-up*) de un algoritmo paralelo a

$$S_r = \frac{\text{Tiempo de ejecución en un sólo procesador}}{\text{Tiempo de ejecución en } r \text{ procesadores}}. \quad (2.1)$$

Obviamente $S_r \leq r$, siendo de esperar que en el caso utópico en que el paralelismo fuera perfecto se obtendría $S_r = r$.

Esta definición compara el mismo algoritmo utilizando uno o r procesadores, sin embargo esto tiene un inconveniente, ya que dicho algoritmo paralelo no tiene porque ser el más eficiente cuando se ejecuta en secuencial. Esto nos lleva a la introducción de una nueva medida de paralelismo,

Definición 2. El *incremento de velocidad* o *speed-up* de un algoritmo paralelo respecto al mejor algoritmo secuencial es

$$S_r = \frac{\text{Tiempo de ejecución en un sólo procesador del algoritmo secuencial más rápido}}{\text{Tiempo de ejecución en } r \text{ procesadores del algoritmo paralelo}} \quad (2.2)$$

Relacionada con las anteriores medidas están las de eficiencia, que miden el grado de utilización de los procesadores del sistema al ejecutar en él un algoritmo. Para un sistema con r procesadores damos la siguiente definición.

Definición 3. Se llama *eficiencia* de un algoritmo paralelo, respecto a sí mismo a

$$E_r = \frac{S_r}{r},$$

siendo la eficiencia respecto al mejor algoritmo secuencial

$$E_r' = \frac{S_r'}{r},$$

Obviamente $E_r' \leq E_r \leq 1$. El objetivo es conseguir la mayor eficiencia (o incremento de velocidad) posible, aunque difícilmente se podrá obtener el óptimo $E_r' = 1$.

Hacemos notar, que para obtener los parámetros definidos en esta sección es necesario realizar el experimento numérico, pues estamos hablando de tiempos

reales de ejecución. Por otra parte no siempre va a ser posible determinar el mejor algoritmo secuencial en términos absolutos.

Para finalizar esta sección, decir que para maximizar la eficiencia de un algoritmo paralelo, hay que tener muy en cuenta el *equilibrio de la carga*, es decir, la distribución de las tareas entre los distintos procesadores de forma que todos ellos tengan una cantidad de trabajo similar, más aún, debe procurarse que este reparto equitativo se realice entre todos los puntos de sincronización a fin de evitar que algunos procesadores se mantengan inactivos o *perezosos*. Esto supone que el problema de asignación de tareas a los distintos procesadores sea clave a la hora de mejorar la eficiencia. Esta asignación en los multiprocesadores de memoria distribuida suele ser estática, es decir, que se ha de hacer antes de comenzar con la ejecución del algoritmo, sin embargo en los de memoria compartida suele ser dinámica, lo que permite que dependiendo de cómo va desarrollándose la ejecución se asignen las tareas a los distintos procesadores.

2.2 Antecedentes y estado actual

En esta sección vamos a estudiar métodos iterativos en paralelo, para la resolución del sistema de ecuaciones $Ax = b$. En el Capítulo 1 hemos dado condiciones suficientes o necesarias y suficientes, para que un método iterativo secuencial converja. Estos métodos se pueden utilizar, sin demasiada dificultad, en un multiprocesador. Para esto, distribuimos la carga de trabajo, mediante la asignación de incógnitas, entre los distintos procesadores. Esta asignación se

puede realizar de cuatro formas diferentes:

- de forma directa,
- mediante reordenamiento de incógnitas,
- utilizando descomposición en bloques, y
- con la introducción de las multiparticiones.

La asignación se puede hacer de forma directa cuando el método permite que los procesadores trabajen independientemente, como ocurre en el método de Jacobi, que puede ejecutarse en paralelo, haciendo que cada procesador actualice sus incógnitas y después, en paralelo también, envíe las variables necesarias a los otros procesador. Si no se tuviese en cuenta el tiempo de comunicación se alcanzaría un speed-up teórico igual al número de procesadores, ya que en este caso, el grado de paralelismo es perfecto.

La reordenación adecuada de incógnitas, permite distribuir la carga entre procesadores de forma independiente, en algunos algoritmos.

Sin embargo la distribución de un grupo de incógnitas a cada procesador induce una idea más general, que consiste en dividir la matriz A en bloques y asignar estos bloques a los distintos procesadores. Es decir, si la matriz A está dividida en bloques, con bloques diagonales invertibles de tamaño $n_i \times n_i$, de la siguiente forma

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{bmatrix},$$

se pueden construir los siguientes sub-sistemas que se asignan cada uno a un procesador

$$A_{ii}y_i^{(k)} = - \sum_{j=1, j \neq i}^r A_{ij}y_j^{(k-1)} + b_i, \quad i = 1, 2, \dots, r, \quad k = 1, 2, \dots, \quad (2.3)$$

donde y_i y b_i son bloques de los vectores y y b , respectivamente, de tamaño conforme con el tamaño de los correspondientes bloques de A .

La ecuación (2.3) corresponde al esquema iterativo denominado método de Jacobi por bloques (un estudio exhaustivo de métodos iterativos secuenciales por bloques se puede seguir en Young [83] y en Bertsekas y Tsitsiklis [8]). En determinados casos, los métodos por bloques son al menos tan rápidos como los métodos por filas, esto ocurre por ejemplo cuando la matriz de coeficientes es una M -matriz. Será interesante, por tanto, no sólo la elección del método utilizado para la resolución de los distintos subsistemas planteados en (2.3), sino la elección de los bloques de la matriz A . En White [80] puede verse un ejemplo numérico basado en la ecuación de Poisson discretizada mediante diferencias finitas de segundo orden y con la utilización de las técnicas de reordenación y descomposición en bloques.

La técnica de descomposición en bloques puede verse como un caso particular de la técnica de multipartición. Esta técnica fue introducida por O'Leary y White en [61].

Nos centraremos en la Subsección 2.1, en el estudio de los métodos basados en multiparticiones, ya que en ellos se fundamentan los resultados obtenidos en esta memoria.

2.2.1 Métodos iterativos basados en la técnica de multipartición

Sea A la matriz del sistema lineal $Ax = b$; en el Capítulo 1 se definió el concepto de partición de una matriz $A = M - N$. Esto permitió construir el esquema iterativo (1.8), que como recordamos aquí, se escribe como

$$x^{(k)} = M^{-1}Nx^{(k-1)} + M^{-1}b, \quad k = 1, 2, \dots \quad (2.4)$$

Claramente, se podrían haber construido distintas particiones de la matriz A , y obtener para cada una de ellas un esquema iterativo diferente.

Bajo estas condiciones, O'Leary y White [61] introducen, en 1985, la siguiente técnica de multipartición para la resolución de sistemas de ecuaciones lineales en paralelo.

Definición 4. Sea A una matriz cuadrada. Se llama *multipartición* de A al conjunto $\{M_i, N_i, E_i\}_{i=1}^r$ tal que

- (i) $A = M_i - N_i$, es una partición de A , para todo $i = 1, 2, \dots, r$.
- (ii) $0 \leq E_i \leq I$, E_i no negativa y diagonal, para todo $i = 1, 2, \dots, r$.
- (iii) $\sum_{i=1}^r E_i = I$.

O'Leary y White, a partir del concepto de multipartición, plantean el siguiente procedimiento iterativo

$$y^{(k)} = \sum_{i=1}^r E_i M_i^{-1} N_i y^{(k-1)} + \sum_{i=1}^r E_i M_i^{-1} b, \quad k = 1, 2, \dots, \quad (2.5)$$

que tiene la misma estructura que el esquema iterativo secuencial (2.4).

Suponen que se dispone de r procesadores conectados entre sí y que cada uno de ellos puede ejecutar una sucesión diferente de instrucciones sobre sus datos locales. En cada iteración el procesador i -ésimo calcula solamente las componentes del vector

$$M_i^{-1} N_i y^{(k-1)} + M_i^{-1} b,$$

que corresponden a los elementos no nulos de E_i . El procesador entonces pondera estas componentes para que sea capaz de distribuir el vector

$$E_i M_i^{-1} N_i y^{(k-1)} + E_i M_i^{-1} b$$

a uno de los procesadores, que denominaremos *central*, donde se suman las soluciones de cada uno de los esquemas iterativos asignados a los distintos procesadores, para así actualizar el vector y en la iteración k , es decir para obtener $y^{(k)}$.

Notar que si un elemento diagonal de la matriz E_i es nulo, entonces el procesador i -ésimo no necesita actualizar dicha componente. En consecuencia, si las matrices E_i son tales que si una de ellas tiene el j -ésimo elemento de la diagonal no nulo y todas las demás E_t , $t \neq i$ tienen su elemento diagonal j -ésimo nulo, se puede asignar la fila j de A sólo al procesador i -ésimo. Si este razonamiento se piensa con bloques de filas, sólo se asignaría el correspondiente bloque a un único procesador. Así, con una única partición se puede construir una multipartición y repartir el trabajo entre procesadores, mediante el cálculo de bloques de componentes del vector $y^{(k)}$.

Si denotamos

$$H = \sum_{i=1}^r E_i M_i^{-1} N_i \quad \text{y} \quad C = \sum_{i=1}^r E_i M_i^{-1} b,$$

la ecuación (2.5) se puede escribir de la siguiente manera

$$y^{(k)} = Hy^{(k-1)} + Cb, \quad k = 1, 2, \dots, \quad (2.6)$$

donde H es la matriz de iteración.

Al estudiar este método iterativo, podríamos pensar que si las r particiones de la multipartición son convergentes entonces el esquema (2.5) o equivalentemente (2.6) converge. Sin embargo, esto no es cierto como se refleja en el siguiente ejemplo (ver O'Leary y White [61]).

Ejemplo 1. Sea

$$A = \begin{bmatrix} \frac{3}{4} & 0 \\ 0 & \frac{3}{4} \end{bmatrix}$$

y consideremos las particiones

$$A = M_1 - N_1 = \begin{bmatrix} \frac{1}{2} & -1 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} -\frac{1}{4} & -1 \\ 1 & \frac{13}{4} \end{bmatrix}$$

y

$$A = M_2 - N_2 = \begin{bmatrix} 4 & 1 \\ -1 & \frac{1}{2} \end{bmatrix} - \begin{bmatrix} \frac{13}{4} & 1 \\ -1 & -\frac{1}{4} \end{bmatrix}.$$

Entonces

$$M_1^{-1}N_1 = \begin{bmatrix} 0 & -0.25 \\ 0.25 & 0.875 \end{bmatrix}$$

y

$$M_2^{-1}N_2 = \begin{bmatrix} 0.875 & 0.25 \\ -0.25 & 0 \end{bmatrix}.$$

Ambas matrices son convergentes ya que su radio espectral es 0.7965.

Si consideramos las matrices

$$E_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

entonces la matriz de iteración de la multipartición es

$$H = E_1 M_1^{-1} N_1 + E_2 M_2^{-1} N_2 = \begin{bmatrix} 0.875 & 0.25 \\ 0.25 & 0.875 \end{bmatrix},$$

cuyo radio espectral es $\rho(H) = 1.125$ y por tanto no hay convergencia.

Sin embargo si las matrices E_1 y E_2 se definen

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad E_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

entonces la matriz H resultante tiene radio espectral 0.25. Luego en este caso hay convergencia.

O'Leary y White [61], investigan también clases de matrices A y tipos de multiparticiones que dan lugar a que el procedimiento iterativo (2.5) o (2.6) converja, demostrando los siguientes resultados.

Teorema 1. (a) Sea A una matriz monótona, y sea $\{M_i, N_i, E_i\}_{i=1}^r$ una multipartición de A , tal que todas las particiones $A = M_i - N_i$ son débilmente regulares. Entonces el esquema iterativo (2.5) es convergente.

(b) Sea $\{M_i, N_i, E_i\}_{i=1}^r$ una multipartición de la matriz A . Si se cumple que $\|M_i^{-1} N_i\|_{\infty} < 1$ para todo $i = 1, 2, \dots, r$, entonces el esquema iterativo (2.5) es convergente.

La técnica de multipartición ha sido ampliamente estudiada. Así, O'Leary y White [61], y posteriormente, en 1990, White en [80] estudian la convergencia del esquema (2.5) cuando la matriz A es simétrica y definida positiva. En [80] también se dan resultados numéricos de la implementación de este método sobre un multiprocesador de memoria compartida; la conclusión más visible es que disminuye el número de iteraciones cuando aumenta el solapamiento, pero sin embargo el tiempo no disminuye sensiblemente. En 1987, Neumann y Plemmons [57] estudian con detalle la convergencia del método iterativo basado en una multipartición de A , cuando ésta es una matriz monótona. Dicho estudio se basa en una norma monótona (ver expresión (1.4) del Capítulo 1), usando el vector Perron, de una matriz no negativa.

Teorema 2. *Sea A una matriz no negativa e irreducible. Entonces*

- (i) $\rho(A)$ es un valor propio de A .
- (ii) $\rho(A)$ aumenta cuando los elementos de A aumentan.
- (iii) $\rho(A)$ es un valor propio simple.
- (iv) Existe un vector propio u positivo asociado a $\rho(A)$, es decir $Au = \rho(A)u$, con $u > 0$.

Este resultado se debe a Frobenius y es una generalización del que estableció Perron en 1907 para matrices positivas. Al vector propio u se le conoce con el nombre de *vector Perron* de A .

Teorema 3. (Lema 2.2 de [57]). *Sea A una matriz monótona. Sean*

$$\{M_i, N_i, E_i\}_{i=1}^r \quad \text{y} \quad \{\widehat{M}_i, \widehat{N}_i, \widehat{E}_i\}_{i=1}^r$$

dos multiparticiones de A tales que $\widehat{M}_i^{-1} \geq M_i^{-1}$, para todo $i = 1, 2, \dots, r$. Entonces si $Au_i \geq 0$, para todo i , donde u_i es el vector Perron positivo de $M_i^{-1}N_i$, $i = 1, 2, \dots, r$, se cumple que $\rho(\widehat{H}) \leq \rho(H)$.

Nótese que en el Teorema 3 no se ha impuesto ninguna condición sobre las matrices E y \widehat{E}_i , $i = 1, 2, \dots, r$. Ello sugiere que el radio de convergencia de un esquema iterativo basado en multiparticiones de una matriz monótona, es una propiedad de las particiones de la multipartición y no de las matrices E_i . Anteriormente hemos discutido cómo se puede distribuir el trabajo entre procesadores mediante las matrices E_i . Por tanto, se puede decir que el radio de convergencia no depende de la distribución del trabajo entre procesadores y en consecuencia se puede equilibrar la carga con la sola idea de minimizar el tiempo de cálculo de una iteración.

En este mismo artículo, [57] Neumann y Plemmons dan especial atención al caso en que las particiones están basadas en las clásicas particiones de Jacobi y Gauss Seidel.

Suponen, sin pérdida de generalidad, que la matriz de coeficientes A se puede escribir como

$$A = I - L - U,$$

donde L y U son matrices estrictamente triangular inferior y superior respectivamente. Entonces construyen las siguientes matrices iteración de los métodos de

$$\text{Jacobi} \quad J = L + U \quad (2.7)$$

$$\text{Gauss-Seidel} \quad \mathcal{L} = (I - L)^{-1}U. \quad (2.8)$$

Sean L_i y U_i matrices que satisfacen las condiciones

$$0 \leq L_i \leq L, \quad 0 \leq U_i \leq U, \quad i = 1, 2, \dots, r,$$

entonces se pueden construir las multiparticiones siguientes

$$\text{Jacobi} \quad \{(I - L_i), (L + U - L_i), E_i\}_{i=1}^r \quad (2.9)$$

$$\text{Gauss-Seidel} \quad \{(I - L_i - U_i), (L + U - L_i - U_i), E_i\}_{i=1}^r. \quad (2.10)$$

Con estas multiparticiones obtienen el siguiente teorema de comparación (ver Neumann y Plemmons [57]).

Teorema 4. *Sea $A = I - L - U$ una M -matriz invertible e irreducible. Entonces el radio de convergencia del esquema iterativo asociado a las multiparticiones (2.9)-(2.10) es superior o igual al radio de convergencia del método de Jacobi.*

Otros resultados de comparación de radios de convergencia con multiparticiones se pueden estudiar en Elsner [22].

Es conocido, que una forma de acelerar la convergencia de un método iterativo es usar un factor de relajación, como comentamos en la Sección 1.3. En [61], O'Leary y White construyen algunos ejemplos de multiparticiones de matrices convergentes, discutiendo su uso sobre computadores paralelos, uno de estos ejemplos corresponde a un método de multipartición relajado en el que se utiliza el parámetro de relajación ω como en el método de Jacobi relajado. Este método está definido por la siguiente iteración

$$\mathbf{x}^{(l+1)} = \omega \sum_{i=1}^r E_i M_i^{-1} (N_i \mathbf{x}^{(l)} + \mathbf{b}) + (1 - \omega) \mathbf{x}^{(l)}, \quad l = 0, 1, \dots \quad (2.11)$$

Claramente el esquema iterativo (2.11) puede escribirse de la forma $x^{(l+1)} = Hx^{(l)} + c$, $l = 0, 1, \dots$, con

$$H = \omega \sum_{i=1}^r E_i M_i^{-1} N_i + (1 - \omega)I, \quad l = 0, 1, \dots$$

Prommer y Mayer en [29] demuestran la convergencia del esquema (2.11) cuando A es una H -matriz (ver Definición 1.2 de la Sección 1.4). La condición que exigen a la multipartición de A , para dicha convergencia, es que $\text{diag}(M_i) = \text{diag}(A)$ y $(A) = (M_i) - |N_i|$, $i = 1, 2, \dots, r$. Denotando $\text{diag}(A) = D$ y $A = D - B$ obtenemos el siguiente teorema.

Teorema 5. *Sea A una H -matriz, y sea el conjunto $\{M_i, N_i, E_i\}_{i=1}^r$ una multipartición de A con $\text{diag}(M_i) = D$ y $(A) = (M_i) - |N_i|$, $i = 1, 2, \dots, r$. Si $0 < \omega < \frac{1}{1 + \sqrt{2} \|B\|}$ entonces el método de multipartición relajado (2.11) converge para cualquier vector inicial $x^{(0)}$.*

Otra posibilidad, estudiada por estos mismos autores consiste en suponer que $A = D - L_k - V_k$, $k = 1, 2, \dots, r$ donde D es la parte diagonal de A , L_k es estrictamente triangular inferior, y V_k es tal que $A = D - L_k - V_k$ (en general V_k no es triangular superior). Supongamos que D es no singular y denotemos los elementos de A , D , L_k y V_k por a_{ij} , d_{ij} , l_{ij}^k y v_{ij}^k respectivamente. Consideremos la multipartición $\{D - L_k, V_k, E_k\}_{k=1}^r$. El esquema que plantean es el siguiente

$$x^{(l+1)} = \sum_{k=1}^r E_k y^{(k)}, \quad l = 0, 1, \dots$$

donde

$$y_i^{(k)} = \frac{\omega}{2} \left(\sum_{j=1}^{i-1} l_{ij}^k x_j^{(l)} + \sum_{j=i}^n v_{ij}^k x_j^{(l)} + b_i \right) + (1 - \omega)x_i^{(l)}, \quad i = 1, 2, \dots, n. \quad (2.12)$$

Claramente los cálculos de $\mathbf{y}^{l,k}$ para $k = 1, 2, \dots, r$, son independientes y se pueden realizar en paralelo.

En este caso el esquema iterativo (2.12) se puede escribir como

$$\mathbf{z}^{(l+1)} = H\mathbf{z}^{(l)} + \mathbf{c}, \quad l = 0, 1, \dots,$$

con

$$H = \sum_{k=1}^r E_k (D - \omega L_k)^{-1} ((1 - \omega)D + \omega V_k).$$

La convergencia del esquema (2.12) se estudia en [29], también para H -matrices, probando la convergencia para los mismos parámetros de relajación ω que en el Teorema 5, y asumiendo que $\langle A \rangle = |D| - |L_k| - |V_k|$, $k = 1, 2, \dots, r$.

Teorema 6. *Sea A una H -matriz, y para $k = 1, 2, \dots, r$ sea L_k una matriz estrictamente triangular inferior. Definimos las matrices V_k tal que $A = D - L_k - V_k$. Asumimos que $\langle A \rangle = |D| - |L_k| - |V_k|$ para $k = 1, 2, \dots, r$. Entonces si $0 < \omega < \frac{1}{1 + \rho(|D|^{-1}|D|)}$ el esquema (2.12) converge a la solución del sistema $A\mathbf{z} = \mathbf{b}$.*

Hacemos notar que estos métodos estudiados por Frommer y Mayer [29], para $\omega = 1$ se reducen a los métodos de multipartición no relajados, por lo que obtenemos nuevos criterios de convergencia para el caso no relajado.

La condición de que A sea una H -matriz cubre varios casos interesantes, tales como el de M -matriz, estricta o irreduciblemente diagonal dominante y el caso en que (A) sea simétrica y definida positiva (ver Teorema 9 de la Sección 1.4).

Deren en [18] presenta una clase de algoritmos relajados basados en multiparticiones, llamados algoritmos de multipartición en paralelo AOR, para resolver

grandes sistemas de ecuaciones. Este tipo de algoritmo es una generalización del modelo (2.12). La convergencia de estos modelos se discute bajo la condición de que A sea una matriz monótona. Haremos un estudio más detallado de los esquemas (2.11) y (2.12) en las Secciones 3.1 y 4.1 respectivamente. Por otra parte en la Sección 4.1 introduciremos los algoritmos estudiados por Deren.

En 1990, Szyld y Jones [77] estudiaron la relación existente entre los métodos iterativos en dos etapas (ver Sección 1.4) y los basados en la técnica de multipartición, analizando también los radios de convergencia. Concluyen que el método de multipartición es asintóticamente más rápido.

2.2.2 Métodos caóticos basados en la técnica de multipartición

En la Subsección 2.2.1 se han estudiado diferentes procesos iterativos en paralelo para la resolución de un sistema de ecuaciones lineales.

En general los diferentes procesos iterativos estudiados resuelven sistemas o subsistemas del sistema de ecuaciones lineales original en cada procesador, actualizando en cada iteración global el vector solución. En estos casos, cada procesador obtiene solamente una solución de su subsistema asociado.

Puesto que los diferentes procesadores de un multiprocesador *MIMD* son independientes, se puede pensar de forma natural en producir un esquema iterativo tal que cada procesador pueda actualizar más de una vez su solución, en cada iteración global. En este caso se obtiene una sucesión de vectores diferente de las obtenidas con los métodos de la Subsección 2.2.1 y por tanto se debe



realizar el estudio teórico de la convergencia de estas nuevas sucesiones.

Denominaremos a esta clase de esquemas, *métodos caóticos*, en el sentido en que cada procesador actualiza su solución un número variable de veces, dependiendo de la iteración global que se está realizando. La iteración global se construye a partir de los vectores de cada uno de los procesadores. La idea de los modelos caóticos surge ante el problema de equilibrar la carga de trabajo entre los distintos procesadores. Así, si suponemos que se ha de resolver un sistema en el que hay una gran diferencia entre los tamaños de los distintos bloques de filas que son asignados a distintos procesadores, podrá ocurrir que en un cierto instante algunos procesadores ya hayan actualizado su bloque de vector iterado, mientras que otros continúan trabajando. Parece entonces lógico, que para evitar esta situación los procesadores ociosos recalculen varias veces su bloque de vector. Esto lograría un equilibrio de la carga de trabajo, además de esperar una mejora en la velocidad de convergencia.

Dentro de los métodos caóticos, distinguiremos dos posibles casos, los *métodos caóticos síncronos* y los *métodos caóticos asíncronos*. En el primero, cada iteración global se construye a partir de las soluciones de todos los procesadores. Claramente se necesita una "sincronización" entre todos los procesadores para construir el vector iterado en cada iteración global. En el caso "asíncrono" cada iteración global se construye a partir de la solución de algún procesador, concretamente el que acaba de actualizar su solución. En este caso existe una gran libertad entre los procesadores, ya que no tienen que sincronizarse para construir la sucesión de vectores iterados.

En 1968, Bru, Elsner y Neumann, motivados por los trabajos de Ostrowski [64] en 1961 y de Chazan y Miranker [13] en 1969, estudian en [10] dos modelos

iterativos caóticos en paralelo, basados en la técnica de multipartición, para la resolución de grandes sistemas no singulares de ecuaciones lineales.

El primero de estos modelos, que corresponde a un esquema caótico síncrono, puede ser considerado como una generalización del método de Jacobi por bloques en paralelo. En el segundo modelo planteado en [10], cada procesador puede actualizar el vector iterado global, o recuperar algún subconjunto de las componentes de dicho vector iterado, que resida en un procesador *central*, en cualquier momento. Estamos hablando de un modelo caótico asíncrono.

Para la formulación matemática de estos modelos suponen $A = M_i - N_i$, $1 \leq i \leq r$, r particiones de la matriz A y para cada i se define el operador $F_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ como sigue,

$$F_i x = M_i^{-1} N_i x + M_i^{-1} b. \quad (2.13)$$

Esta expresión representa el trabajo de una iteración local y se realiza por cada procesador un número determinado de veces para intentar, de esta forma equilibrar la carga entre los distintos procesadores.

Además, para un entero no negativo q , consideramos la siguiente notación para la composición del operador F_i :

$$F_i^q = \begin{cases} F_i \circ F_i \circ \dots \circ F_i & q > 0 \\ I & q = 0. \end{cases} \quad (2.14)$$

Con la notación que acabamos de introducir, la versión matemática del modelo caótico síncrono es la siguiente.

Sean E_i , $1 \leq i \leq r$, r matrices diagonales no negativas verificando que $\sum_{i=1}^r E_i = I$, y sean $A = M_i - N_i$, r particiones. Iniciando con un vector

arbitrario $z^{(0)}$, se realiza la iteración

$$z^{(l)} = \sum_{i=1}^r E_i F_i^{q(l,i)} z^{(l-1)}, \quad l = 1, 2, \dots, \quad (2.15)$$

donde F_i son los operadores dados en (2.13) y $q(l, i)$ son enteros positivos que dependen de i , el procesador, y también de l , el índice de la etapa de iteración global. Obsérvese que $q(l, i)$ indica el número de iteraciones locales del procesador i en la iteración global l .

Tal y como se ha definido el esquema (2.15), no es necesario que cada procesador calcule todo el vector $z^{(l)}$ mediante las operaciones inducidas por los operadores (2.13). El procesador i -ésimo necesita calcular aquellas componentes que le son necesarias para sus sucesivos cálculos y aquellas para las que el correspondiente elemento diagonal de E_i es no nulo. Además, en la práctica se puede esperar que el número de iteraciones locales que cada procesador realiza entre dos iteraciones globales del algoritmo se fije y dependa sólo de A y de la cantidad de trabajo que implica sus cálculos. Bru, Elsner y Neumann [10] obtienen, para este modelo, el siguiente resultado de convergencia.

Teorema 7. *Supongamos que A es una matriz monótona de tamaño $n \times n$ y $\{M_i, N_i, E_i\}_{i=1}^r$, una multipartición de A débilmente regular. Entonces el esquema iterativo (2.15) converge para cualquier vector inicial $z^{(0)}$, siempre que $q(l, i) \geq 1$, $l = 1, 2, \dots$, $i = 1, 2, \dots, r$.*

Para la construcción del modelo asíncrono, Bru, Elsner y Neumann [10], suponen que dos procesadores no pueden acceder al mismo tiempo al último vector iterado global, evitando así los posibles conflictos. Con esta suposición dividen la tarea típica del procesador i -ésimo en tres subtareas que son cíclicamente realizadas. Estas tareas pueden describirse como sigue:

(i) Si w' es la aproximación a la solución global que reside en el procesador al comienzo del ciclo, el procesador actualiza alguna o todas las componentes de w' con las correspondientes componentes de la aproximación actual a la solución que reside en un procesador central, y obtiene la aproximación revisada w .

(ii) El procesador calcula en la iteración local, el vector

$$u = E_i F_i w = E_i M_i^{-1} N_i w + E_i M_i^{-1} b. \quad (2.16)$$

(iii) Si y es la aproximación a la solución presente en el procesador central cuando el procesador completa el cálculo de (2.16), el procesador actualiza el procesador principal como

$$z = u + (I - E_i)y. \quad (2.17)$$

y este vector pasa a ser la nueva actualización para que el mismo procesador i empiece la primera tarea con $w' = z$.

Denotamos por i_j un procesador cualquiera. Nótese que $1 \leq i_j \leq r$. Para formalizar el modelo asíncrono se necesita trabajar con sucesiones $\{i_j\}_{j=1}^{\infty}$ que poseen ciertas propiedades que se describen a continuación.

Definición 5. Sea $\{i_j\}_{j=1}^{\infty}$, $1 \leq i_j \leq r$, una sucesión de enteros positivos.

- (i) Se dice que la sucesión es *admisible* si cada uno de los enteros $1, 2, \dots, r$, aparece infinitas veces en la sucesión.
- (ii) Se dice que la sucesión es *regulada* si existe un entero positivo K tal que los enteros $1, 2, \dots, r$, aparecen al menos una vez cada K elementos consecutivos de la sucesión.

Claramente, toda sucesión regulada es admisible.

El esquema asíncrono dado en [10] y que definimos a continuación se basa en esta clase de sucesiones.

Se considera el proceso iterativo

$$x^{(j+r_j)} = (I - E_{i_j}) x^{(j+r_j-1)} + E_{i_j} M_{i_j}^{-1} [N_{i_j} x^{(j)} + b], \quad (2.18)$$

$$j = -r+1, -r+2, \dots, 0, 1, 2, \dots$$

Cuando $j \leq 0$, entonces $i_j = -j+1$, mientras que si $j > 0$ entonces i_j indica el procesador que ha actualizado el vector global con el vector $x^{(j)}$.

El escalar r_j indica el número de veces que se ha actualizado el vector de la iteración global por otros procesadores distintos del procesador i_j -ésimo, durante el tiempo que ha necesitado este último procesador para completar el trabajo de las tres etapas descritas anteriormente. Se supone que

$$x^{(-r+1)} = x^{(-r+2)} = \dots = x^{(-1)} = x^{(0)}.$$

Bru, Elsner y Neumann demuestran en [10] la convergencia de este modelo asíncrono. Las condiciones se basan en la monotonía de la matriz A y en la regularidad de la multipartición.

Teorema 8. *Supongamos que la matriz del sistema lineal $Ax = b$ es una matriz monótona. Sea $\{M_i, N_i, E_i\}_{i=1}^r$ una multipartición débilmente regular de A . Entonces la iteración definida en (2.18) converge a la solución exacta de dicho sistema, para cualquier sucesión regulada de enteros $\{i_j\}_{j=1}^{\infty}$ siempre que $0 \leq r_j - 1$ y $j \geq -r + 1$.*

siendo $B_{ij} = -A_{ij}^{-1}A_{ij}$, $j \neq i$, $1 \leq j \leq r$, y el vector

$$\bar{b}_i = \begin{bmatrix} 0 \\ \vdots \\ b_i \\ \vdots \\ 0 \end{bmatrix}.$$

Con las propiedades de una multipartición de A , el método de Jacobi caótico por bloques se puede escribir de la forma

$$z^{(l)} = \sum_{i=1}^r E_i F_i^{q(l,i)} z^{(l-1)}, \quad l = 1, 2, \dots, \quad (2.20)$$

donde $q(l,i) \geq 1$, $i = 1, 2, \dots, r$, y $l = 1, 2, \dots$. Nótese que este esquema es sincrónico.

Es fácil observar que las matrices H_i definidas por (2.19) son idempotentes y en consecuencia

$$F_i^{q(l,i)} z = F_i z.$$

Por tanto el método caótico de Jacobi por bloques definido en (2.20) se reduce

$$z^{(l)} = \sum_{i=1}^r E_i F_i z^{(l-1)}, \quad l = 1, 2, \dots,$$

que es el método de Jacobi por bloques.

Por otro lado, Avdelas, de Pillis, Hadjidimos y Neumann en [4] estudian determinados casos, en los que el método extrapolado de Jacobi tiene un mejor radio de convergencia que el no extrapolado.

Estos resultados, sugirieron a Bru y Fuster en [11] el estudio de esquemas caóticos basados en la partición de Jacobi, pero relajados. Para esto, suponen

que la matriz A del sistema lineal $Ax = b$, esté dividida en bloques, con los bloques diagonales igual a la identidad, es decir, consideran el siguiente sistema lineal

$$\begin{bmatrix} I_{11} & \cdots & A_{1j} & \cdots & A_{1r} \\ \vdots & & \vdots & & \vdots \\ A_{j1} & \cdots & I_{jj} & \cdots & A_{jr} \\ \vdots & & \vdots & & \vdots \\ A_{r1} & \cdots & A_{rj} & \cdots & I_{rr} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_r \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_j \\ \vdots \\ b_r \end{bmatrix}, \quad (2.21)$$

donde los vectores x y b están divididos de acuerdo con el tamaño de los bloques de A .

El esquema paralelo caótico que plantean es el siguiente:

$$x^{(l+1)} = \sum_{j=1}^r E_j F_j^{q(l,j)} x^{(l)}, \quad l = 0, 1, \dots, \quad (2.22)$$

donde la matriz por bloques E_j , $1 \leq j \leq r$ está dividida de acuerdo con el tamaño de la matriz A , con el bloque diagonal j -ésimo igual a la matriz identidad I_{jj} y los demás elementos nulos, es decir,

$$E_j = \begin{bmatrix} O & \cdots & O & \cdots & O \\ \vdots & & \vdots & & \vdots \\ O & \cdots & I_{jj} & \cdots & O \\ \vdots & & \vdots & & \vdots \\ O & \cdots & O & \cdots & O \end{bmatrix}. \quad (2.23)$$

El parámetro $q(l, j)$ indica el número de veces que el procesador j -ésimo va a actualizar su bloque de vector en la iteración l -ésima, antes de que cualquier otro procesador conozca sus cálculos.

Por F_j , $1 \leq j \leq r$, representamos r operadores que asignan el trabajo al j -ésimo procesador. Estos operadores están definidos como

$$F_j x = J_j x + \omega v^{(j)}, \quad 1 \leq j \leq r, \quad (2.24)$$

donde ω es el parámetro de relajación, el vector $v^{(j)}$ tiene el bloque j -ésimo igual al del vector de términos independientes b y los demás son nulos, es decir,

$$v^{(j)} = (0, \dots, b_j^t, \dots, 0)^t, \quad 1 \leq j \leq r,$$

y la matriz J_j es

$$J_j = \begin{bmatrix} I_{11} & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ \omega B_{j1} & \dots & (1 - \omega)I_{jj} & \dots & \omega B_{jr} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & I_{rr} \end{bmatrix}, \quad 1 \leq j \leq r,$$

con

$$B = [B_{ij}] : B_{ij} = \begin{cases} -A_{ij} & i \neq j \\ 0 & i = j. \end{cases}$$

De acuerdo con el esquema de iteración (2.22), para calcular el vector iterado $x^{(l+1)}$, el procesador j -ésimo actualiza $q(l, j)$ veces el correspondiente vector $x^{(l)}$ mediante

$$F_j x^{(l)} = J_j x^{(l)} + \omega v^{(j)}.$$

Una vez realizadas estas operaciones se premultiplica en (2.22) por la matriz E_j , definida en (2.23). Esto supone que dicha actualización se reduce únicamente a la actualización del bloque j -ésimo, ya que sólo éste va a ser necesario,

según la forma de las matrices E_j , para la actualización del vector iterado global. Esta actualización global se produce cuando todos estos cálculos han sido completados por todos los procesadores; desde este punto de vista el esquema (2.22) es síncrono.

En [11] se dan resultados de convergencia para el esquema (2.22) cuando el factor de relajación óptimo está acotado entre 0 y 1, bajo condiciones bastante generales que dependen de la matriz de Jacobi. Posteriormente en [33] se extienden estos resultados de convergencia para el caso en que el factor de relajación toma valores en un cierto intervalo $(0, \omega_0)$, con $\omega_0 > 1$.

Teorema 9. *Consideremos el sistema lineal no singular (2.21). Supongamos que el radio espectral de la matriz de Jacobi por bloques B es menor que 1. Si $0 < \omega < \frac{1}{1+\rho(B)}$ y $q(l, i) = q(l, j) \geq 1$, $l = 0, 1, \dots$, $1 \leq i, j \leq r$, entonces el esquema iterativo (2.22) converge a la solución del sistema lineal (2.21).*

En [33] también se estudia la convergencia del esquema (2.22) debilitando las condiciones que deben cumplir los factores caóticos $q(l, j)$ respecto al teorema anterior, exigiéndoles únicamente que sean mayores o igual que uno. Sin embargo las hipótesis relacionadas con la matriz de Jacobi por bloques son relativas a la norma infinito, exigiéndole a ésta, que sea menor que 1.

Teorema 10. *Supongamos que $\|B\|_\infty < 1$. Si $0 < \omega < \frac{1}{1+\|B\|_\infty}$ y $q(l, j) \geq 1$, $l = 0, 1, \dots$, $1 \leq j \leq r$, entonces el esquema de iteración (2.22) converge a la solución del sistema lineal (2.21).*

Además, en [33] se plantea la versión asincrónica, del esquema planteado en [11] demostrando la convergencia bajo las mismas hipótesis para las que se

prueba el Teorema 10. Los experimentos numéricos que se presentan en [12] y [33] reflejan la aceleración de los modelos caóticos o síncronos, en los que a priori se ha equilibrado la carga entre los procesadores, y de los modelos asíncronos frente al correspondiente modelo paralelo no caótico.

Por otro lado el efecto de los procesadores sobre la convergencia de métodos iterativos fue objeto de estudio por Elsner, Neumann y Vemmer en [24] y posteriormente por Elsner y Neumann en [23].

En artículos previos a [24] sobre métodos de multipartición, tanto síncronos como asíncronos, se asume tácitamente que el número de particiones coincide con el número de procesadores. Este es el caso en [10], [57] y [61]. Mientras en el caso síncrono tal suposición es natural, es innecesaria en el caso asíncrono.

Elsner, Neumann y Vemmer en [24] y posteriormente Elsner y Neumann [23] examinan el efecto que produce sobre la aceleración de la convergencia, la variación del número de procesadores en relación con el número de particiones. Concluyen que los procesos iterativos asíncronos que utilizan aproximaciones más recientes para construir el vector iterado global convergen más rápidamente que los procesos que utilizan aproximaciones menos recientes.

Capítulo 3

Método de multipartición en dos etapas

3.1 Introducción

Consideremos el sistema lineal

$$Az = b, \quad (3.1)$$

donde A es una matriz cuadrada de tamaño $n \times n$, invertible.

Supongamos que para la resolución del sistema lineal (3.1) se dispone de r procesadores, conectados en paralelo. Como ya se explicó en el Capítulo 2, la literatura existente sobre métodos iterativos paralelos para resolver dicho sistema, indica que se consigue mejorar la eficiencia de los algoritmos paralelos, cuando se obtiene un equilibrio en la carga de trabajo entre los procesadores.

Así, en [12] se define un método iterativo paralelo y caótico, basado en la técnica en dos etapas, estudiando la versión síncrona y asíncrona. La técnica en dos etapas ha sido desarrollada desde el punto de vista secuencial en la Sección 1.4.

Para el planteamiento del esquema iterativo, definido en [12], se supone que la matriz A del sistema lineal (3.1) está dividida en bloques y los bloques diagonales son cuadrados. De esta forma, el sistema (3.1) se escribe como

$$\begin{bmatrix} D_{11} & \cdots & A_{1j} & \cdots & A_{1r} \\ \vdots & & \vdots & & \vdots \\ A_{j1} & \cdots & D_{jj} & \cdots & A_{jr} \\ \vdots & & \vdots & & \vdots \\ A_{r1} & \cdots & A_{rj} & \cdots & D_{rr} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_r \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_j \\ \vdots \\ b_r \end{bmatrix}, \quad (3.2)$$

donde los vectores x y b están divididos de acuerdo con el tamaño de los bloques de A .

Con estas hipótesis, los autores de [12] plantean un procedimiento iterativo que consiste en construir la siguiente sucesión de vectores

$$x^{(l+1)} = \begin{bmatrix} x_1^{(l+1)} \\ \vdots \\ x_j^{(l+1)} \\ \vdots \\ x_r^{(l+1)} \end{bmatrix}, \quad l = 0, 1, \dots, \quad (3.3)$$

donde cada $x_j^{(l+1)}$, $1 \leq j \leq r$, es una solución aproximada del sistema lineal

$$D_{jj}x_j = b_j + \sum_{i=1, i \neq j}^r B_{ji}x_i^{(l)}, \quad (3.4)$$

que se obtiene realizando $q(l, j)$ pasos "internos" del esquema iterativo relajado

$$z_j^{(k+1)} = \omega M_j^{-1} N_j z_j^{(k)} + \omega M_j^{-1} \left(b_j + \sum_{i=1, i \neq j}^r B_{ji} x_i^{(k)} \right) + (1 - \omega) z_j^{(k)}, \quad (3.5)$$

$$k = 0, 1, \dots, q(l, j) - 1, \quad z_j^{(0)} = x_j^{(l)},$$

siendo B la siguiente matriz dividida en bloques

$$B = [B_{ij}] : B_{ij} = \begin{cases} -A_{ij} & i \neq j \\ 0 & i = j \end{cases} \quad 1 \leq i, j \leq r, \quad (3.6)$$

y $D_{jj} = M_j - N_j$ son r particiones, $1 \leq j \leq r$,

Finalizadas estas $q(l, j)$ etapas, para todo $j = 1, 2, \dots, r$, se define la $(l+1)$ iteración "externa" como

$$x_j^{(l+1)} = z_j^{(q(l, j))}.$$

Claramente, si se dispone de r procesadores, cada cálculo de los vectores $x_j^{(l+1)}$, $1 \leq j \leq r$, mediante la expresión (3.5), puede asignarse a un procesador, realizándose así, todos los cálculos en paralelo.

En este modelo cada procesador j , $1 \leq j \leq r$, puede realizar un número arbitrario de iteraciones locales para calcular el vector $x_j^{(l+1)}$, antes de que se forme la nueva aproximación a la solución global (3.3). Este es, por tanto, un esquema caótico extrapolado, pero síncrono, que se puede expresar, teniendo en cuenta la expresión (3.5), como

$$x^{(l+1)} = \sum_{j=1}^r E_j F_j^{q(l, j)} x^{(l)}, \quad l = 0, 1, \dots, \quad (3.7)$$

donde las matrices E_j están divididas según el tamaño de los bloques de A y tienen el bloque diagonal j -ésimo igual a I_{jj} y los demás bloques son cero;

$q(l, j)$ es el número de veces que el procesador j -ésimo actualiza el bloque correspondiente de vector iterado y F_j son r operadores, cada uno de los cuales asigna el trabajo al j -ésimo procesador, definidos como

$$F_j x = J_j x + \omega M_j^{-1} v^{(j)}, \quad (3.8)$$

donde ω es el factor de relajación, $v^{(j)} = (0, \dots, b_j^t, \dots, 0)^t$, y

$$J_j = \begin{bmatrix} I_{11} & \dots & O & \dots & O \\ \vdots & & \vdots & & \vdots \\ \omega M_j^{-1} B_{j1} & \dots & H_j & \dots & \omega M_j^{-1} B_{jr} \\ \vdots & & \vdots & & \vdots \\ O & \dots & O & \dots & I_{rr} \end{bmatrix}, \quad (3.9)$$

con

$$H_j = (1 - \omega)I_{jj} + \omega M_j^{-1} N_j. \quad (3.10)$$

A continuación, resumimos los resultados de convergencia, obtenidos para el esquema (3.7), en [12].

El siguiente teorema muestra que para cualquier partición convergente de $D = \text{diag}(D_{11}, D_{22}, \dots, D_{rr})$, el método caótico síncrono construido en [12] converge si $q(l, j)$ tiende a infinito, para todo $j = 1, 2, \dots, r$. Se exige, además, que la norma infinito de $D^{-1}B$ sea menor que 1 y que el valor de ω esté comprendido en un cierto intervalo.

Teorema 1. *Sea A la matriz del sistema lineal (3.2), con $A = D - B$ y $D = \text{diag}(D_{11}, \dots, D_{rr})$. Sean $D_{jj} = M_j - N_j$, $1 \leq j \leq r$, particiones convergentes y supongamos que $\|D^{-1}B\|_\infty < 1$. Si $0 < \omega < \frac{1}{1+\rho}$, con $\rho = \max_{1 \leq j \leq r} \rho(M_j^{-1}N_j)$ y $\lim_{l \rightarrow \infty} q(l, j) = \infty$, entonces la sucesión de vectores generada por el esquema iterativo caótico (3.7) converge a la única solución del sistema lineal (3.2).*

En el artículo citado, también se demuestra la convergencia del esquema (3.7), sin exigir que los factores caóticos tiendan a infinito, sino que sean mayores o iguales que 1. Las hipótesis exigidas son que la partición externa sea convergente y regular. Además las particiones internas deben ser débilmente regulares.

Teorema 2. *Sea A la matriz del sistema lineal (3.2), con $A = D - B$ una partición regular y convergente, y $D = \text{diag}(D_{11}, \dots, D_{rr})$. Sean $D_{jj} = M_j - N_j$, $1 \leq j \leq r$, particiones débilmente regulares. Si $0 < \omega \leq 1$ y $q(l, j) \geq 1$, $1 \leq j \leq r$, $l = 0, 1, \dots$, entonces la sucesión de vectores generada por el esquema iterativo caótico (3.7) converge a la única solución del sistema lineal (3.2).*

Una vez estudiado el caso síncrono, los autores de [12] presentan el estudio de una versión asíncrona del esquema (3.7), probando la convergencia bajo las mismas hipótesis que en el caso síncrono. Los experimentos numéricos que obtienen muestran la buena eficiencia de los algoritmos en dos etapas paralelos caóticos.

Todo esto, junto a los antecedentes estudiados sobre los métodos basados en multiparticiones, en el Capítulo 2, nos ha sugerido la generalización del esquema (3.7), de forma que los sistemas internos no se resuelvan necesariamente sobre los bloques diagonales de la matriz A . Para ello consideraremos una multipartición de A ,

$$A = F_j - G_j, \quad j = 1, 2, \dots, r,$$

y aplicaremos $q(l, j)$ etapas de un método iterativo relajado (también basado en

particiones) para resolver en cada iteración l , r subsistemas

$$F_j z = G_j z^{(l)} + b, \quad j = 1, 2, \dots, r.$$

En la Sección 3.2 se planteará el esquema caótico síncrono, y se estudiará la convergencia. La Sección 3.3 estará dedicada a la versión asíncrona del nuevo esquema. Finalizaremos el capítulo con una sección dedicada a los experimentos numéricos, y otra en la que se resumirán los resultados más importantes obtenidos.

3.2 Modelo síncrono

Consideremos el sistema lineal

$$Ax = b, \quad (3.11)$$

donde A es una matriz cuadrada de tamaño $n \times n$, no singular y x y b son vectores de tamaño n .

Definición 1. Diremos que el conjunto de matrices

$$\{F_j, G_j, M_j, N_j, E_j\}_{j=1}^r,$$

es una multipartición en dos etapas de la matriz A , si verifica las siguientes condiciones.

- (i) $A = F_j - G_j$, son particiones de A , para todo $j = 1, 2, \dots, r$.

$$(ii) \sum_{j=1}^r E_j = I \text{ y } E_j \geq O, \text{ para todo } j = 1, 2, \dots, r.$$

$$(iii) F_j = M_j - N_j, \text{ es una partición de } F_j, \text{ para todo } j = 1, 2, \dots, r.$$

Consideremos $\{F_j, G_j, M_j, N_j, E_j\}_{j=1}^r$, una multipartición en dos etapas de la matriz A y sea $x^{(0)}$ un vector arbitrario. A partir de este vector construimos una sucesión de vectores $\{x^{(l)}\}_{l=0}^{\infty}$ de la siguiente forma.

Sea $x_j^{(l+1)}$ una aproximación a la solución del sistema lineal

$$F_j z = G_j x^{(l)} + b, \quad l = 0, 1, 2, \dots,$$

obtenida realizando $q(l, j)$ etapas, que llamaremos *internas*, de un método iterativo relajado basado en la partición $F_j = M_j - N_j$.

Una vez calculados, a partir del vector $x^{(l)}$, todos los vectores $x_j^{(l+1)}$, se construye el siguiente elemento de la sucesión $\{x^{(l)}\}_{l=0}^{\infty}$, como

$$x^{(l+1)} = \sum_{j=1}^r E_j x_j^{(l+1)}. \quad (3.12)$$

Así pues, el cálculo de los vectores $x_j^{(l+1)}$, $j = 1, 2, \dots, r$, en cada iteración externa l , implica la realización de las siguientes $q(l, j)$ etapas

$$z_j^{(k+1)} = \omega M_j^{-1} N_j z_j^{(k)} + \omega M_j^{-1} (b_j + G_j x^{(l)}) + (1 - \omega) z_j^{(k)}, \quad (3.13)$$

$$k = 0, 1, \dots, q(l, j) - 1, \quad z_j^{(0)} = x^{(l)},$$

Finalizadas estas $q(l, j)$ etapas, se define el vector x_j en la $(l + 1)$ iteración externa como

$$x_j^{(l+1)} = z_j^{(q(l, j))}. \quad (3.14)$$

Claramente, si se dispone de r procesadores, cada cálculo de los vectores $x_j^{(l+1)}$, $1 \leq j \leq r$, mediante la expresión (3.13), puede asignarse a un procesador,

realizándose así, todos los cálculos en paralelo. Además, no es necesario que el procesador j -ésimo calcule todas las componentes del vector $\mathbf{x}_j^{(l+1)}$, sino aquellas para las que se verifica que el correspondiente elemento diagonal de la matriz E_j es no nulo. Sólo estas componentes serán relevantes para el cálculo del vector $\mathbf{x}^{(l+1)}$ en (3.12).

En este modelo, cada procesador j , $1 \leq j \leq r$, puede realizar un número arbitrario de iteraciones locales para calcular el vector $\mathbf{x}_j^{(l+1)}$, antes de que se forme la nueva aproximación a la solución global (3.12). Por tanto, este es un esquema caótico pero síncrono.

Tal y como se indica en las expresiones (3.12), (3.13) y (3.14), cada vector $\mathbf{x}_j^{(l+1)}$ se calcula aplicando $q(l, j)$ veces el siguiente operador sobre el vector $\mathbf{x}^{(l)}$

$$P_{lj} : \mathbf{R}^n \longrightarrow \mathbf{R}^n$$

$$P_{lj}\mathbf{x} = H_j\mathbf{x} + \omega M_j^{-1} (G_j\mathbf{x}^{(l)} + \mathbf{b}), \quad (3.15)$$

siendo

$$H_j = (1 - \omega)I + \omega M_j^{-1} N_j \quad (3.16)$$

y ω el factor de relajación.

Por tanto, el esquema paralelo caótico extrapolado que planteamos, puede expresarse como

$$\mathbf{x}^{(l+1)} = \sum_{j=1}^r E_j P_{lj}^{q(l,j)} \mathbf{x}^{(l)}, \quad l = 0, 1, 2, \dots, \quad (3.17)$$

donde las matrices E_j son no negativas y verifican $\sum_{j=1}^r E_j = I$, $q(l, j)$ es el número de veces que el procesador j -ésimo actualiza parte de las componentes del vector $\mathbf{x}^{(l+1)}$ y P_{lj} son r operadores, para cada iteración externa l , definidos en la expresión (3.15), que asignan el trabajo al j -ésimo procesador.

Para analizar la convergencia del esquema iterativo (3.17) vamos a detallar más explícitamente dicho esquema. Así, por la forma de los operadores P_j en (3.15) podemos escribir las siguientes igualdades, a partir del esquema (3.17)

$$\begin{aligned}
 \mathbf{x}^{(l+1)} &= \sum_{j=1}^r E_j P_j^{q^{(l,j)}} \mathbf{x}^{(l)} \\
 &= \sum_{j=1}^r E_j P_j^{q^{(l,j)}-1} \left[H_j \mathbf{x}^{(l)} + \omega M_j^{-1} (G_j \mathbf{x}^{(l)} + \mathbf{b}) \right] \\
 &= \sum_{j=1}^r E_j P_j^{q^{(l,j)}-2} \left[H_j (H_j \mathbf{x}^{(l)} + \omega M_j^{-1} (G_j \mathbf{x}^{(l)} + \mathbf{b})) \right. \\
 &\quad \left. + \omega M_j^{-1} (G_j \mathbf{x}^{(l)} + \mathbf{b}) \right] \\
 &= \sum_{j=1}^r E_j P_j^{q^{(l,j)}-2} \left[H_j^2 \mathbf{x}^{(l)} + \omega (H_j + I) M_j^{-1} (G_j \mathbf{x}^{(l)} + \mathbf{b}) \right] \\
 &\quad \dots \\
 &= \sum_{j=1}^r E_j \left[H_j^{q^{(l,j)}} \mathbf{x}^{(l)} + \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} (G_j \mathbf{x}^{(l)} + \mathbf{b}) \right]. \quad (3.18)
 \end{aligned}$$

Por tanto, el esquema (3.17) es equivalente al esquema

$$\mathbf{x}^{(l+1)} = \sum_{j=1}^r E_j \left[H_j^{q^{(l,j)}} \mathbf{x}^{(l)} + \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} (G_j \mathbf{x}^{(l)} + \mathbf{b}) \right]. \quad (3.19)$$

Para analizar la convergencia del esquema (3.17) o equivalentemente (3.19), utilizaremos el análisis del error

$$\mathbf{e}^{(l+1)} = \mathbf{x}^{(l+1)} - \boldsymbol{\xi},$$

donde $\boldsymbol{\xi}$ es la solución exacta del sistema lineal $A\mathbf{x} = \mathbf{b}$. Bastará con demostrar que la sucesión de vectores error $\{\mathbf{e}^{(l)}\}_{l=0}^{\infty}$, converge al vector \mathbf{o} .

En los siguientes lemas damos una expresión del vector error $e^{(l)}$ en cada iteración, en función del vector error en la iteración inicial y de un producto de matrices que denominaremos *matrices de iteración*.

Lema 1. *La solución ξ del sistema lineal $Ax = b$, es un punto fijo del esquema iterativo (3.19), es decir,*

$$\xi = \sum_{j=1}^r E_j \left[H_j^{q^{(l,j)}} \xi + \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} (G_j \xi + b) \right].$$

Demostración. Teniendo en cuenta las particiones $A = F_j - G_j$, $F_j = M_j - N_j$, $j = 1, 2, \dots, r$ y la definición de H_j , $j = 1, 2, \dots, r$ en (3.16), obtenemos, para cada $j = 1, 2, \dots, r$

$$\begin{aligned} \omega M_j^{-1} (G_j \xi + b) &= \omega M_j^{-1} F_j \xi = \omega M_j^{-1} (M_j - N_j) \xi \\ &= \omega (I - M_j^{-1} N_j) \xi \\ &= (I - I + \omega I - \omega M_j^{-1} N_j) \xi \\ &= [I - ((1 - \omega)I + \omega M_j^{-1} N_j)] \xi \\ &= (I - H_j) \xi. \end{aligned} \tag{3.20}$$

Teniendo en cuenta la expresión (3.18), pero sustituyendo $x^{(l)}$ por ξ , y considerando la igualdad (3.20) que acabamos de demostrar, podemos establecer las siguientes igualdades.

$$\begin{aligned} &\sum_{j=1}^r E_j \left[H_j^{q^{(l,j)}} \xi + \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} (G_j \xi + b) \right] \\ &= \sum_{j=1}^r E_j \left[H_j^{q^{(l,j)}} \xi + \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) (I - H_j) \xi \right] \end{aligned}$$



$$\begin{aligned}
 &= \sum_{j=1}^r E_j \left[H_j^{q(l,j)} \xi + (I - H_j^{q(l,j)}) \xi \right] \\
 &= \sum_{j=1}^r E_j \xi \\
 &= \xi.
 \end{aligned}$$

Con lo que queda probado el lema. ■

Lema 2. Sea ξ la solución del sistema lineal $Ax = b$ y $\{x^{(l)}\}_{l=0}^{\infty}$ la sucesión de vectores generada por el esquema (3.19). Entonces, la sucesión de vectores error $\{e^{(l)}\}_{l=0}^{\infty}$, con $e^{(l)} = x^{(l)} - \xi$, satisface

$$e^{(l+1)} = \sum_{j=1}^r E_j \left[H_j^{q(l,j)} + \omega \left(\sum_{i=0}^{q(l,j)-1} H_j^i \right) M_j^{-1} G_j \right] e^{(l)}, \quad l = 0, 1, \dots$$

Demostración. Teniendo en cuenta la expresión (3.19) y el Lema 1, podemos analizar la expresión del error como sigue.

$$\begin{aligned}
 e^{(l+1)} &= x^{(l+1)} - \xi \\
 &= \sum_{j=1}^r E_j \left[H_j^{q(l,j)} x^{(l)} + \omega \left(\sum_{i=0}^{q(l,j)-1} H_j^i \right) M_j^{-1} (G_j x^{(l)} + b) \right] \\
 &\quad - \sum_{j=1}^r E_j \left[H_j^{q(l,j)} \xi + \omega \left(\sum_{i=0}^{q(l,j)-1} H_j^i \right) M_j^{-1} (G_j \xi + b) \right] \\
 &= \sum_{j=1}^r E_j \left[H_j^{q(l,j)} (x^{(l)} - \xi) + \omega \left(\sum_{i=0}^{q(l,j)-1} H_j^i \right) M_j^{-1} (G_j x^{(l)} + b) \right. \\
 &\quad \left. - \omega \left(\sum_{i=0}^{q(l,j)-1} H_j^i \right) M_j^{-1} (G_j \xi + b) \right] \\
 &= \sum_{j=1}^r E_j \left[H_j^{q(l,j)} e^{(l)} + \omega \left(\sum_{i=0}^{q(l,j)-1} H_j^i \right) M_j^{-1} G_j e^{(l)} \right]
 \end{aligned}$$

$$= \sum_{j=1}^r E_j \left[H_j^{q^{(l,j)}} + \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} G_j \right] e^{(l)}.$$

■

Observación 1. Por el Lema 2, la sucesión de vectores error verifica

$$e^{(l+1)} = \sum_{j=1}^r E_j T_j^{(l)} e^{(l)}, \quad l = 0, 1, \dots,$$

donde las matrices $T_j^{(l)}$ están definidas como

$$T_j^{(l)} = H_j^{q^{(l,j)}} + \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} G_j. \quad (3.21)$$

Podemos escribir estas matrices $T_j^{(l)}$ de otra forma equivalente y que nos será de utilidad con posterioridad. Para ello, es necesario reescribir las expresiones $\omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} G_j$, para todo $j = 1, 2, \dots, r$, $l = 0, 1, \dots$, como sigue

$$\begin{aligned} \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} G_j &= \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} F_j F_j^{-1} G_j \\ &= \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} (M_j - N_j) F_j^{-1} G_j \\ &= \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) (I - M_j^{-1} N_j) F_j^{-1} G_j \\ &= \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) (I - I + \omega I - \omega M_j^{-1} N_j) F_j^{-1} G_j \\ &= \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) (I - H_j) F_j^{-1} G_j \\ &= (I - H_j^{q^{(l,j)}}) F_j^{-1} G_j. \end{aligned} \quad (3.22)$$

Luego, por la igualdad (3.22) y la expresión de las matrices $T_j^{(l)}$ en (3.21), obtenemos la siguiente nueva expresión para dichas matrices.

$$T_j^{(l)} = H_j^{q(l,j)} + (I - H_j^{q(l,j)}) F_j^{-1} G_j. \quad (3.23)$$

A las matrices

$$T^{(l)} = \sum_{j=1}^r E_j T_j^{(l)}, \quad l = 0, 1, \dots, \quad (3.24)$$

donde $T_j^{(l)}$ vienen definidas en (3.21) o equivalentemente en (3.23), las llamaremos *matrices de iteración* del esquema iterativo (3.19).

Con estas matrices, la sucesión de vectores error verifica, según el Lema 2

$$\mathbf{e}^{(l+1)} = T^{(l)} \mathbf{e}^{(l)} = \prod_{i=0}^l T^{(i)} \mathbf{e}^{(0)}, \quad l = 0, 1, \dots$$

Por tanto, como la convergencia del esquema iterativo (3.19) es equivalente a la convergencia de la sucesión $\{\mathbf{e}^{(l)}\}_{l=0}^{\infty}$ al vector cero, podemos afirmar que el esquema (3.19) converge, si y sólo si

$$\lim_{l \rightarrow \infty} \prod_{i=0}^l T^{(i)} = O.$$

El siguiente lema nos será de utilidad en los resultados de convergencia y es debido a Bru y Fuster [11].

Lema 3. *Sea (E, d) un espacio métrico completo y sean $A^{(l)}$, $l = 0, 1, \dots$, aplicaciones contractivas con el mismo punto fijo \mathbf{x} y constantes contractivas $\alpha^{(l)}$ respectivamente. Si existe una constante α tal que $\alpha^{(l)} \leq \alpha < 1$, $l = 0, 1, \dots$, entonces el esquema iterativo $\mathbf{x}^{(l+1)} = A^{(l)} \mathbf{x}^{(l)}$ converge a \mathbf{x} para cualquier $\mathbf{x}^{(0)}$. En particular, si $A^{(l)}$ son matrices complejas cuadradas y existe una norma matricial compatible $\|\cdot\|$ tal que*

$$\|A^{(l)}\| \leq \alpha < 1, \quad l = 0, 1, \dots,$$

entonces

$$\lim_{l \rightarrow \infty} \prod_{i=1}^l A^{(i)} = O.$$

Los autores destacan que la condición $\|A^{(l)}\| \leq \alpha < 1$, $l = 0, 1, \dots$, puede ser aligerada, suponiendo que existe una subsucesión $\{l_i\}_{i=1}^{\infty}$ satisfaciendo

$$\|A^{(l_i)}\| \leq \alpha < 1, \quad i = 1, 2, \dots$$

Hacemos notar también que la acotación $\|A^{(l)}\| \leq \alpha < 1$, $l = 0, 1, \dots$, no puede ser aligerada exigiendo que $\|A^{(l)}\| < 1$, $l = 0, 1, \dots$.

Para estudiar la convergencia del esquema iterativo (3.17) o (3.19) utilizaremos el lema que enunciamos a continuación. Este lema es la versión matricial del Lema de Neumann para series convergentes. Su demostración puede verse en Berman y Plemmons [7].

Lema 4. *La matriz no negativa T es convergente, es decir, $\rho(T) < 1$, si y sólo si, $(I - T)^{-1}$ existe y en este caso*

$$(I - T)^{-1} = \sum_{i=0}^{\infty} T^i \geq O.$$

A continuación, nos dedicamos a probar explícitamente la convergencia del método de multipartición en dos etapas extrapolado, definido en (3.17) o análogamente en (3.19) bajo distintas condiciones. En el primer resultado, demostramos la convergencia cuando las particiones tanto internas como externas son convergentes y se realiza un número infinito de iteraciones internas.

Teorema 3. *Sea $\{F_j, G_j, M_j, N_j, E_j\}_{j=1}^r$ una multipartición en dos etapas de la matriz del sistema lineal $Ax = b$. Supongamos que las particiones internas $F_j = M_j - N_j$, $j = 1, 2, \dots, r$ son convergentes y que las externas verifican*

$\|F_j^{-1}G_j\|_\infty < 1$, $j = 1, 2, \dots, r$. Si $0 < \omega \leq \frac{1}{1+\rho}$ con $\rho = \max_{1 \leq j \leq r} \rho(M_j^{-1}N_j)$ y $\lim_{l \rightarrow \infty} q(l, j) = \infty$, entonces el método de multipartición en dos etapas (3.19) converge a la solución ξ del sistema lineal $Ax = b$ para cualquier vector inicial $x^{(0)}$.

Demostración. Según la Observación 1 y el Lema 3, para probar la convergencia del esquema (3.19) será suficiente demostrar que $\|T^{(l)}\|_\infty \leq \gamma$, $l = 0, 1, \dots$, para alguna constante real $\gamma < 1$.

Veamos en primer lugar que el radio espectral de las matrices H_j , $j = 1, 2, \dots, r$, es menor que 1.

Teniendo en cuenta la definición de H_j , $j = 1, 2, \dots, r$, en (3.16) y las propiedades del radio espectral, obtenemos

$$\begin{aligned} \rho(H_j) &= \rho((1-\omega)I + \omega M_j^{-1}N_j) \\ &\leq |1-\omega| + \omega \rho(M_j^{-1}N_j) \\ &\leq |1-\omega| + \omega \max_{1 \leq j \leq r} \rho(M_j^{-1}N_j). \end{aligned}$$

Por tanto

$$\rho(H_j) \leq |1-\omega| + \omega \rho, \quad j = 1, 2, \dots, r,$$

con $\rho = \max_{1 \leq j \leq r} \rho(M_j^{-1}N_j) < 1$, ya que las particiones $F_j = M_j - N_j$, $j = 1, 2, \dots, r$, son convergentes.

Distinguimos dos casos atendiendo al valor de ω .

(a) Si $0 < \omega < 1$,

$$\begin{aligned} \rho(H_j) &\leq |1-\omega| + \omega \rho \\ &= 1 - \omega + \omega \rho \end{aligned}$$

$$< 1, \quad j = 1, 2, \dots, r.$$

(b) Si $1 \leq \omega < \frac{2}{1+\rho}$,

$$\begin{aligned} \rho(H_j) &\leq |1 - \omega| + \omega\rho \\ &= \omega - 1 + \omega\rho \\ &= \omega(1 + \rho) - 1 \\ &< \frac{2}{1 + \rho}(1 + \rho) - 1 \\ &= 1, \quad j = 1, 2, \dots, r. \end{aligned}$$

Para acotar la norma infinito de las matrices de iteración, utilizaremos la expresión de las matrices $T_j^{(l)}$, $l = 0, 1, 2, \dots$, en (3.23)

$$\begin{aligned} \|T^{(l)}\|_\infty &\leq \max_{1 \leq j \leq r} \left\{ \|H_j^{q(l,j)} + (I - H_j^{q(l,j)})F_j^{-1}G_j\|_\infty \right\} \\ &\leq \max_{1 \leq j \leq r} \left\{ \|H_j^{q(l,j)}\|_\infty + (1 + \|H_j^{q(l,j)}\|_\infty) \|F_j^{-1}G_j\|_\infty \right\}. \end{aligned}$$

Tal y como hemos demostrado, sabemos que el radio espectral de las matrices H_j es menor que 1, lo que implica, por el Teorema 3 de la Sección 1.2, que

$$\lim_{p \rightarrow \infty} H_j^p = \mathcal{O}, \quad j = 1, 2, \dots, r,$$

y por tanto, dado un $\epsilon > 0$

$$\exists p_0 \in \mathbb{N} : \|H_j^p\|_\infty \leq \epsilon, \quad \forall p \geq p_0, \quad j = 1, 2, \dots, r.$$

Ahora bien, como $\lim_{l \rightarrow \infty} q(l, j) = \infty$,

$$\exists l_0 \in \mathbb{N} : \|H_j^{q(l,j)}\|_\infty \leq \epsilon, \quad \forall l \geq l_0, \quad j = 1, 2, \dots, r.$$



3.2 Modelo síncrono

77

Por hipótesis existe un real β tal que

$$\|F_j^{-1}G_j\|_\infty \leq \beta < 1, \quad j = 1, 2, \dots, r.$$

Así, para $l \geq l_0$

$$\begin{aligned} \|T^{(l)}\|_\infty &\leq \max_{1 \leq j \leq r} \{ \epsilon + (1 + \epsilon) \|F_j^{-1}G_j\|_\infty \} \\ &\leq \epsilon + (1 + \epsilon)\beta \\ &= \epsilon(1 + \beta) + \beta \equiv \alpha_\epsilon. \end{aligned}$$

Tomando $\epsilon < \frac{1-\beta}{1+\beta}$ se verifica que,

$$\|T^{(l)}\|_\infty \leq \alpha_\epsilon < 1.$$

Y por tanto el esquema iterativo de multipartición en dos etapas (3.19) converge a la única solución del sistema lineal $Ax = b$. ■

En el siguiente teorema aplicamos el hecho de que la matriz A sea monótona, es decir, $A^{-1} \geq O$, además de exigir las condiciones de partición regular y débilmente regular sobre las particiones externas e internas, respectivamente.

Teorema 4. *Sea A tal que $A^{-1} \geq O$. Supongamos que las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$, son particiones regulares y que $F_j = M_j - N_j$ son particiones débilmente regulares. Si $0 < \omega \leq 1$, el método de multipartición en dos etapas (3.19) converge a la solución ξ del sistema lineal $Ax = b$, para cualquier vector inicial $x^{(0)}$ y cualquier sucesión de iteraciones internas $q(l, j) \geq 1$, $j = 1, 2, \dots, r$, $l = 0, 1, \dots$.*

Demostración. Si demostramos que las matrices de iteración $T^{(l)}$, $l = 0, 1, \dots$, definidas en (3.24) satisfacen

- (i) $T^{(l)} \geq O$, $l = 0, 1, \dots$,
- (ii) existe un vector $\mathbf{x} > \mathbf{o}$ y una constante real θ , con $0 \leq \theta < 1$, independiente de la iteración l , tal que

$$T^{(l)} \mathbf{x} \leq \theta \mathbf{x},$$

entonces, considerando la norma matricial monótona inducida por la norma vectorial definida en la expresión (1.3), se tendría que

$$\|T^{(l)}\|_{\mathbf{x}} \leq \theta < 1,$$

y por el Lema 3 obtendríamos

$$\lim_{l \rightarrow \infty} \prod_{i=0}^l T^{(i)} = O,$$

Teniendo en cuenta que la sucesión de vectores error $\{\mathbf{e}^{(l)}\}_{l=0}^{\infty}$ satisface

$$\mathbf{e}^{(l+1)} = \prod_{i=0}^l T^{(i)} \mathbf{e}^{(0)}, \quad l = 0, 1, \dots,$$

la convergencia al vector cero estaría probada, es decir

$$\lim_{l \rightarrow \infty} \mathbf{e}^{(l)} = \mathbf{o}.$$

En consecuencia, el esquema (3.19) sería convergente.

Veamos, pues, (i) y (ii).

- (i) Al ser las particiones $F_j = M_j - N_j$, $j = 1, 2, \dots, r$, débilmente regulares, $M_j^{-1} N_j \geq O$ y $M_j^{-1} \geq O$, $j = 1, 2, \dots, r$. Así, si $0 < \omega \leq 1$, entonces

$$H_j = (1 - \omega)I + \omega M_j^{-1} N_j \geq O, \quad j = 1, 2, \dots, r.$$

Además, como las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$, son regulares, las matrices G_j son no negativas, y por tanto

$$H_j^{q^{(l,j)}} + \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} G_j \geq O, \quad j = 1, 2, \dots, r, \quad l = 0, 1, \dots,$$

luego

$$T^{(l)} = \sum_{j=1}^r E_j \left[H_j^{q^{(l,j)}} + \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} G_j \right] \geq O, \quad l = 0, 1, \dots$$

(ii) Por la expresión de $T_j^{(l)}$ en (3.23) y las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$, obtenemos

$$\begin{aligned} T_j^{(l)} &= H_j^{q^{(l,j)}} + (I - H_j^{q^{(l,j)}}) F_j^{-1} G_j \\ &= I - (I - H_j^{q^{(l,j)}}) (I - F_j^{-1} G_j) \\ &= I - (I - H_j^{q^{(l,j)}}) F_j^{-1} (F_j - G_j) \\ &= I - (I - H_j^{q^{(l,j)}}) F_j^{-1} A. \end{aligned} \quad (3.25)$$

Consideremos \mathbf{e} un vector positivo, por ejemplo $\mathbf{e} = (1, 1, \dots, 1)^t$, y $\mathbf{x} = A^{-1} \mathbf{e}$. Como $A^{-1} \geq O$ y ninguna fila de A^{-1} tiene todos sus elementos nulos, entonces $\mathbf{x} > \mathbf{o}$.

Calculemos $T_j^{(l)} \mathbf{x}$, donde las matrices $T_j^{(l)}$ están evaluadas en (3.25) y $\mathbf{x} = A^{-1} \mathbf{e}$.

$$\begin{aligned} T_j^{(l)} \mathbf{x} &= \mathbf{x} - (I - H_j^{q^{(l,j)}}) F_j^{-1} A \mathbf{x} \\ &= \mathbf{x} - (I - H_j^{q^{(l,j)}}) F_j^{-1} \mathbf{e}. \end{aligned} \quad (3.26)$$

Como las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$, son regulares, se tiene que $F_j^{-1} \geq O$, $j = 1, 2, \dots, r$, y al ser $F_j = M_j - N_j$ particiones débilmente regulares, se tiene, por el Lema 1 de la Sección 1.2, que $\rho(M_j^{-1}N_j) < 1$, $j = 1, 2, \dots, r$. Luego

$$\begin{aligned}\rho(H_j) &= |1 - \omega| + \omega\rho(M_j^{-1}N_j) \\ &= (1 - \omega) + \omega\rho(M_j^{-1}N_j) \\ &< 1,\end{aligned}$$

y así, utilizando el Lema 4

$$(I - H_j)^{-1} = \sum_{i=0}^{\infty} H_j^i, \quad j = 1, 2, \dots, r. \quad (3.27)$$

Ahora bien, como para todo $j = 1, 2, \dots, r$, se tiene

$$F_j = M_j - N_j = \omega^{-1}M_j - \omega^{-1}[(1 - \omega)M_j + \omega N_j],$$

multiplicando ambas expresiones por ω y considerando la expresión de las matrices H_j , $j = 1, 2, \dots, r$ en (3.16), obtenemos

$$\omega F_j = M_j - [(1 - \omega)M_j + \omega N_j] = M_j (I - H_j).$$

Teniendo en cuenta ahora, la existencia de las matrices M_j^{-1} y de $(I - H_j)^{-1}$, estas últimas expresadas en (3.27), se tiene

$$F_j^{-1} = \omega (I - H_j)^{-1} M_j^{-1} = \omega \left(\sum_{i=0}^{\infty} H_j^i \right) M_j^{-1}.$$

Sustituyendo la expresión que hemos obtenido de las matrices F_j^{-1} en la igualdad dada en (3.26)

$$T_j^{(l)} \mathbf{x} = \mathbf{x} - (I - H_j^{q(l,j)}) \omega \left(\sum_{i=0}^{\infty} H_j^i \right) M_j^{-1} \mathbf{e}$$

$$\begin{aligned}
 &= \mathbf{x} - \omega \left(\sum_{i=0}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} \mathbf{e} \\
 &= \mathbf{x} - \omega M_j^{-1} \mathbf{e} - \omega \left(\sum_{i=1}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} \mathbf{e}.
 \end{aligned}$$

Como las matrices H_j y M_j^{-1} , $j = 1, 2, \dots, r$, son no negativas, se tiene que

$$\omega \left(\sum_{i=1}^{q^{(l,j)}-1} H_j^i \right) M_j^{-1} \mathbf{e} \geq \mathbf{0},$$

y por tanto

$$T_j^{(l)} \mathbf{x} \leq \mathbf{x} - \omega M_j^{-1} \mathbf{e}.$$

Ahora bien, como las matrices M_j^{-1} son no negativas y ninguna fila es totalmente nula, $M_j^{-1} \mathbf{e} > \mathbf{0}$. Si además tenemos en cuenta que $T_j^{(l)} \mathbf{x} \geq \mathbf{0}$, entonces podemos afirmar que existe una constante real θ_j , con $0 \leq \theta_j < 1$ tal que

$$T_j^{(l)} \mathbf{x} \leq \theta_j \mathbf{x}, \quad j = 1, 2, \dots, r.$$

Tomando $\theta = \max\{\theta_1, \theta_2, \dots, \theta_r\}$, se obtiene

$$\begin{aligned}
 T^{(l)} \mathbf{x} &= \sum_{j=1}^r E_j T_j^{(l)} \mathbf{x} \\
 &\leq \sum_{j=1}^r E_j \theta_j \mathbf{x} \\
 &\leq \sum_{j=1}^r E_j \theta \mathbf{x} \\
 &\leq \theta \mathbf{x}.
 \end{aligned}$$

Con lo que queda demostrado (ii).

De (i) y (ii) se deduce, tal y como hemos dicho al inicio de la demostración, la convergencia del esquema iterativo (3.19). ■

El siguiente teorema prueba la convergencia del esquema (3.19) cuando la matriz A es una H -matriz. Previo al siguiente resultado de convergencia será necesario conocer algunas propiedades sobre matrices que recogemos en los lemas que enunciamos a continuación.

Lema 5. Sean A y B dos matrices reales de orden n .

- (a) Si A es una H -matriz, entonces $|A^{-1}| \leq (A)^{-1}$.
- (b) Si $|A| \leq B$, entonces $\rho(A) \leq \rho(B)$.

Demostración. La parte (a) puede verse en Frommer y Mayer [29] mientras que la demostración de (b) puede seguirse en Ortega y Rheinboldt [63]. ■

Lema 6. Sea $A = M - N$ una partición de A .

- (a) Si la partición es una H -partición, entonces A y M son H -matrices y $\rho(M^{-1}N) \leq \rho((M)^{-1}|N|) < 1$.
- (b) Si la partición es una partición H -compatible y A es una H -matriz, entonces es una H -partición y por tanto convergente.

Demostración. La demostración de (a) puede verse, por ejemplo, en Mayer [55] o en Frommer y Szyld [31]. En este último también se da la demostración de (b). ■

Con estos resultados, podemos estudiar en el siguiente teorema, la convergencia del esquema (3.19) cuando la matriz A es una H -matriz.

Teorema 5. Supongamos que A es una H -matriz. Consideremos $A = F_j - G_j$, $j = 1, 2, \dots, r$, r particiones H -compatibles de A y $F_j = M_j - N_j$, particiones H -compatibles. Entonces, si $0 < \omega \leq 1$, el método de multipartición en dos etapas (3.19) converge a la solución ξ del sistema lineal $Ax = b$, para cualquier vector inicial $x^{(0)}$ y cualquier sucesión de iteraciones internas $q(l, j) \geq 1$, $j = 1, 2, \dots, r$, $l = 0, 1, \dots$.

Demostración. Teniendo en cuenta la expresión de las matrices H_j en (3.16)

$$H_j = (1 - \omega)I + \omega M_j^{-1} N_j,$$

y que $0 < \omega \leq 1$

$$|H_j| \leq (1 - \omega)I + \omega |M_j^{-1}| |N_j|. \quad (3.28)$$

Considerando el módulo de las matrices $T_j^{(l)}$ en su expresión en (3.21)

$$|T_j^{(l)}| \leq |H_j|^{q(l, j)} + \omega \left(\sum_{i=0}^{q(l, j)-1} |H_j|^i \right) |M_j^{-1}| |G_j|. \quad (3.29)$$

Como $F_j = M_j - N_j$ son particiones H -compatibles, por el Lema 6, apartado (b), deducimos que $F_j = M_j - N_j$, $j = 1, 2, \dots, r$, son H -particiones y por el apartado (a) del mismo lema obtenemos que las matrices M_j son H -matrices. Teniendo en cuenta ahora, el Lema 5, apartado (a), se verifica la siguiente acotación

$$|M_j^{-1}| \leq (M_j)^{-1}, \quad j = 1, 2, \dots, r. \quad (3.30)$$

Luego, por (3.28) y (3.30)

$$|H_j| \leq (1 - \omega)I + \omega (M_j)^{-1} |N_j| =: \bar{H}_j. \quad (3.31)$$

Teniendo en cuenta las expresiones (3.29), (3.30) y (3.31)

$$\begin{aligned} |T_j^{(l)}| &\leq |H_j|^{\nu^{(l,j)}} + \omega \left(\sum_{i=0}^{\nu^{(l,j)}-1} |H_j|^i \right) |M_j^{-1}| |G_j| \\ &\leq H_j^{\nu^{(l,j)}} + \omega \left(\sum_{i=0}^{\nu^{(l,j)}-1} H_j^i \right) \langle M_j \rangle^{-1} |G_j| \\ &=: T_j^{(l)}. \end{aligned}$$

Así, las matrices

$$T^{(l)} = \sum_{j=1}^r E_j T_j^{(l)}, \quad (3.32)$$

son las matrices de iteración del método de multipartición en dos etapas, para la matriz $\langle A \rangle = \langle F_j \rangle - |G_j|$, con particiones externas $\langle F_j \rangle - |G_j|$, $j = 1, 2, \dots, r$, y particiones internas $\langle F_j \rangle = \langle M_j \rangle - |N_j|$, $j = 1, 2, \dots, r$, con $0 < \omega \leq 1$.

Estas particiones verifican las siguientes condiciones.

(i) Como A es una H -matriz, por el apartado (a) del Lema 5

$$\langle A \rangle^{-1} \geq |A^{-1}| \geq O.$$

(ii) Como $A = F_j - N_j$, son particiones H -compatibles y A es una H -matriz, por el apartado (b) del Lema 6, $A = F_j - G_j$, son H -particiones y teniendo en cuenta el apartado (a) del mismo lema, deducimos que F_j , $j = 1, 2, \dots, r$, son H -matrices. Luego por el apartado (a) del Lema 5

$$\langle F_j \rangle^{-1} \geq |F_j^{-1}| \geq O, \quad j = 1, 2, \dots, r.$$

Como además $|G_j| \geq O$, $j = 1, 2, \dots, r$, se tiene que

$$\langle A \rangle = \langle F_j \rangle - |G_j|, \quad j = 1, 2, \dots, r,$$

son particiones regulares.

(iii) En (3.30) hemos demostrado que $|M_j| \leq \langle M_j \rangle^{-1}$, $j = 1, 2, \dots, r$. Por tanto

$$\langle M_j \rangle^{-1} \geq 0, \quad j = 1, 2, \dots, r,$$

luego, $\langle M_j \rangle^{-1} |N_j| \geq 0$, $j = 1, 2, \dots, r$ y como además $|N_j| \geq 0$, obtenemos que las particiones

$$\langle F_j \rangle = \langle M_j \rangle - |N_j|, \quad j = 1, 2, \dots, r,$$

son débilmente regulares.

De (i), (ii) y (iii) podemos afirmar que las matrices $T^{(l)}$, $l = 0, 1, \dots$, son las matrices de iteración de un método de multipartición en dos etapas que satisfacen todas las hipótesis del Teorema 4.

Por tanto, según la demostración del Teorema 4 podemos afirmar

$$\lim_{l \rightarrow \infty} \prod_{i=0}^l T^{(i)} = 0.$$

Ahora bien, si tenemos en cuenta las desigualdades

$$0 \leq \left| \prod_{i=0}^l T^{(i)} \right| \leq \prod_{i=0}^l |T^{(i)}| \leq \prod_{i=0}^l T^{(i)},$$

podemos afirmar que

$$\lim_{l \rightarrow \infty} \prod_{i=0}^l T^{(i)} = 0$$

y por tanto, la sucesión de vectores error converge al vector cero. Es decir, el esquema (3.19) es convergente. \blacksquare

Es evidente que el esquema planteado en [12] y que se ha expuesto en la introducción de este Capítulo, es un caso particular del esquema iterativo (3.19).

Para ello, basta considerar las particiones externas $A = F_j - G_j$, todas iguales entre sí, en concreto

$$F_j = D, \quad G_j = B,$$

donde la matriz $D = \text{diag}(D_{11}, \dots, D_{rr})$ es la matriz diagonal por bloques de manera que los bloques diagonales coinciden con los de A (ver expresión (3.2)), y B está definida en (3.6).

La solución aproximada a los subsistemas definidos en (3.4) se calcularía utilizando las particiones internas $D_{jj} = M_j - N_j$, $j = 1, 2, \dots, r$.

En este caso las matrices de iteración son

$$T^{(l)} = \sum_{j=1}^r E_j \left[H^{q(l,j)} + \omega \left(\sum_{i=0}^{q(l,j)-1} H^i \right) M^{-1} B \right], \quad (3.33)$$

donde E_j , $1 \leq j \leq r$, son matrices diagonales por bloques que siguen la estructura de D y tienen el bloque j -ésimo igual a la identidad y el resto son nulos; la matriz H es la siguiente

$$H = (1 - \omega)I + \omega M^{-1} N$$

y

$$M = \text{diag}(M_1, M_2, \dots, M_r) \quad \text{y} \quad N = \text{diag}(N_1, N_2, \dots, N_r).$$

El Teorema 5 nos da nuevas condiciones de convergencia para este caso particular. Además podemos enunciar el correspondiente teorema suponiendo que la partición externa sólo sea H -partición.

Teorema 6. *Supongamos que A es una H -matriz, con $A = D - B$ una H -partición, y $D = \text{diag}(D_{11}, \dots, D_{rr})$. Sean $D_{jj} = M_j - N_j$, $1 \leq j \leq r$, particiones H -compatibles. Si $0 < \omega \leq 1$ y $q(l, j) \geq 1$, $1 \leq j \leq r$, $l = 0, 1, \dots$,*

entonces la sucesión de vectores generada por el esquema iterativo caótico (3.7) converge a la única solución del sistema $Ax = b$.

Demostración. Al igual que hemos hecho en la demostración del Teorema 5, las matrices de iteración dadas en (3.33) pueden ser analizadas de la siguiente forma

$$|T^{(l)}| \leq \sum_{j=1}^r E_j \left[|H|^{r^{(l,j)}} + \omega \left(\sum_{i=0}^{r^{(l,j)}-1} |H|^i \right) \langle M \rangle^{-1} |B| \right] =: T^{(l)},$$

con

$$H = (1 - \omega)I + \omega \langle M \rangle^{-1} |N|.$$

Las matrices $T^{(l)}$ corresponden a las matrices de iteración de un proceso iterativo equivalente al planteado en la expresión (3.7), para la matriz monótona $\langle D \rangle - |B|$ con la partición externa $\langle D \rangle - |B|$ y las particiones internas $\langle D_{jj} \rangle = \langle M_j \rangle - |N_j|$, $j = 1, 2, \dots, r$.

La demostración de que estas particiones y matrices verifican las condiciones del Teorema 2 (o equivalentemente, del Teorema 4, al ser éste una generalización) es análoga a la realizada en el Teorema 5. Por tanto, tenemos demostrada la convergencia. ■

Observación 2. Los teoremas anteriores se pueden demostrar considerando r parámetros de relajación $\omega_1, \omega_2, \dots, \omega_r$, asociados a cada operador P_j , $1 \leq j \leq r$, $l = 0, 1, \dots$. Esta generalización no afecta a los fundamentos teóricos de las respectivas demostraciones.

3.3 Modelo asíncrono

En esta sección desarrollamos el estudio del modelo asíncrono correspondiente al esquema (3.19) o equivalentemente (3.17) que fue expuesto en la sección anterior. El procedimiento iterativo que planteábamos venía expresado mediante el siguiente esquema

$$\mathbf{x}^{(l+1)} = \sum_{j=1}^r E_j P_{l_j}^{q(l,j)} \mathbf{x}^{(l)}, \quad l = 0, 1, 2, \dots,$$

donde las matrices E_j son no negativas y satisfacen $\sum_{j=1}^r E_j = I$, el parámetro $q(l, j)$ indica el número de veces que el procesador j -ésimo actualiza parte de las componentes del vector $\mathbf{x}^{(l+1)}$ y P_{l_j} son r operadores, para cada iteración externa l , que asignan el trabajo al j -ésimo procesador y definidos como

$$P_{l_j} : \mathbf{R}^n \longrightarrow \mathbf{R}^n$$

$$P_{l_j} \mathbf{x} = H_j \mathbf{x} + \omega M_j^{-1} (G_j \mathbf{x}^{(l)} + \mathbf{b}),$$

siendo

$$H_j = (1 - \omega)I + \omega M_j^{-1} N_j,$$

y ω el factor de relajación.

Si suponemos que $\mathcal{F}_j = \{i_j\}_{j=1}^{\infty}$, $1 \leq i_j \leq r$, es una sucesión regulada (ver Definición 5 del Capítulo 2), para un vector inicial, podemos plantear el siguiente modelo asíncrono basado en el esquema caótico de multipartición en dos etapas (3.17)

$$\mathbf{x}^{(j+r)} = (I - E_{i_j}) \mathbf{x}^{(j+r,-1)} + E_{i_j} \sum_{i=1}^r E_i P_{i_j}^{q(i,j)} \mathbf{x}^{(j)}, \quad j = 0, 1, 2, \dots \quad (3.34)$$



siendo r_j el menor entero positivo tal que

$$i_j = i_{j+r_j}.$$

Por la propia definición, como $\{i_j\}_{j=1}^{\infty}$ es regulada debe existir un entero positivo K que verifique

$$0 \leq r_j - 1 < K, \quad (3.35)$$

El valor de $r_j - 1$ es igual al número de veces que la aproximación que reside en un procesador, que denominamos *central*, es actualizada por otros procesadores distintos del i_j -ésimo, durante el intervalo de tiempo en el que dicho procesador calcula su iteración local.

A continuación estudiamos la convergencia de este modelo asíncrono bajo las mismas hipótesis que en el caso síncrono.

Teorema 7. Sea $\{F_i, G_i, M_i, N_i, E_i\}_{i=1}^r$ una multipartición en dos etapas de la matriz del sistema lineal $Ax = b$ y $\{i_j\}_{j=0}^{\infty}$, $1 \leq i_j \leq r$, una sucesión de enteros regulada. Supongamos que existe un vector positivo x y una constante real β tal que las matrices de iteración $T^{(j)}$ definidas en (3.24) verifican

$$|T^{(j)}|x \leq \beta x, \quad 0 \leq \beta < 1, \quad \forall j = 0, 1, \dots \quad (3.36)$$

Entonces, el método de multipartición en dos etapas caótico asíncrono (3.34) converge a la solución ξ del sistema lineal $Ax = b$, para cualquier vector inicial.

Demostración. Para analizar la convergencia del esquema asíncrono (3.34) es conveniente sumergir este procedimiento iterativo en un procedimiento iterativo en \mathbf{R}^{nK} , siendo K el entero positivo que regula a la sucesión $\{i_j\}_{j=1}^{\infty}$ en la Definición 5 del Capítulo 2.

Para tal propósito, introduciremos la siguiente notación.

$$\epsilon^{(j)} = \mathbf{x}^{(j)} - \xi, \quad \bar{\epsilon}_j = \begin{bmatrix} \epsilon^{(j)} \\ \epsilon^{(j-1)} \\ \vdots \\ \epsilon^{(j-K+1)} \end{bmatrix} \in \mathbb{R}^{nK},$$

y

$$\bar{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ \mathbf{x} \\ \vdots \\ \mathbf{x} \end{bmatrix} \in \mathbb{R}^{nK},$$

donde \mathbf{x} es el vector positivo dado en el enunciado del teorema y satisface (3.36).

Por hipótesis, la sucesión $\{r_j\}_{j=1}^{\infty}$ es regulada, por tanto

$$0 \leq r_j - 1 < K, \quad j = 0, 1, 2, \dots$$

De la primera desigualdad de esta expresión obtenemos

$$j \leq j + r_j - 1, \quad j = 0, 1, 2, \dots \quad (3.37)$$

Por la segunda desigualdad

$$\begin{aligned} r_j - K &< 1 \\ r_j - K &\leq 0 \\ j + r_j - K &\leq j, \quad j = 0, 1, 2, \dots \end{aligned} \quad (3.38)$$

Uniando las expresiones (3.37) y (3.38) se tiene

$$j + r_j - K \leq j \leq j + r_j - 1, \quad j = 0, 1, 2, \dots,$$

lo que nos permite asegurar que el vector n -dimensional $\epsilon^{(j)}$ es un bloque del vector nK -dimensional $\bar{\epsilon}_{j+r,-1}$, es decir,

$$\bar{\epsilon}_{j+r,-1} = \begin{bmatrix} \epsilon^{(j+r,-1)} \\ \vdots \\ \epsilon^{(j)} \\ \vdots \\ \epsilon^{(j+r,-K)} \end{bmatrix}.$$

Notemos que el vector $\epsilon^{(j)}$ ocupa la posición r_j en el anterior vector, por tanto, si denotamos por S_j , $j = 0, 1, 2, \dots$, a la matriz de tamaño $n \times nK$ dividida en bloques de tamaño $n \times n$ y que tiene un bloque identidad de orden n en la posición r_j y los $K - 1$ bloques restantes son nulos, es decir

$$S_j = \begin{bmatrix} \overbrace{O_n \cdots O_n}^{r_j-1 \text{ bloques}} & I_n & O_n & \cdots & O_n \end{bmatrix}, \quad (3.39)$$

podemos escribir

$$\epsilon^{(j)} = S_j \bar{\epsilon}_{j+r,-1}. \quad (3.40)$$

Analizamos seguidamente la expresión del error

$$\epsilon^{(j+r_j)} = \mathbf{x}^{(j+r_j)} - \xi.$$

Puesto que ξ es un punto fijo del esquema (3.17), la expresión del error quedará como

$$\begin{aligned} \epsilon^{(j+r_j)} &= \mathbf{x}^{(j+r_j)} - \xi \\ &= (I - E_{i_j}) \mathbf{x}^{(j+r_j,-1)} + E_{i_j} \sum_{i=1}^r E_i P_{j_i}^{q(j,i)} \mathbf{x}^{(j)} \end{aligned}$$

$$\begin{aligned}
 & - (I - E_i)\xi - E_i, \sum_{i=1}^r E_i P_{ji}^{q(j,i)} \xi \\
 & = (I - E_i)\epsilon^{(j+r,-1)} + E_i, \sum_{i=1}^r E_i P_{ji}^{q(j,i)} (\mathbf{x}^{(j)} - \xi)
 \end{aligned}$$

Ahora bien, la expresión $\sum_{i=1}^r E_i P_{ji}^{q(j,i)} (\mathbf{x}^{(j)} - \xi)$ fue analizada en el Lema 2, demostrando que

$$\sum_{i=1}^r E_i P_{ji}^{q(j,i)} (\mathbf{x}^{(j)} - \xi) = T^{(j)} \epsilon^{(j)}. \quad (3.41)$$

Teniendo en cuenta las expresiones (3.40) y (3.41), podemos escribir

$$\begin{aligned}
 \epsilon^{(j+r)} & = (I - E_i)\epsilon^{(j+r,-1)} + E_i, \sum_{i=1}^r E_i P_{ji}^{q(j,i)} (\mathbf{x}^{(j)} - \xi) \\
 & = (I - E_i)\epsilon^{(j+r,-1)} + E_i, T^{(j)} \epsilon^{(j)} \\
 & = (I - E_i)\epsilon^{(j+r,-1)} + E_i, T^{(j)} S_j \bar{\epsilon}_{j+r,-1}.
 \end{aligned}$$

Denotemos por B_{j+r} , la matriz cuadrada de orden nK , dada por

$$B_{j+r} = \begin{bmatrix} I - E_i & O & \dots & O & O \\ I & O & \dots & O & O \\ \vdots & \vdots & & \vdots & \vdots \\ O & O & \dots & I & O \end{bmatrix} + \begin{bmatrix} E_i, T^{(j)} S_j \\ O \\ \vdots \\ O \end{bmatrix}, \quad (3.42)$$

donde la matriz del primer sumando está dividida en bloques de tamaño $n \times n$ y la matriz del segundo sumando, en bloques de tamaño $n \times nK$.

Calculemos el producto $B_{j+r} \bar{\epsilon}_{j+r,-1}$.

$$B_{j+r} \bar{\epsilon}_{j+r,-1} = \left\{ \begin{bmatrix} I - E_i & O & \dots & O & O \\ I & O & \dots & O & O \\ \vdots & \vdots & & \vdots & \vdots \\ O & O & \dots & I & O \end{bmatrix} + \begin{bmatrix} E_i, T^{(j)} S_j \\ O \\ \vdots \\ O \end{bmatrix} \right\} \begin{bmatrix} \epsilon^{(j+r,-1)} \\ \vdots \\ \epsilon^{(j+r,-K-1)} \\ \epsilon^{(j+r,-K)} \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} (I - E_{i_j})\epsilon^{(j+r_j-1)} + E_{i_j}T^{(j)}S_j\bar{\epsilon}_{j+r_j-1} \\ \epsilon^{(j+r_j-1)} \\ \vdots \\ \epsilon^{(j+r_j-K+1)} \end{bmatrix} \\
&= \begin{bmatrix} \epsilon^{(j+r_j)} \\ \epsilon^{(j+r_j-1)} \\ \vdots \\ \epsilon^{(j+r_j-K+1)} \end{bmatrix} \\
&= \bar{\epsilon}_{j+r_j}.
\end{aligned}$$

Obtenemos, por tanto, que

$$\bar{\epsilon}_{j+r_j} = B_{j+r_j}\bar{\epsilon}_{j+r_j-1}.$$

Entonces

$$\bar{\epsilon}_{j+2K-1} = B_{j+2K-1}B_{j+2K-2}\cdots B_{j+1}\bar{\epsilon}_j.$$

Estamos interesados en encontrar una constante real γ , con $0 \leq \gamma < 1$, tal que

$$\|B_{j+2K-1}B_{j+2K-2}\cdots B_{j+1}\bar{x}\|_{\bar{x}} \leq \gamma\bar{x}. \quad (3.43)$$

Si esto fuera cierto, considerando la norma vectorial $\|\cdot\|_{\bar{x}}$ definida en (1.3) y su norma matricial inducida, al ser ésta monótona se tiene que

$$\|B_{j+2K-1}B_{j+2K-2}\cdots B_{j+1}\bar{x}\|_{\bar{x}} \leq \|\gamma\bar{x}\|_{\bar{x}},$$

y así

$$\|B_{j+2K-1}B_{j+2K-2}\cdots B_{j+1}\|_{\bar{x}} \leq \gamma,$$



y entonces

$$\lim_{\nu \rightarrow \infty} \tau_\nu = 0,$$

con lo que quedaría demostrada la convergencia del esquema (3.34).

Sea

$$\omega_\nu = |B_{j+\nu} B_{j+\nu-1} \cdots B_{j+1}| \bar{x}.$$

Demostraremos que existe una constante real $0 \leq \gamma < 1$, que satisface

$$\omega_{2K-1} \leq \gamma \bar{x}. \quad (3.44)$$

Claramente

$$\omega_\nu \leq |B_{j+\nu}| |B_{j+\nu-1}| \cdots |B_{j+1}| \bar{x} =: v_\nu = \begin{bmatrix} v_\nu^1 \\ v_\nu^2 \\ \vdots \\ v_\nu^K \end{bmatrix}, \quad (3.45)$$

siendo

$$v_\nu^s \in \mathbf{R}_+, \quad 1 \leq s \leq K.$$

Por tanto, demostrar (3.43) o equivalentemente (3.44) se traduce en probar

$$v_{2K-1}^s \leq \gamma \bar{x}, \quad 1 \leq s \leq K.$$

Demostraremos, en primer lugar que

$$v_\nu^s \leq \gamma \bar{x}, \quad \nu \geq K.$$

Como \bar{x} es un vector que satisface (3.36), de la forma de las matrices S_j en

(3.39) y de la definición de B_i en (3.42) se obtienen las siguientes desigualdades

$$\begin{aligned}
 |B_{j+1}| \bar{x} &\leq \left\{ \begin{bmatrix} I - E_i & O & \dots & O & O \\ I & O & \dots & O & O \\ \vdots & \vdots & & \vdots & \vdots \\ O & O & \dots & I & O \end{bmatrix} + \begin{bmatrix} E_i |T^{(j)}| S_j \\ O \\ \vdots \\ O \end{bmatrix} \right\} \begin{bmatrix} x \\ x \\ \vdots \\ x \end{bmatrix} \\
 &= \begin{bmatrix} (I - E_i)x + E_i |T^{(j)}| S_j \bar{x} \\ x \\ \vdots \\ x \end{bmatrix} \\
 &= \begin{bmatrix} (I - E_i)x + E_i |T^{(j)}| x \\ x \\ \vdots \\ x \end{bmatrix} \\
 &\leq \begin{bmatrix} (I - E_i)x + E_i \beta x \\ x \\ \vdots \\ x \end{bmatrix}
 \end{aligned}$$

Como $0 \leq \beta < 1$, se tiene

$$(I - E_i)x + E_i \beta x \leq (I - E_i)x + E_i x = x,$$

y por tanto

$$|B_{j+\nu}| \bar{x} \leq \begin{bmatrix} x \\ x \\ \vdots \\ x \end{bmatrix} = \bar{x}, \quad \forall j = 0, 1, 2, \dots, \nu \geq 1.$$

Entonces, claramente por la definición de v_ν en (3.45), se deduce

$$v_\nu \leq \bar{x}.$$

A partir de esta última desigualdad, puesto que $v_\nu = |B_{j+\nu}| v_{\nu-1}$, podemos, utilizando la expresión (3.36) y la forma de S_j en (3.39), acotar el primer bloque del vector v_ν de la siguiente forma

$$\begin{aligned} v_\nu^1 &\leq (I - E_p) v_{\nu-1}^1 + E_p |T^{(j)}| S_j v_{\nu-1} & (3.46) \\ &\leq (I - E_p) v_{\nu-1}^1 + E_p |T^{(j)}| S_j \bar{x} \\ &= (I - E_p) v_{\nu-1}^1 + E_p |T^{(j)}| x \\ &\leq (I - E_p) x + \beta E_p x \\ &= x + (\beta - 1) E_p x. & (3.47) \end{aligned}$$

Vamos a analizar cada una de las componentes del vector v_ν^1 . Sea $\sigma \in \{1, 2, \dots, n\}$; distinguiremos dos casos:

(a) Si el elemento diagonal que ocupa la posición σ de E_p es nulo, es decir, si $(E_p)_{\sigma\sigma} = 0$, entonces por la expresión (3.46) obtenemos

$$(v_\nu^1)_\sigma \leq (v_{\nu-1}^1)_\sigma.$$

3.3 Modelo asíncrono

97

(b) Si $(E_\nu)_{\sigma\sigma} = \alpha$, con $0 < \alpha \leq 1$, entonces por la ecuación (3.47) se sigue que

$$(v_\nu^1)_\sigma \leq x_\sigma + (\beta - 1)\alpha x_\sigma = (1 + (\beta - 1)\alpha)x_\sigma.$$

Si $\gamma = 1 + (\beta - 1)\alpha < 1$, entonces

$$(v_\nu^1)_\sigma \leq \gamma x_\sigma.$$

Ahora bien, como la sucesión $\{i_j\}_{j=1}^\infty$ es regulada, todos los enteros $1, 2, \dots, r$, aparecen en el subconjunto $\{i_j + 1, \dots, i_j + \nu\}$, $\nu \geq K$, puesto que las matrices E_j son no negativas y satisfacen que $\sum_{j=1}^r E_j = I$ y teniendo en cuenta los dos casos anteriores, obtenemos, para todo $\sigma = 1, 2, \dots, n$

$$(v_\nu^1)_\sigma \leq \gamma x_\sigma, \text{ si } \nu \geq K.$$

Entonces

$$v_\nu^1 \leq \gamma x, \quad \nu \geq K. \quad (3.48)$$

Teniendo en cuenta que

$$v_{\nu+1} = \begin{bmatrix} v_{\nu+1}^1 \\ v_{\nu+1}^2 \\ \vdots \\ v_{\nu+1}^K \end{bmatrix} = |B_{j+\nu+1}| v_\nu = \begin{bmatrix} * \\ v_\nu^1 \\ \vdots \\ v_\nu^{K-1} \end{bmatrix},$$

se tiene, para todo $\nu = 0, 1, 2, \dots$,

$$v_{\nu+1}^{1+s} = v_\nu^s, \quad s = 1, 2, \dots, K-1,$$

de donde

$$v_{2K-1}^{1+s} = v_{2K-2}^s = v_{2K-3}^{s-1} = \dots = v_{2K-s-1}^1, \quad s = 1, 2, \dots, K-1,$$

y así

$$v_{2K-1}^{1+s} = v_{2K-s-1}^1, \quad s = 1, 2, \dots, K-1. \quad (3.49)$$

Finalmente, como $2K-s-1 \geq K$ para $s = 1, 2, \dots, K-1$, por las ecuaciones (3.48) y (3.49) obtenemos

$$v_{2K-1}^{1+s} \leq \gamma x, \quad s = 0, 1, \dots, K-1.$$

Esto es

$$v_{2K-1} \leq \gamma \bar{x},$$

pero entonces, por (3.45), se cumple

$$\omega_{2K-1} \leq v_{2K-1} \leq \gamma \bar{x}, \quad 0 \leq \gamma < 1.$$

Por tanto la convergencia del esquema asíncrono (3.34) queda demostrada. ■

Observación 3. Con el Teorema 7 tenemos demostrada la convergencia del esquema (3.34) bajo las mismas hipótesis que en el Teorema 3. En este teorema se prueba que la norma infinito de las matrices de iteración $T^{(l)}$, $l = 0, 1, \dots$, es menor que 1. Por tanto, si tomamos el vector $x = (1, 1, \dots, 1)^t$ y consideramos $T^{(l)} = [t_{ij}^{(l)}]_{1 \leq i, j \leq n}$, $l = 0, 1, \dots$, obtenemos

$$|T^{(l)}|x = \begin{bmatrix} \sum_{j=1}^n |t_{1j}^{(l)}| \\ \sum_{j=1}^n |t_{2j}^{(l)}| \\ \vdots \\ \sum_{j=1}^n |t_{nj}^{(l)}| \end{bmatrix} \leq \begin{bmatrix} \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^n |t_{ij}^{(l)}| \right\} \\ \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^n |t_{ij}^{(l)}| \right\} \\ \vdots \\ \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^n |t_{ij}^{(l)}| \right\} \end{bmatrix}$$



$$\begin{aligned}
 &= \begin{bmatrix} \|T^{(0)}\|_{\infty} \\ \|T^{(0)}\|_{\infty} \\ \vdots \\ \|T^{(0)}\|_{\infty} \end{bmatrix} \leq \begin{bmatrix} \beta \\ \beta \\ \vdots \\ \beta \end{bmatrix} \\
 &= \beta \mathbf{x}, \quad 0 \leq \beta < 1, \quad \forall l = 0, 1, \dots,
 \end{aligned}$$

donde β es una constante real que verifica $\|T^{(l)}\|_{\infty} \leq \beta < 1$, $l = 0, 1, \dots$

Esto demuestra la desigualdad (3.36) y por tanto el esquema asíncrono (3.34) converge bajo las hipótesis del Teorema 3.

Las correspondientes versiones asíncronas de los Teoremas 4, 5 y 6 también se verifican ya que en la demostración de estos teoremas se prueba explícitamente la existencia de un vector positivo y una constante real en el intervalo $[0, 1)$ que satisface la condición (3.36).

Observación 4. Hacemos notar, que al igual que en el caso síncrono, los teoremas correspondientes al caso asíncrono pueden ser demostrados considerando r parámetros de relajación $\omega_1, \omega_2, \dots, \omega_r$, asociados a cada uno de los operadores P_{lj} , $1 \leq j \leq r$, $l = 0, 1, \dots$

3.4 Experimentos numéricos

En esta sección estudiamos desde el punto de vista experimental distintos algoritmos caóticos, tanto síncronos como asíncronos para la resolución de un

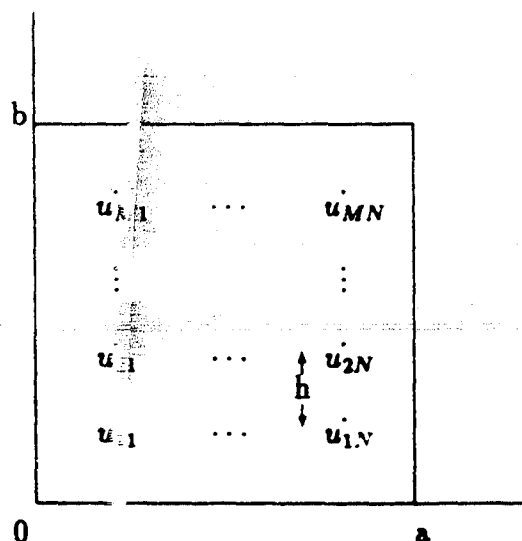
sistema de ecuaciones lineales. Dichos algoritmos están basados en los esquemas teóricos (3.17) y (3.34) estudiados en las Secciones 3.2 y 3.3, respectivamente.

Para ilustrar los esquemas caóticos estudiados tanto en este capítulo como en los posteriores, tomaremos como problema modelo la obtención de una solución aproximada de la ecuación de Laplace, que aparece en multitud de problemas físicos y técnicos. Dicha ecuación relaciona las parciales de segundo orden de una función de dos variables $u(x, y)$ de la siguiente forma

$$\nabla^2 u = u_{xx} + u_{yy} = 0.$$

Buscaremos una solución definida en la región o dominio $\Omega = [0, a] \times [0, b]$ y que verifique determinadas condiciones de contorno, es decir, se conocen los valores de la función u en la frontera de Ω .

Para obtener una aproximación a dicha ecuación se puede utilizar el método de diferencias finitas. Para ello discretizamos el rectángulo Ω con una malla de puntos separados constantemente por h de la forma



Entonces, reemplazando las derivadas por cocientes de diferencias en cada nodo (x_i, y_j) en que se divide la región, se obtiene

$$\begin{aligned} \nabla^2 u(x_i, y_j) &= \frac{u(x_{i+1}, y_j) - 2u(x_i, y_j) + u(x_{i-1}, y_j))}{h^2} \\ &+ \frac{u(x_i, y_{j+1}) - 2u(x_i, y_j) + u(x_i, y_{j-1}))}{h^2} = 0. \end{aligned}$$

Si llamamos $u(x_i, y_j) = u_{ij}$ obtenemos

$$\nabla^2 u_{ij} = \frac{1}{h^2} [u_{i+1,j} + u_{i-1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{ij}] = 0.$$

Entonces se obtiene el sistema

$$Ax = b,$$

donde

$$x = (u_{11} \dots u_{1N} \ u_{21} \dots u_{2N} \dots u_{M1} \dots u_{MN})^t,$$

y la matriz A de tamaño $MN \times MN$, tiene la estructura

$$A = \begin{bmatrix} B & -I & & & \\ -I & B & -I & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -I \\ & & & -I & B \end{bmatrix} \quad (3.50)$$

con los bloques identidad I de tamaño $N \times N$ y las M submatrices B , también de tamaño $N \times N$, son

$$B = \begin{bmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 4 \end{bmatrix} \quad (3.51)$$

El vector \mathbf{b} consiste en los valores conocidos de la función u en la frontera del rectángulo Ω .

Antes de discutir los resultados numéricos obtenidos, vamos a detallar las implementaciones de los esquemas iterativos caóticos tanto síncrono como asíncrono, estudiados en este capítulo.

3.4.1 Algoritmos

Consideremos el sistema lineal $A\mathbf{z} = \mathbf{b}$ y las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$. Se desea resolver este sistema de ecuaciones lineales concurrentemente utilizando los esquemas caóticos paralelos síncrono y asíncrono, descritos en las expresiones (3.17) y (3.34) respectivamente. Supondremos que se dispone de r procesadores conectados en paralelo.

La diferencia básica entre el algoritmo caótico síncrono y el algoritmo caótico asíncrono es la siguiente: en el primero, la parte principal la constituye la etapa 2, denominada *etapa caótica*, en la que cada procesador j actualiza sólo aquellas componentes del vector iterado $\mathbf{z}^{(k+1)}$ que corresponden a elementos no nulos de la matriz E_j . Una vez que todos los procesadores han finalizado sus cálculos se combinan en un mismo instante (fase síncrona), para formar el nuevo vector iterado global $\mathbf{z}^{(k+1)}$; esto corresponde a la etapa 3 del algoritmo síncrono. Por el contrario, en el algoritmo asíncrono, no es necesario que se produzca dicha sincronización entre los procesadores.

ALGORITMO SÍNCRONO**Etapa 0. Lectura de datos:**Matriz de coeficientes: A Término independiente: b Factores caóticos: $q(l, j), 1 \leq j \leq r, l = 0, 1, \dots$ Factor de relajación: ω **Etapa 1. Elegir $x^{(0)}$ arbitrariamente y $l = 0$.****Etapa 2. (Etapa caótica) En paralelo,**Para $j = 1, 2, \dots, r$

$$c_j = G_j x^{(l)} + b_j.$$

Realizar $q(l, j)$ etapas de un método iterativo extrapolado para obtener una solución aproximada del sistema

$$F_j z_j = c_j, \text{ con } z_j^{(0)} = x^{(l)}.$$

Etapa 3. (Etapa sincrónica) $x^{(l+1)} = \sum_{j=1}^r E_j z_j^{(q(l, j))}$.**Etapa 4. Comprobación del criterio de convergencia**

$$\sum_{k=1}^n |x_k^{(l+1)} - x_k^{(l)}| < 5 \cdot 10^{-10}.$$

- Si test de convergencia = VERDADERO, entonces FIN.

- Si test de convergencia = FALSO, entonces

- $l \leftarrow l + 1$.

- Volver a la etapa 2.

ALGORITMO ASÍNCRONO

Etapa 0. Lectura de datos:

Matriz de coeficientes: A Término independiente: b Factores caóticos: $q(l, j), 1 \leq j \leq r, l = 0, 1, \dots$ Factor de relajación: ω

Etapa 1. Elegir $\mathbf{x}^{(0)}$ arbitrariamente y $l_j = 0, 1 \leq j \leq r$. En paralelo, para $j = 1, 2, \dots, r$ (etapas 2-5).

Etapa 2. $\mathbf{x}^{(l_j)} = \mathit{comm}$, donde comm es un vector compartido por todos los procesadores.

Etapa 3. $\mathbf{c}_j = G_j \mathbf{x}^{(l_j)} + \mathbf{b}$.

Realizar $q(l_j, j)$ etapas de un método iterativo extrapolado para obtener una solución aproximada del sistema

$$F_j \mathbf{z}_j = \mathbf{c}_j, \text{ con } \mathbf{z}_j^{(0)} = \mathbf{x}^{(l_j)}.$$

Etapa 4. $\mathit{comm} = (I - E_j) \mathit{comm} + E_j \mathbf{z}_j^{(q(l_j, j))}$, siendo E_j , matrices diagonales no negativas cuya suma es la identidad.

Etapa 5. Comprobación del criterio de convergencia.

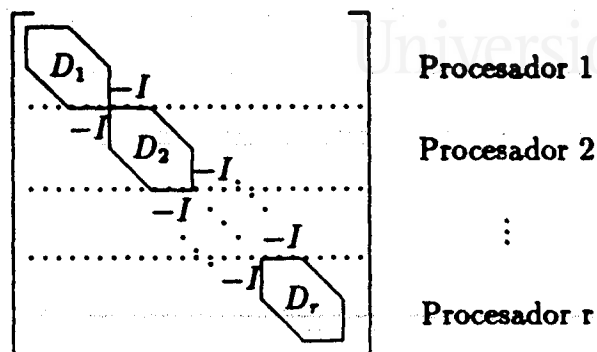
- Si test de convergencia = VERDADERO, entonces FIN.
- Si test de convergencia = FALSO, entonces
 - $l_j \leftarrow l_j + 1$.
 - Volver a la etapa 2.

**3.4.2 Resultados numéricos**

La evaluación tanto de los algoritmos introducidos previamente como los que se estudiarán en los restantes capítulos, se ha realizado atendiendo a varios criterios.

- (i) Encontrar el factor óptimo de relajación, es decir, aquel factor para el que la convergencia es más rápida.
- (ii) Comparar los algoritmos caóticos síncronos con los correspondientes asíncronos.
- (iii) Comparar los algoritmos caóticos tanto síncronos como asíncronos, respecto de los no caóticos.
- (iv) Obtener el incremento de velocidad o “speed-up” en los modelos caóticos planteados.
- (v) Estudio de la variación del tiempo de ejecución atendiendo al equilibrio de la carga asignada a cada uno de los procesadores, mediante los factores caóticos.

Se han utilizado distintos tamaños de la matriz A y de sus bloques diagonales B . La estructura de las matrices A y B viene definida en las expresiones (3.50) y (3.51) respectivamente. Según la interpretación en términos de teoría de grafos del concepto de irreducibilidad que se da en la Sección 4.1, es fácil ver que la matriz A dada en (3.50) es irreduciblemente diagonal dominante (ver Definición 10 de la Sección 1.3). Por tanto, teniendo en cuenta el Teorema 9 de la Sección 1.4, la matriz A es una H -matriz.



La implementación de los algoritmos síncrono y asíncrono se ha realizado obteniendo en paralelo la solución aproximada de los sistemas $F_j z_j = c_j$, $1 \leq j \leq 8$ (ver etapas 2 y 3 de los algoritmos síncrono y asíncrono, respectivamente), realizando $q(l, j)$ pasos del método *SOR*, eligiendo por tanto, la partición correspondiente.

Las particiones así contruidas satisfacen las hipótesis de Teorema 5 para $0 < \omega \leq 1$.

La Tabla 1 explica la forma de las matrices de prueba dando el orden de la matriz A , el orden de las matrices diagonales B , el número de bloques B asignados a cada procesador y el número de filas con las que cada procesador trabaja, o lo que es lo mismo, el orden de las matrices D_i , $i = 1, 2, \dots, 8$. Obsérvese que sólo para la matriz de orden 1024 existe un equilibrio de la carga de trabajo entre los procesadores.

La Tabla 2 presenta los resultados de los distintos algoritmos caóticos síncronos, ejecutados en un sólo procesador y con los datos relativos al factor de relajación óptimo. Como cabía esperar este factor óptimo no coincide para los distintos modelos caóticos de una matriz concreta.

Orden A	Orden B	Bloques por proc.	Filas por proc.
384	32	2,2,2,2,1,1,1,1	64,64,64,64,32,32,32,32
1024	32	4,4,4,4,4,4,4,4	128,128,128,128,128,128,128,128
2816	256	2,2,2,1,1,1,1,1	512,512,512,256,256,256,256,256
5632	512	2,2,2,1,1,1,1,1	1024,1024,1024,512,512,512,512,512

Tabla 1: Matrices de prueba.

Los correspondientes tiempos de ejecución se utilizan como referencia para el cálculo de los incrementos de velocidad de los algoritmos paralelos. Los factores $q(l, j)$ indican el número de iteraciones internas mediante las que el bloque j -ésimo de vector iterado es actualizado en una iteración externa l . En los ejemplos que presentamos dichos factores permanecen constantes en cada iteración externa, es decir, $q(l, j) = q(j)$, $j = 1, 2, \dots, r$.

Observando la Tabla 2, vemos que aunque los algoritmos caóticos secuenciales consiguen reducir el número de iteraciones respecto del no caótico, los tiempos no siempre son menores. Esto es debido a que un excesivo esfuerzo computacional como consecuencia del número de iteraciones internas, en los algoritmos caóticos secuenciales, no compensa la reducción en tiempo que supone disminuir el número de iteraciones externas. Sin embargo, cuando estos algoritmos se ejecutan con más procesadores, tiene sentido el realizar más iteraciones internas por parte de aquellos procesadores que a priori tengan menos trabajo. Así pues, el uso de los algoritmos caóticos tiene, en un principio, más interés desde el punto de vista paralelo que secuencial.



Universitat d'Alacant
Universidad de Alicante

Orden	$q(l, j), 1 \leq j \leq 8$	ω_{opt}	Iter.	Tiempo (seg.)
384	1,1,1,1,2,2,2,2	1.575	480	4.498
	2,2,2,2,4,4,4,4	1.500	426	5.138
	2,2,2,2,2,2,2,2	1.500	431	4.672
	3,3,3,3,6,6,6,6	1.475	432	6.237
	1,1,1,1,1,1,1,1	1.375	536	4.861
1024	3,3,3,3,3,3,3,3	1.725	759	25.106
	2,2,2,2,2,2,2,2	1.700	729	20.875
	1,1,1,1,1,1,1,1	1.650	933	22.530
2816	2,2,2,4,4,4,4,4	1.450	499	44.370
	1,1,1,2,2,2,2,2	1.500	536	37.788
	2,2,2,2,2,2,2,2	1.450	504	39.205
	1,1,1,1,1,1,1,1	1.350	571	37.863
5632	2,2,2,4,4,4,4,4	1.450	515	91.816
	1,1,1,2,2,2,2,2	1.500	555	79.105
	2,2,2,2,2,2,2,2	1.425	522	81.163
	1,1,1,1,1,1,1,1	1.350	591	80.175

Tabla 2: Algoritmo secuencial.

La Tabla 3 resume los resultados obtenidos para los esquemas caóticos síncronos ejecutados en paralelo, utilizando los 8 procesadores de los que dispone el Alliant FX/80. Como en la Tabla 2, se ha reflejado el factor óptimo de relajación.

Se han utilizado varias directivas de compilación. T_1 indica el tiempo de CPU utilizando ejecución concurrente y T_2 el tiempo de CPU cuando se utiliza ejecución vectorial-concurrente. Estos tipos de ejecuciones se han explicado en la Sección 2.1.

Para estos tiempos se ha calculado dos tipos de speed-up. El primero de ellos compara el algoritmo secuencial no caótico (factores caóticos iguales a 1) con el algoritmo caótico paralelo para los correspondientes factores óptimos de relajación, es decir

$$S_i = \frac{T_s}{T_i}, \quad i = 1, 2,$$

donde T_s es el tiempo de CPU del algoritmo secuencial no caótico (ver Tabla 2). Se ha calculado este incremento de velocidad debido a la razón ya mencionada, de que en general los esquemas caóticos están pensados para equilibrar la carga de trabajo entre los procesadores y por tanto ser ejecutados en paralelo y no en secuencial. Puesto que comparamos distintos algoritmos, este speed-up, en principio no está acotado superiormente por el número de procesadores utilizado.

También se ha calculado el incremento de velocidad de acuerdo con la Definición 1 de la Sección 2.1. Dicho speed-up compara, en este caso, un algoritmo caótico paralelo con el mismo algoritmo caótico pero ejecutado en un sólo procesador. En la Tabla 3 este speed-up está denotado por S'_i , $i = 1, 2$, es decir

$$S'_i = \frac{T_{cs}}{T_i}, \quad i = 1, 2,$$

donde T_{cs} es el tiempo del algoritmo caótico secuencial. Nótese que S'_1 si es menor o igual que el número de procesadores utilizado en la ejecución concurrente.

Se observa, en la Tabla 3, que el factor de relajación óptimo para los modelos caóticos es mayor que para el modelo no caótico. Además el rango de convergencia del modelo caótico es más amplio. Esto puede observarse en la Tabla 4, donde además se han comparado distintos esquemas caóticos, utilizando distintas particiones de la matriz de tamaño 1024. Los datos en blanco indican que el esquema no converge para el correspondiente factor de relajación.

Orden	$q(i, j), 1 \leq j \leq 8$	ω_{opt}	Iter	Tiempo (seg.)		Speed-up			
				T_1	T_2	S_1	S_2	S'_1	S'_2
384	1,1,1,1,2,2,2,2	1.575	486	0.825	0.373	5.89	13.02	5.45	12.06
	2,2,2,2,4,4,4,4	1.500	426	0.872	0.427	5.57	11.38	5.89	12.03
	2,2,2,2,2,2,2,2	1.500	431	0.876	0.425	5.59	11.43	5.33	11.00
	3,3,3,3,6,6,6,6	1.475	432	1.029	0.689	4.72	7.05	6.06	9.05
	1,1,1,1,1,1,1,1	1.375	536	0.925	0.426	5.25	11.41	5.25	11.41
1024	3,3,3,3,3,3,3,3	1.725	759	4.098	2.295	5.49	9.82	6.13	10.94
	2,2,2,2,2,2,2,2	1.700	729	3.470	1.660	6.49	13.57	6.02	12.58
	1,1,1,1,1,1,1,1	1.650	933	3.946	1.656	5.70	13.60	5.70	13.60
2816	2,2,2,4,4,4,4,4	1.450	499	7.621	4.242	4.97	8.93	5.82	10.46
	1,1,1,2,2,2,2,2	1.500	536	6.880	2.578	5.50	14.68	5.49	14.66
	2,2,2,2,2,2,2,2	1.450	504	7.566	3.517	5.01	10.77	5.18	11.15
	1,1,1,1,1,1,1,1	1.350	573	7.295	3.650	5.19	10.37	5.19	10.37
5632	2,2,2,4,4,4,4,4	1.450	511	16.379	10.004	4.89	8.01	5.61	9.18
	1,1,1,2,2,2,2,2	1.500	551	14.854	5.438	5.40	14.74	5.33	14.55
	2,2,2,2,2,2,2,2	1.425	523	16.230	7.974	4.93	10.05	5.00	9.92
	1,1,1,1,1,1,1,1	1.350	585	15.826	5.706	5.07	14.05	5.07	14.05

Tabla 3: Resultados de los esquemas caóticos síncronos para ω óptimo.

La Tabla 4 muestra además, la mejora en el tiempo de ejecución de los algoritmos caóticos síncronos cuando la carga de trabajo se intenta equilibrar entre

los distintos procesadores. Cuando los bloques diagonales son de distinto tamaño, se obtienen mejores resultados utilizando factores caóticos que equilibran la carga de trabajo, pero no suponen la realización de un excesivo número de iteraciones internas. Esta situación se refleja de forma clara en las Figuras 1, 2, 3 y 4, donde se compara para cada matriz de prueba distintos esquemas caóticos síncronos, calculando el tiempo de *CPU* en ejecución concurrente atendiendo al factor de relajación.

Bloques por proc.	4,4,4,4,4,4,4		6,6,4,4,4,2,2		8,8,4,4,2,2,2,2		6,6,6,6,2,2,2,2	
Factores caóticos	1,1,1,1,1,1,1		1,1,2,2,2,2,3,3		1,1,2,2,4,4,4,4		1,1,1,1,3,3,3,3	
ω	Iter.	Tiempo	Iter.	Tiempo	Iter.	Tiempo	Iter.	Tiempo
0.5	7949	35.327	5460	26.543	6197	32.735	7466	34.188
0.6	6319	28.118	4355	21.163	4924	26.019	5924	27.123
0.7	5143	22.914	3559	17.318	4007	21.163	4812	22.046
0.8	4252	18.957	2957	14.384	3314	17.429	3977	18.293
0.9	3553	15.791	2487	12.102	2771	14.399	3311	15.216
1.0	2990	13.298	2110	10.250	2336	12.117	2779	12.745
1.1	2525	11.230	1801	8.753	1979	10.278	2342	10.758
1.2	2134	9.491	1544	7.506	1683	8.740	1975	9.069
1.3	1800	8.006	1327	6.449	1434	7.417	1664	7.636
1.4	1512	6.721	1144	5.557	1225	6.354	1398	6.418
1.5	1259	5.619	967	4.781	1051	5.443	1170	5.368
1.55	1144	5.022	917	4.441	975	5.051	1069	4.901
1.6	1036	4.627	853	4.132	907	4.701	977	4.468
1.65	933	3.916	793	3.840	845	4.377	934	4.287
1.675	5140	21.267	765	3.921	816	4.1553	860	3.933
1.7	—	—	737	3.573	789	4.092	2592	11.9012
1.75	—	—	—	—	735	3.8131	—	—
1.8	—	—	—	—	1504	7.7394	—	—

Tabla 4: Comparación de esquemas caóticos síncronos con distintas particiones, $n = 1024$.

Tanto en estas figuras como en las demás de este capítulo y los sucesivos, utilizamos una notación abreviada para los factores caóticos. Así por ejemplo, el esquema caótico $1^4 2^4$ corresponde al esquema caótico 1, 1, 1, 1, 2, 2, 2, 2, es decir, al esquema caótico en el que, en cada iteración externa, para actualizar los 4 primeros bloques del vector iterado se realiza una única iteración interna, mientras que para los 4 últimos se realizan dos.

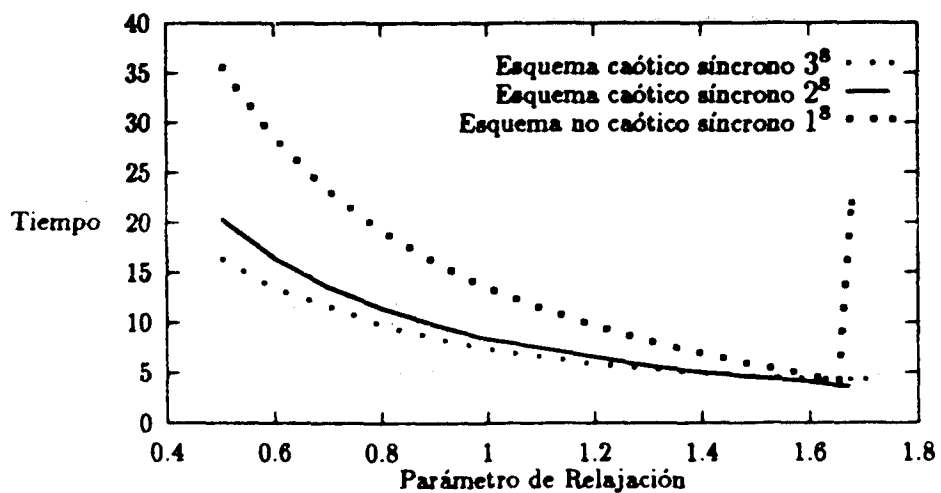


Figura 1: Comparación esquemas paralelos caóticos sincrónicos, $n = 1024$.

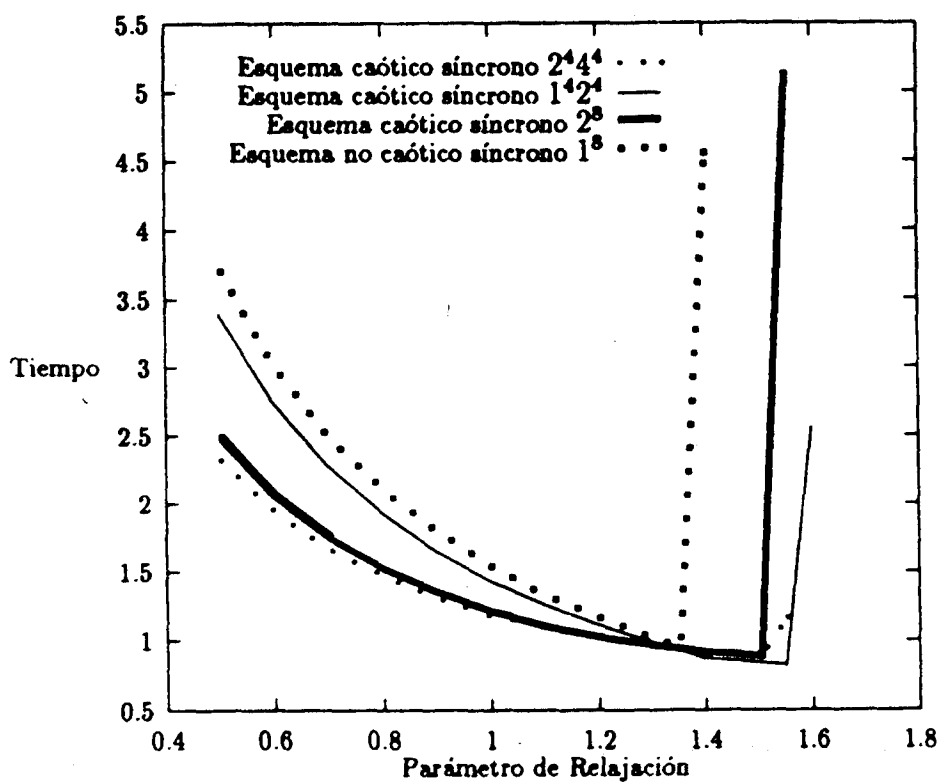


Figura 2: Comparación esquemas paralelos caóticos síncronos, $n = 384$.

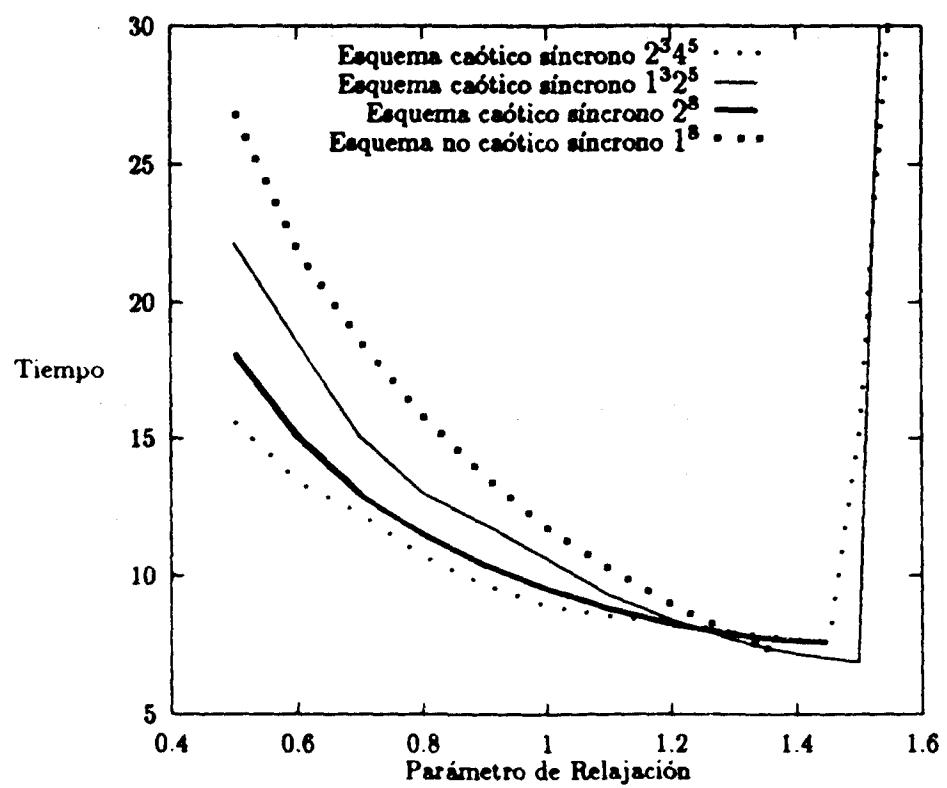


Figura 3: Comparación esquemas paralelos caóticos sincrónicos, $n = 2816$.

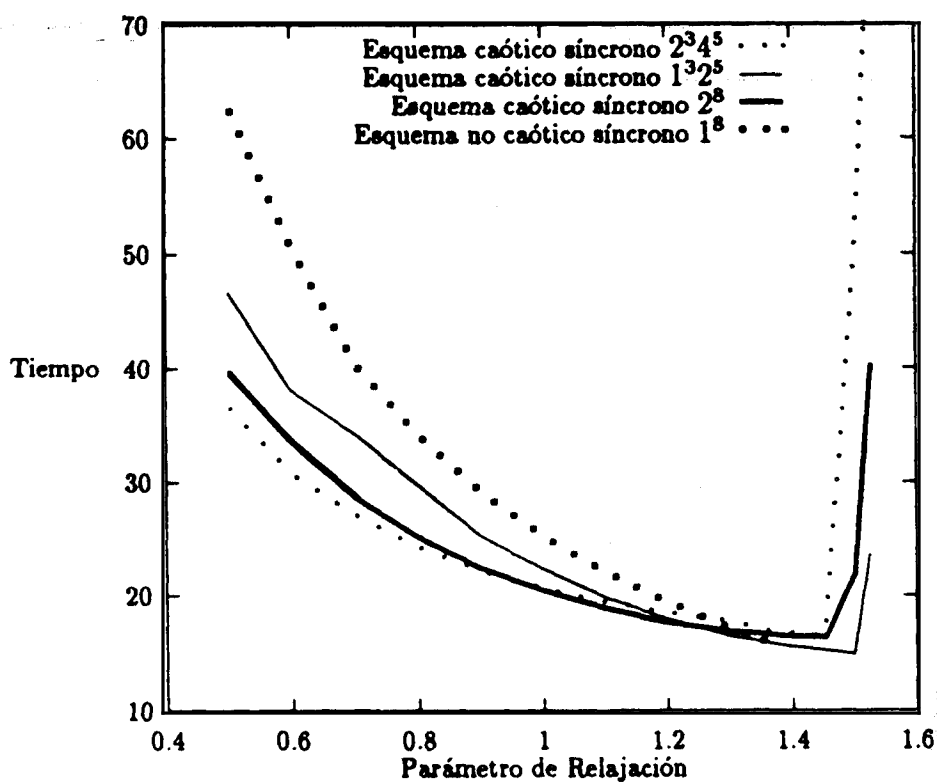


Figura 4: Comparación esquemas paralelos caóticos sincros, $n = 5632$.

Los experimentos realizados han mostrado que los modelos caóticos asíncronos se comportan, independientemente del factor de relajación, de forma más eficiente que el correspondiente modelo caótico síncrono. Esto queda claro en las gráficas expuestas de la Figura 5 a la Figura 19. En estas figuras se compara, para las distintas matrices de prueba, el tiempo de *CPU* en ejecución concurrente, de las versiones síncrona y asíncrona de un mismo esquema caótico, para un amplio intervalo del factor de relajación.

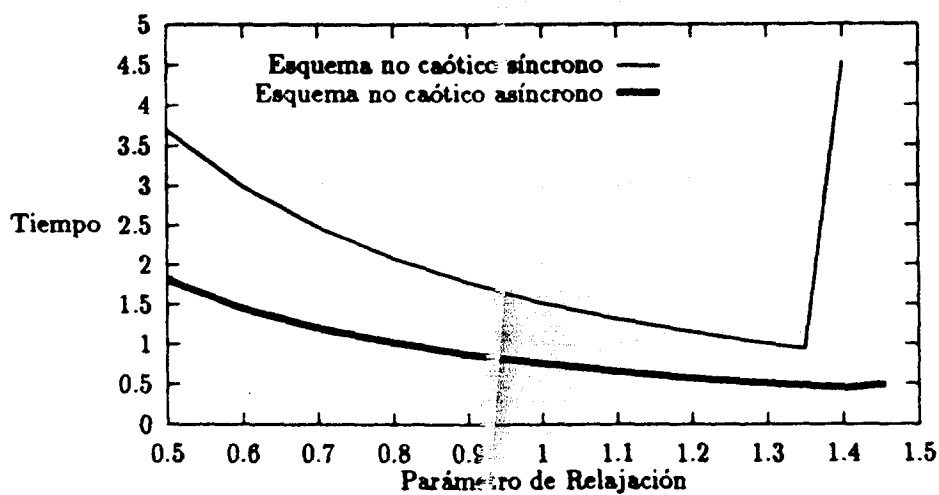


Figura 5: Comparación esquemas paralelos síncrono y asíncrono, $n = 384$.

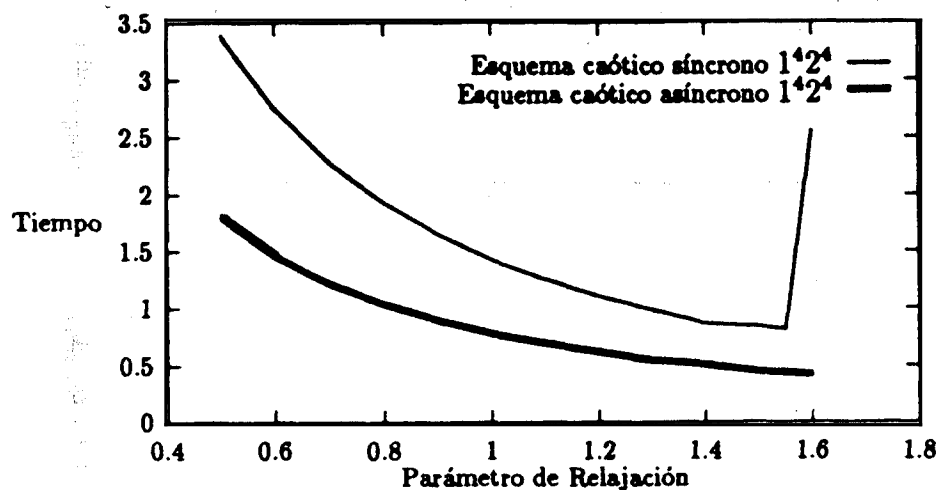


Figura 6: Comparación esquemas paralelos síncrono y asíncrono, $n = 384$.

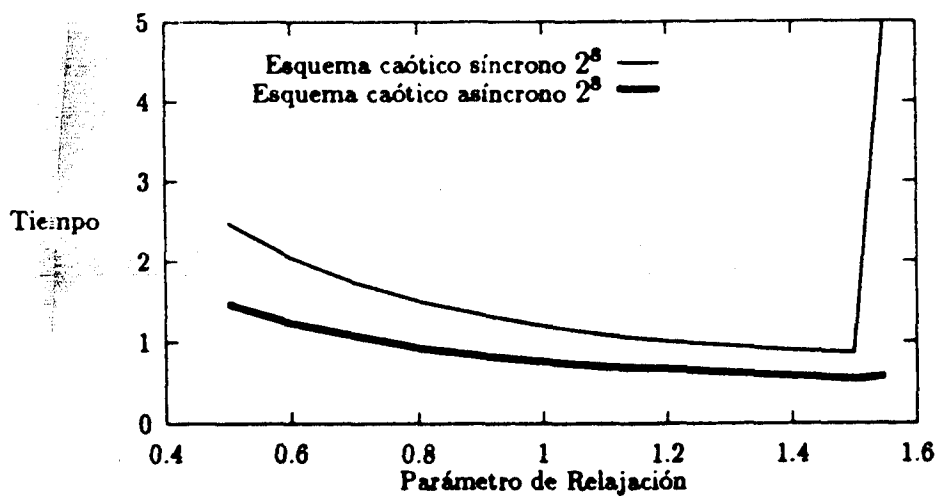


Figura 7: Comparación esquemas paralelos síncrono y asíncrono, $n = 384$.

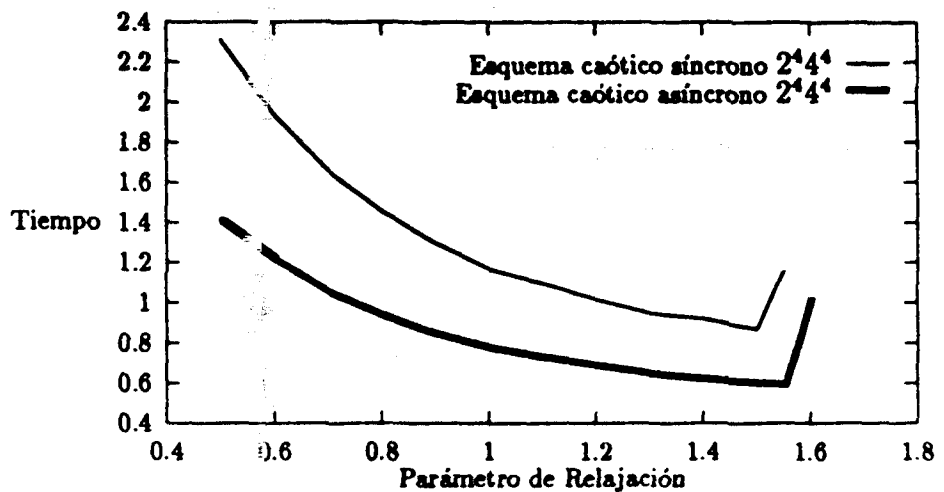


Figura 8: Comparación esquemas paralelos síncrono y asíncrono, $n = 384$.

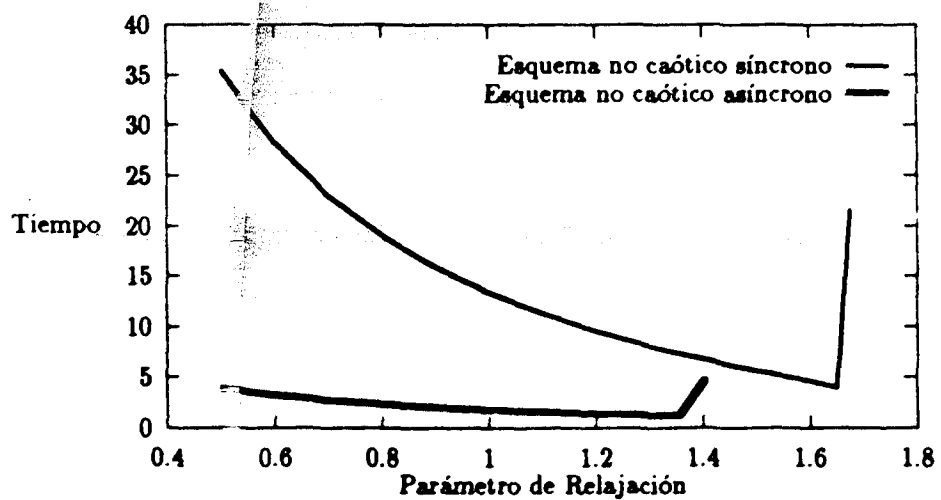


Figura 9: Comparación esquemas paralelos síncrono y asíncrono, $n = 1024$.

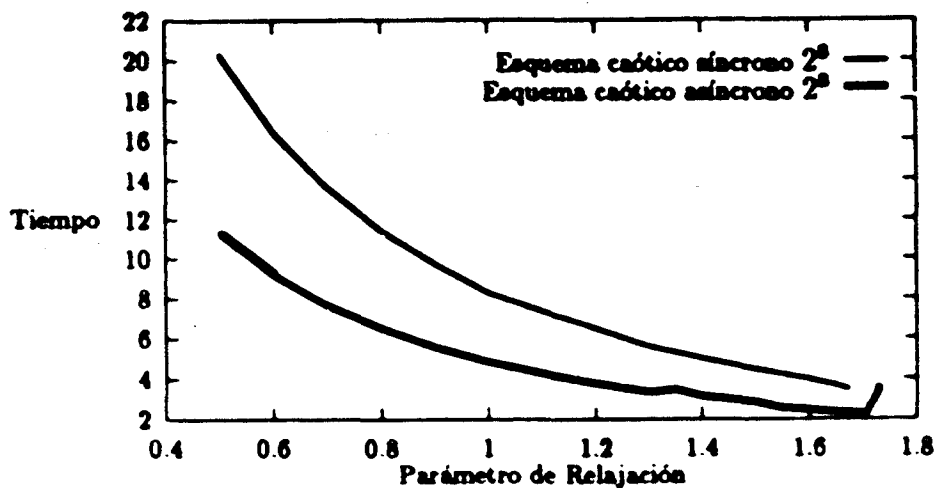


Figura 10: Comparación esquemas paralelos síncrono y asíncrono, $n = 1024$.

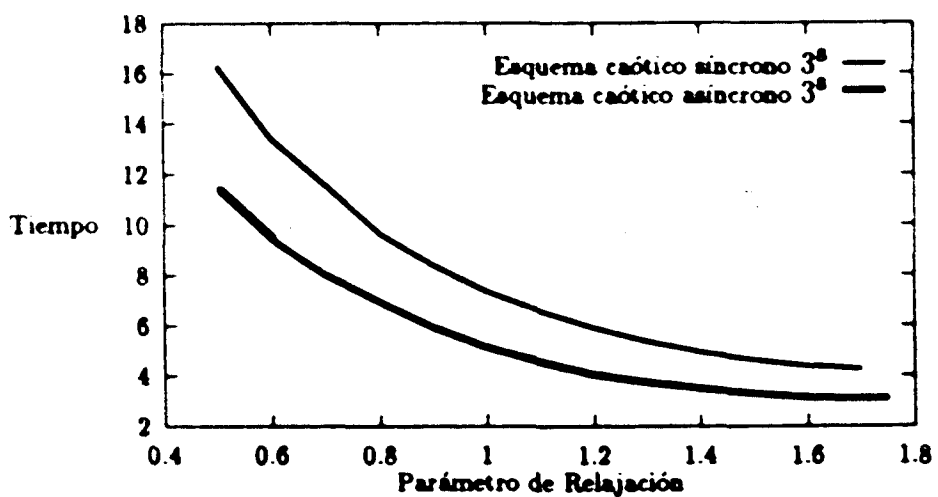


Figura 11: Comparación esquemas paralelos síncrono y asíncrono, $n = 1024$.



3.4 Experimentos numéricos

121

Universitat d'Alacant
Universidad de Alicante

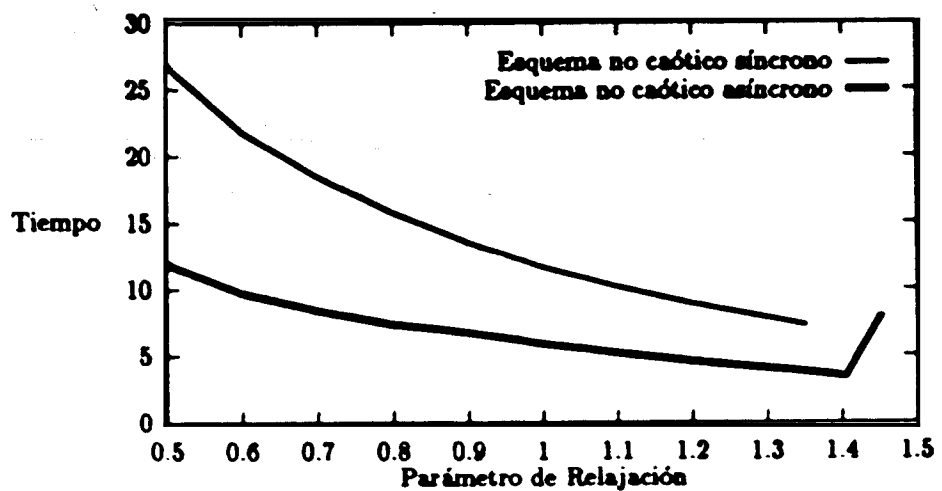


Figura 12: Comparación esquemas paralelos síncrono y asíncrono, $n = 2816$.

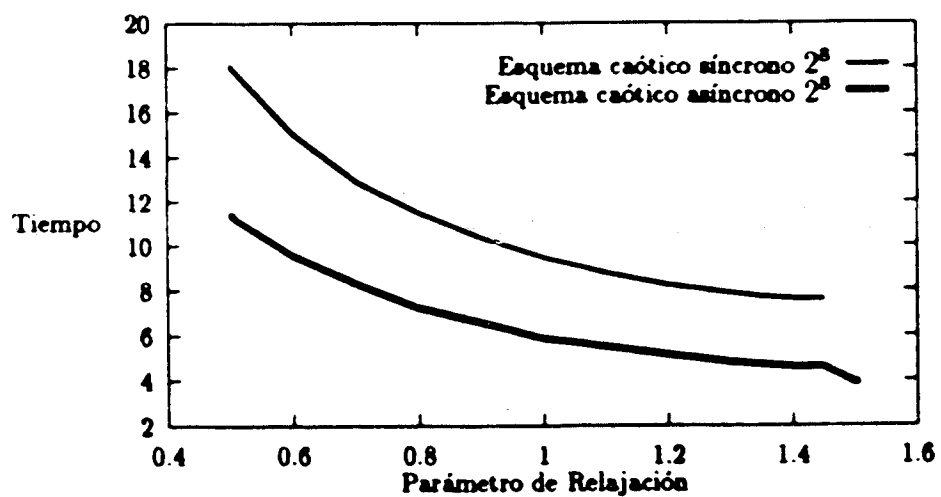


Figura 13: Comparación esquemas paralelos síncrono y asíncrono, $n = 2816$.

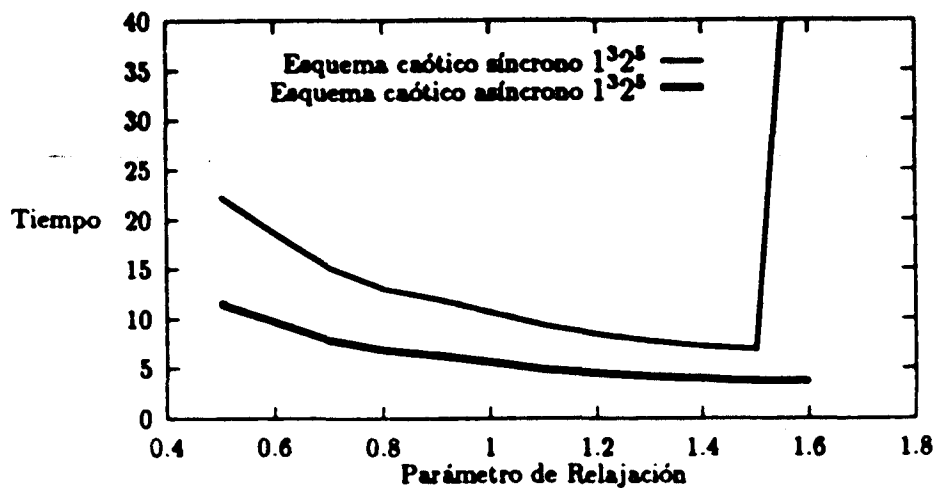


Figura 14: Comparación esquemas paralelos sincrónico y asíncrónico, $n = 2816$.

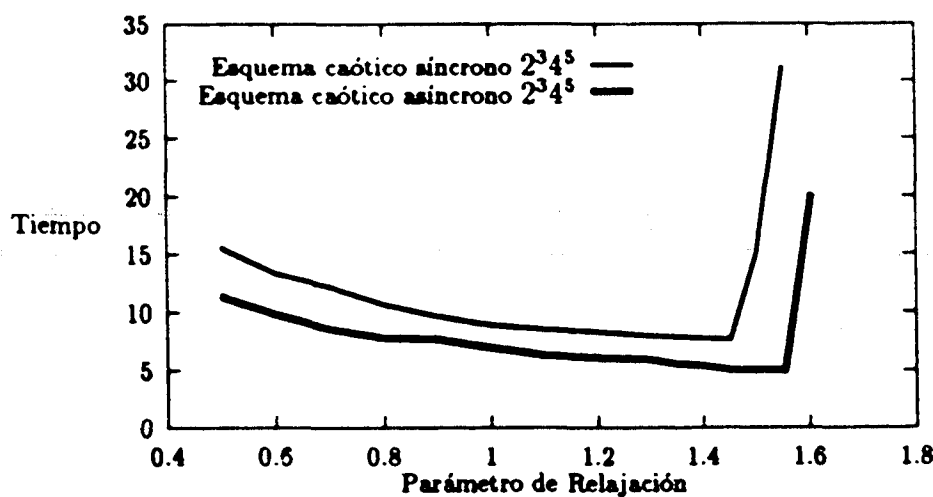
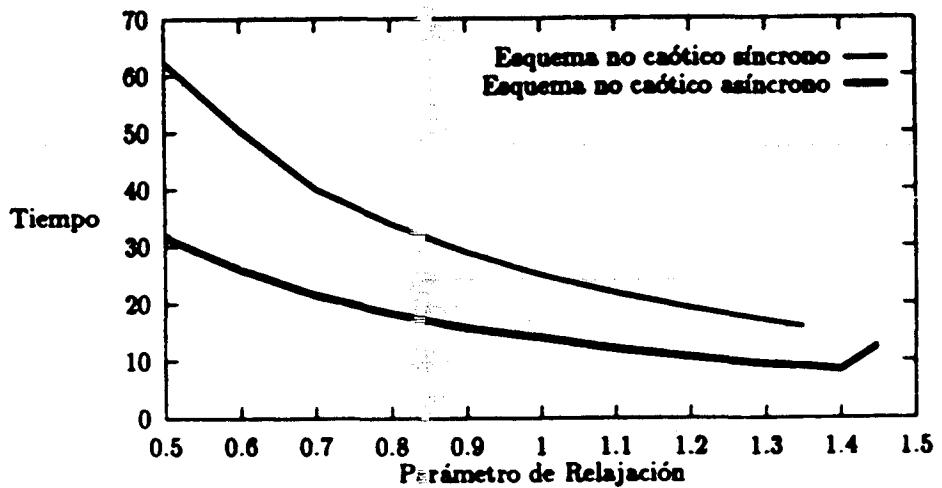
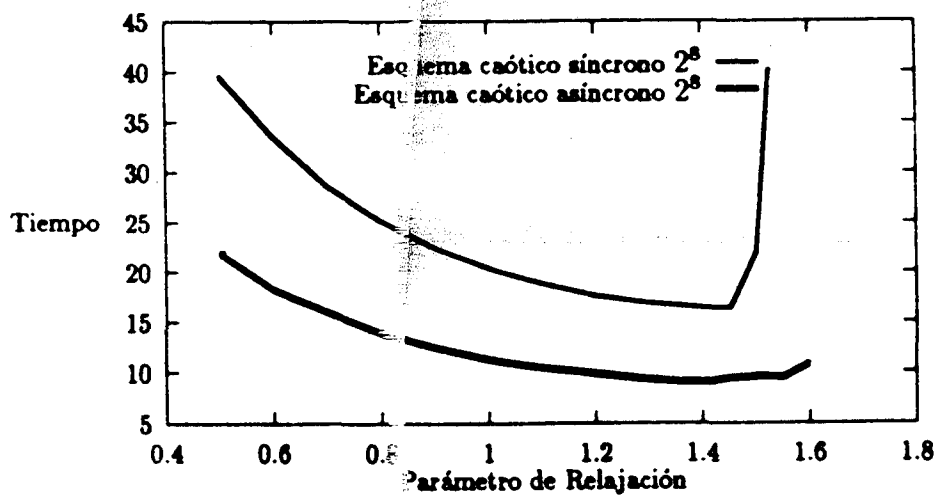


Figura 15: Comparación esquemas paralelos sincrónico y asíncrónico, $n = 816$.

3.4 Experimentos numéricos

123

Figura 16: Comparación esquemas paralelos síncrono y asíncrono, $n = 5632$.Figura 17: Comparación esquemas paralelos síncrono y asíncrono, $n = 5632$.

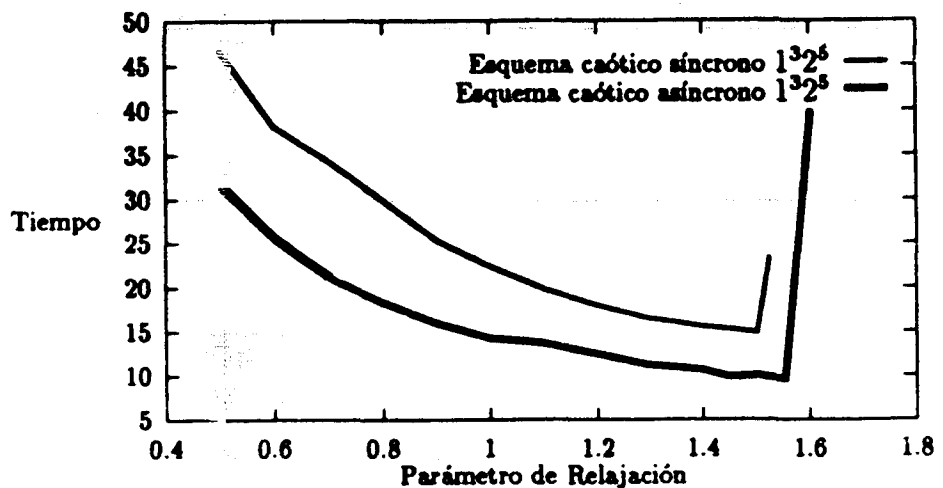


Figura 18: Comparación esquemas paralelos síncrono y asíncrono, $n = 5632$.

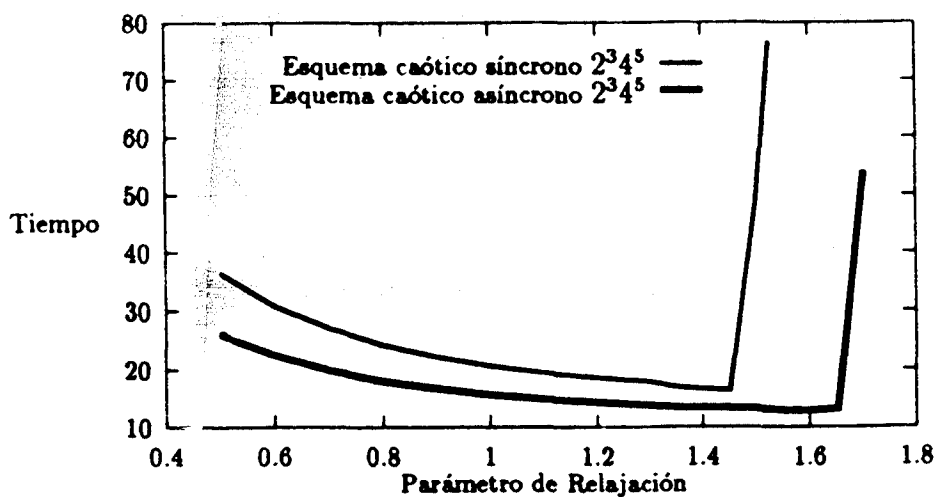


Figura 19: Comparación esquemas paralelos síncrono y asíncrono, $n = 5632$.

La Tabla 5 muestra algunos esquemas caóticos asíncronos correspondientes al factor de relajación óptimo, obtenido de forma experimental. T_1 representa el tiempo de *CPU* cuando se utiliza ejecución concurrente y T_2 cuando se utiliza ejecución vectorial-concurrente.

Se ha calculado el speed-up correspondiente a cada una de las ejecuciones, tomando como algoritmo secuencial de referencia el mejor algoritmo secuencial de entre los obtenidos experimentalmente en la Tabla 2. Esto se ha hecho así, debido a que tal y como se ha construido el esquema caótico asíncrono paralelo, no es factible obtener la correspondiente versión secuencial. Téngase en cuenta que incluso en distintas ejecuciones de un mismo algoritmo asíncrono no se obtienen exactamente los mismos resultados, aunque es de esperar y de hecho así ocurre, que se comporten de manera similar. Esto se observa en las ejecuciones concurrente y vectorial-concurrente, en las que el número de iteraciones por procesador ($iter(j)$, $1 \leq j \leq 8$) y el factor de relajación óptimo no tienen por qué coincidir.

Nótese que obtenemos buenos valores del speed-up, por encima incluso del número de procesadores; este hecho no es contradictorio debido a que aunque el esquema caótico paralelo asíncrono se ha comparado con un "buen" secuencial, no es ni el mejor secuencial ni la versión secuencial del modelo asíncrono.

En la Tabla 5 se observa además que el mejor modelo asíncrono de entre los comparados experimentalmente, cuando se utiliza el factor de relajación óptimo, es el no caótico. Este comportamiento no se puede generalizar a todos los factores de relajación. Las Figuras 20, 21, 22 y 23 presentan para las distintas matrices que son objeto de estudio, el comportamiento de distintos esquemas caóticos asíncronos. Los comentarios realizados respecto al equilibrio de la carga en el caso síncrono, siguen siendo válidos, por lo general, para el caso asíncrono, tal y como se refleja en dichas figuras.

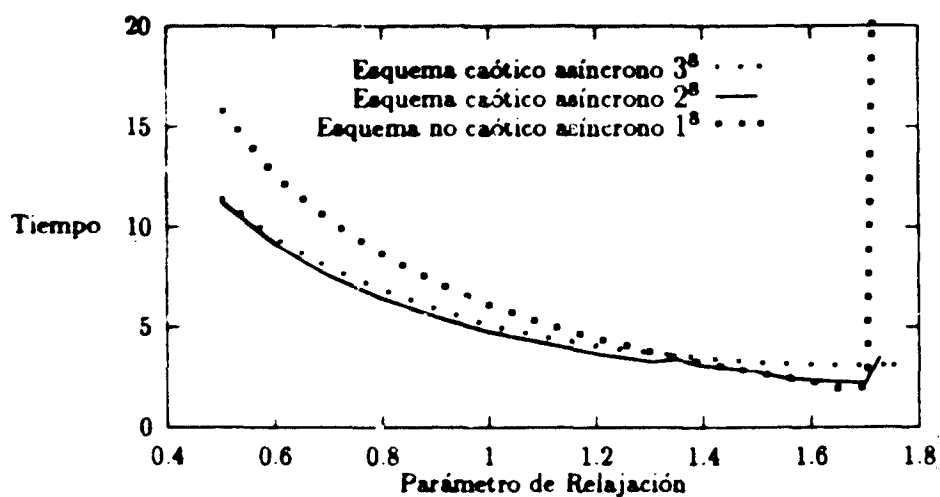


Figura 20: Comparación esquemas paralelos caóticos asíncronos, $n = 1024$.

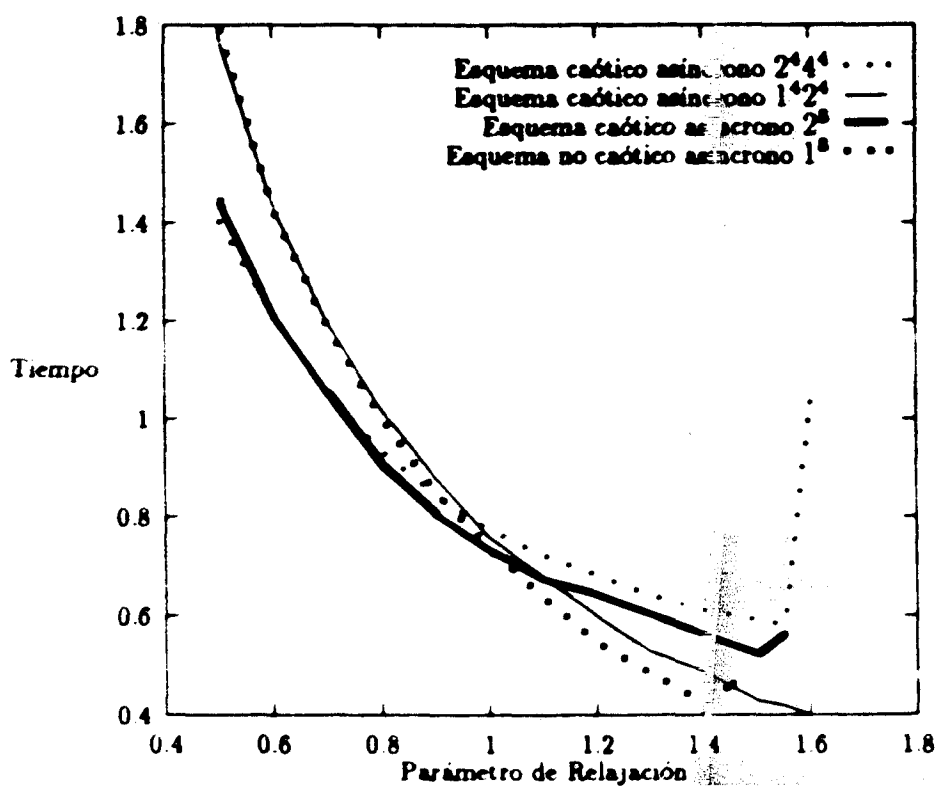


Figura 21: Comparación esquemas paralelos cúbicos asincrónicos, $n = 384$.

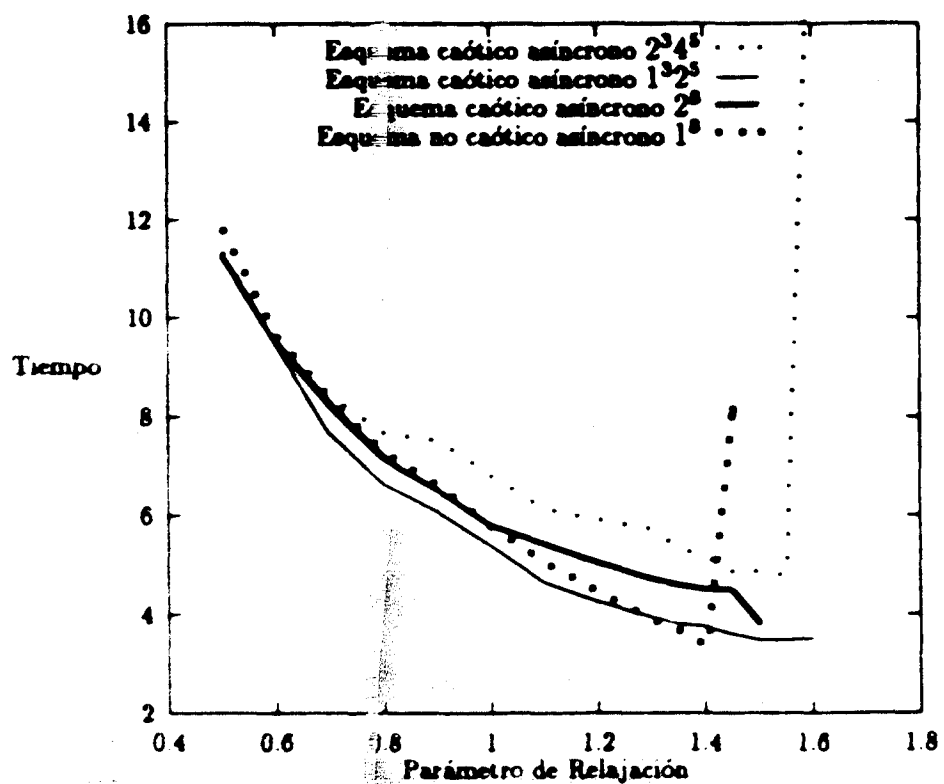


Figura 22: Comparación esquemas paralelos caóticos asíncronos, $n = 2816$.

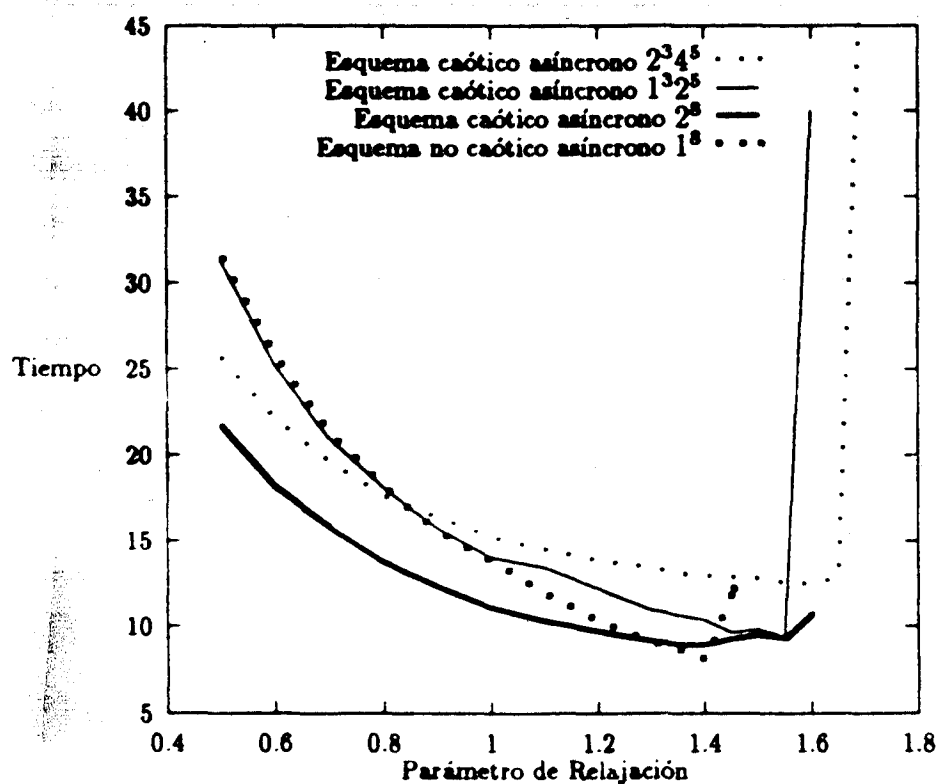


Figura 23: Comparación esquemas paralelos caóticos asincrónicos, $n = 5632$.



3.5 Conclusiones

Los estudios realizados en [12], donde se construye un método iterativo paralelo y caótico, basado en la técnica en dos etapas, motivaron el planteamiento de un esquema caótico más general, en el que las etapas internas no se aplican necesariamente sobre subsistemas construidos a partir de los bloques diagonales de la matriz A del sistema lineal. El planteamiento de este esquema se recoge en la expresión (3.17), en el caso síncrono y en la expresión (3.34), en el caso asíncrono. En las Secciones 3.2 y 3.3 se demuestran los resultados teóricos de convergencia para las versiones síncrona y asíncrona respectivamente.

El Teorema 3 demuestra la convergencia del esquema caótico síncrono (3.17), cuando se realiza un número infinito de iteraciones internas, las particiones tanto interna como externa son convergentes y el factor de relajación pertenece a un intervalo $(0, \omega_0)$ con $\omega_0 > 1$.

La convergencia de este esquema también se demuestra, en el Teorema 4, para matrices monótonas, suponiendo que las particiones externas sean regulares y las internas débilmente regulares; se exige que el factor de relajación esté en el intervalo $(0, 1]$.

También se ha estudiado la convergencia del esquema (3.17) cuando la matriz del sistema es una H -matriz. Este resultado corresponde al Teorema 5. Las hipótesis que se exigen son que las particiones tanto internas como externas sean H -compatibles.

Análogos resultados de convergencia se obtienen, a partir del Teorema 7, para el esquema caótico asíncrono (3.34).

Los experimentos numéricos realizados han mostrado la gran eficiencia de los modelos caóticos frente a los no caóticos. Además, cabe destacar que los modelos caóticos asíncronos se comportan, con independencia del factor de relajación utilizado, mejor que el correspondiente esquema caótico síncrono y por supuesto mejor que los esquemas no caóticos.



Universitat d'Alacant
Universidad de Alicante

Capítulo 4

Método de multipartición relajado

4.1 Introducción

En el Capítulo 3 se ha introducido una técnica de resolución de sistemas de ecuaciones lineales de la forma $A\mathbf{x} = \mathbf{b}$, en la que a partir de la multipartición en dos etapas $\{F_j, G_j, M_j, N_j, E_j\}_{j=1}^r$, de la matriz A y un vector inicial $\mathbf{x}^{(0)}$, se calculan independientemente r soluciones aproximadas de los subsistemas de ecuaciones

$$F_j \mathbf{z} = G_j \mathbf{x}^{(l)} + \mathbf{b}, \quad 1 \leq j \leq r, \quad l = 0, 1, \dots, \quad (4.1)$$

aplicando $q(l, j)$ etapas de un método iterativo relajado basado, también, en multiparticiones. Una vez calculadas en paralelo, las r aproximaciones a la

solución de los subsistemas dados en (4.1), éstas se combinan, en el caso síncrono, para formar el nuevo vector iterado. En el caso asíncrono, no es necesario que todos los procesadores realicen la actualización en el mismo instante.

Por otro lado, en la técnica de multipartición, que introdujeron O'Leary y White [61], en 1985, se obtiene la solución exacta de los r subsistemas que aparecen en la expresión (4.1), de forma independiente. Esta técnica ha sido desarrollada en la Subsección 2.2.1.

Posteriormente, Bru, Elsner y Neumann [10], estudian dos modelos caóticos, basados en la técnica de multipartición, demostrando su convergencia cuando la matriz A es monótona y todas las particiones son regulares. Hemos abordado el estudio de dichos modelos en la Subsección 2.2.2.

Ahora bien, en el contexto de los métodos relajados, O'Leary y White [61] consideran un método de multipartición, en el que el parámetro de relajación ω se usa de la misma forma que en el método extrapolado de Jacobi. Este método, como ya se dijo en la Subsección 2.2.1, (ver expresión (2.11) de Capítulo 2) viene dado por el siguiente esquema iterativo

$$\mathbf{x}^{(l+1)} = \omega \sum_{j=1}^r E_j F_j^{-1}(G_j \mathbf{x}^{(l)} + \mathbf{b}) + (1 - \omega) \mathbf{x}^{(l)}. \quad (4.2)$$

Fromer y Mayer en [29] demuestran la convergencia de este esquema, cuando A es una H -matriz. Hemos recogido este resultado de convergencia en el Teorema 5 del Capítulo 2.

Esto nos ha sugerido plantear un modelo caótico, en el sentido de Bru, Elsner y Neumann [10], basado en el esquema (4.2). Es decir, planteamos un esquema caótico, donde se ha introducido un factor de relajación, como lo hacen O'Leary y White [61].

En la Sección 4.2, plantearémos el nuevo esquema y estudiaremos su convergencia en el contexto de las H -matrices. La Sección 4.3 se dedica al planteamiento y estudio de la versión asíncrona. En las Secciones 4.4 y 4.5 se presentan, respectivamente, los resultados experimentales obtenidos y las conclusiones más relevantes que se extraen del contenido de este capítulo.

Previamente, vamos a introducir una serie de resultados, que nos serán de utilidad en las siguientes secciones.

Un concepto que nos interesa en este capítulo es el de matriz irreducible, que fue introducido en la Sección 1.3 (ver Definición 9). La interpretación geométrica del concepto de irreducibilidad mediante teoría de grafos es sencilla. Tal interpretación, requiere de algunas nociones elementales de dicha teoría. Sea $A = [a_{ij}]$ una matriz compleja de tamaño $n \times n$ y consideremos n puntos distintos en el plano v_1, v_2, \dots, v_n que llamaremos *vértices*. Para cada elemento a_{ij} , no nulo, de la matriz A , se conecta el vértice v_i con el v_j mediante un arco $\overrightarrow{v_i v_j}$, dirigido del vértice v_i al v_j . De esta forma, a cada matriz A de tamaño $n \times n$ se le puede asociar un grafo finito dirigido $G(A)$.

Definición 1. Un grafo dirigido es *fuertemente conexo* si para cualquier par ordenado de vértices v_i y v_j , existe un *camino dirigido*

$$\overrightarrow{v_i v_{k_1}}, \overrightarrow{v_{k_1} v_{k_2}}, \dots, \overrightarrow{v_{k_{r-1}} v_j},$$

conectando el vértice v_i con el v_j .

El siguiente teorema (ver Varga [78]) muestra la equivalencia existente entre el concepto de matriz irreducible de la Definición 9 de la Sección 1.3, y el concepto de grafo dirigido fuertemente conexo de dicha matriz.

Teorema 1. *Una matriz A compleja de tamaño $n \times n$ es irreducible, si y sólo si, su grafo dirigido asociado $G(A)$ es fuertemente conexo.*

Claramente a partir de este resultado, se sigue que si una matriz es irreducible no pueden ser nulos todos los elementos no diagonales de una misma fila o columna, ya que esto significaría que el vértice asociado a dicha fila o columna no se conecta con ningún otro.

Obviamente, si una matriz tiene todos sus elementos no diagonales no nulos, entonces su grafo asociado es fuertemente conexo y por tanto, irreducible.

El siguiente lema, se refiere a H -matrices, y su demostración puede verse en Frommer y Mayer [29].

Lema 1. *Sea $A \in \mathcal{M}_n$ una H -matriz, $D = \text{diag}(A)$ y $A = D - B$. Entonces $|D|$ es no singular y $\rho(|D|^{-1}|B|) < 1$.*

Una vez introducidos los conceptos y propiedades que se utilizan en este capítulo, vamos a construir, en la siguiente sección, un modelo caótico para la resolución de sistemas lineales en paralelo.

4.2 Modelo síncrono

En esta sección vamos a plantear un modelo caótico para la resolución del sistema lineal

$$Ax = b, \quad (4.3)$$

donde A es una matriz cuadrada de tamaño $n \times n$ invertible.

Este modelo está basado en la técnica de multipartición, introduciendo un factor de relajación y unos factores caóticos, tal y como se explica a continuación.

Supongamos que se dispone de un sistema multiprocesador, y sea la multipartición de A , $\{F_j, G_j, E_j\}_{j=1}^r$, es decir, $A = F_j - G_j$, $j = 1, 2, \dots, r$, con F_j invertible y las matrices E_j son matrices no negativas cuya suma es la matriz identidad.

Para cada $j = 1, 2, \dots, r$, definimos los operadores P_j como sigue

$$P_j : \mathbf{R}^n \longrightarrow \mathbf{R}^n$$

$$P_j \mathbf{x} = F_j^{-1} G_j \mathbf{x} + F_j^{-1} \mathbf{b}, \quad j = 1, 2, \dots, r. \quad (4.4)$$

Estos operadores representan, como veremos más adelante, el trabajo que cada procesador debe realizar. Nótese que, aplicar el operador P_j sobre un vector cualquiera \mathbf{x}^* , equivale a resolver el sistema lineal $F_j \mathbf{z} = G_j \mathbf{x}^* + \mathbf{b}$.

A partir de un vector $\mathbf{x}^{(0)}$ arbitrario construiremos una sucesión de vectores $\{\mathbf{x}^{(l)}\}_{l=0}^{\infty}$, de la siguiente forma

$$\mathbf{x}^{(l+1)} = \omega \left(\sum_{j=1}^r E_j \mathbf{x}_j^{(l+1)} \right) + (1 - \omega) \mathbf{x}^{(l)}, \quad \omega \in \mathbf{R}, \quad l = 0, 1, \dots \quad (4.5)$$

El cálculo, en la iteración l , de los vectores $\mathbf{x}_j^{(l+1)}$ se obtiene mediante la siguiente expresión

$$\mathbf{x}_j^{(l+1)} = P_j^{q(j)} \mathbf{x}^{(l)}, \quad j = 1, 2, \dots, r. \quad (4.6)$$

Claramente, los cálculos de los vectores $\mathbf{x}_j^{(l+1)}$, $j = 1, 2, \dots, r$, implicados en la expresión (4.6), son independientes y pueden ser realizados en paralelo. Haremos referencia a estos cálculos denominándolos iteraciones *locales* realizadas por los distintos procesadores.



Los factores $q(l, j)$ indican el número de veces que el procesador j -ésimo realiza los cálculos indicados por los operadores P_j , definidos en (4.4), en la iteración l . Llamaremos a estos factores, *factores caóticos*.

Claramente, sustituyendo la expresión (4.6) en (4.5), el esquema relajado, que se plantea, puede escribirse como

$$\mathbf{x}^{(l+1)} = \omega \sum_{j=1}^r E_j P_j^{q(l,j)} \mathbf{x}^{(l)} + (1 - \omega) \mathbf{x}^{(l)}, \quad l = 0, 1, 2, \dots \quad (4.7)$$

Teniendo en cuenta la construcción del esquema iterativo (4.7), en cada iteración local, no es necesario calcular todas las componentes de los vectores $\mathbf{x}_j^{(l+1)}$, sino sólo aquellas que sean necesarias para los correspondientes cálculos locales y aquellas para las que el correspondiente elemento diagonal de la matriz E_j es no nulo. El papel de estas matrices puede ser visto como el medio que se utiliza para distribuir el trabajo entre los procesadores.

Cuando los factores caóticos $q(l, j)$ son iguales a la unidad, el esquema (4.7) se reduce al esquema (4.2) introducido por O'Leary y White en [61] y cuya convergencia fue estudiada por Frommer y Mayer en [29]. El esquema (4.7) constituye una versión caótica y síncrona del esquema (4.2).

El uso de estos factores caóticos iguales a 1 tiene sentido cuando las matrices E_j reparten por igual el trabajo entre los distintos procesadores. Por el contrario, cuando este trabajo no está distribuido homogéneamente, es apropiado, como se observará en los experimentos numéricos, el uso de factores caóticos mayores que 1, con el fin de equilibrar la carga de trabajo entre los procesadores.

Para analizar la convergencia del esquema iterativo (4.7) escribiremos dicho esquema de una forma alternativa. Para esto, nos basamos en la forma que tienen los operadores P_j definidos en la expresión (4.4).

$$\begin{aligned}
\mathbf{x}^{(l+1)} &= \omega \sum_{j=1}^r E_j P_j^{q(l,j)} \mathbf{x}^{(l)} + (1 - \omega) \mathbf{x}^{(l)} \\
&= \omega \sum_{j=1}^r E_j P_j^{q(l,j)-1} [F_j^{-1} G_j \mathbf{x}^{(l)} + F_j^{-1} \mathbf{b}] + (1 - \omega) \mathbf{x}^{(l)} \\
&= \omega \sum_{j=1}^r E_j P_j^{q(l,j)-2} [F_j^{-1} G_j (F_j^{-1} G_j \mathbf{x}^{(l)} + F_j^{-1} \mathbf{b}) + F_j^{-1} \mathbf{b}] + (1 - \omega) \mathbf{x}^{(l)} \\
&= \omega \sum_{j=1}^r E_j P_j^{q(l,j)-2} [(F_j^{-1} G_j)^2 \mathbf{x}^{(l)} + (F_j^{-1} G_j) F_j^{-1} \mathbf{b} + F_j^{-1} \mathbf{b}] + (1 - \omega) \mathbf{x}^{(l)} \\
&= \omega \sum_{j=1}^r E_j P_j^{q(l,j)-2} [(F_j^{-1} G_j)^2 \mathbf{x}^{(l)} + (F_j^{-1} G_j + I) F_j^{-1} \mathbf{b}] + (1 - \omega) \mathbf{x}^{(l)} \\
&\dots \\
&= \left[\omega \sum_{j=1}^r E_j (F_j^{-1} G_j)^{q(l,j)} + (1 - \omega) I \right] \mathbf{x}^{(l)} \\
&\quad + \omega \sum_{j=1}^r E_j \left(\sum_{i=0}^{q(l,j)-1} (F_j^{-1} G_j)^i \right) F_j^{-1} \mathbf{b}. \tag{4.8}
\end{aligned}$$

A partir de la expresión (4.8), podemos escribir el esquema (4.7) como sigue

$$\mathbf{x}^{(l+1)} = H^{(l)} \mathbf{x}^{(l)} + \mathbf{r}^{(l)}, \quad l = 0, 1, \dots, \tag{4.9}$$

donde

$$H^{(l)} = \omega \sum_{j=1}^r E_j (F_j^{-1} G_j)^{q(l,j)} + (1 - \omega) I, \quad l = 0, 1, \dots, \tag{4.10}$$

y

$$\mathbf{r}^{(l)} = \omega \sum_{j=1}^r E_j \left(\sum_{i=0}^{q(l,j)-1} (F_j^{-1} G_j)^i \right) F_j^{-1} \mathbf{b}, \quad l = 0, 1, \dots. \tag{4.11}$$

A las matrices $H^{(l)}$, $l = 0, 1, \dots$, las denominaremos *matrices de iteración* del esquema (4.9) o equivalentemente del esquema (4.7).

Para probar la convergencia del esquema (4.7) o equivalentemente (4.9) analizaremos la expresión del error

$$e^{(l+1)} = x^{(l+1)} - \xi, \quad l = 0, 1, \dots,$$

siendo ξ la solución exacta del sistema lineal (4.3). El esquema (4.7) será convergente si la sucesión de vectores error $\{e^{(l)}\}_{l=0}^{\infty}$ converge al vector o .

En el siguiente lema expresamos los vectores error $e^{(l+1)}$, $l = 0, 1, \dots$, en función de las matrices de iteración $H^{(l)}$.

Lema 2. Sea ξ la solución del sistema lineal (4.3) y $\{x^{(l)}\}_{l=0}^{\infty}$ la sucesión de vectores generada por el esquema (4.7). Entonces, la sucesión de vectores error $\{e^{(l)}\}_{l=0}^{\infty}$, con $e^{(l)} = x^{(l)} - \xi$, satisface

$$e^{(l+1)} = H^{(l)}e^{(l)}, \quad l = 0, 1, \dots,$$

donde las matrices de iteración $H^{(l)}$ están definidas en (4.10).

Demostración. Teniendo en cuenta las particiones $A = F_j - G_j$, $1 \leq j \leq r$, y que ξ es la solución exacta del sistema lineal (4.3), para cada j , $1 \leq j \leq r$ se tiene

$$(F_j - G_j)\xi = b,$$

$$F_j\xi = G_j\xi + b,$$

$$\xi = F_j^{-1}G_j\xi + F_j^{-1}b,$$

y por la definición de los operadores P_j , $1 \leq j \leq r$ en la expresión (4.4), se tiene que ξ es un punto fijo de los operadores P_j , $1 \leq j \leq r$, es decir,

$$\xi = P_j\xi, \quad 1 \leq j \leq r.$$



4.2 Modelo síncrono

141

Universitat d'Alacant
Universidad de Alicante

Por tanto

$$\begin{aligned}
 \xi &= \omega \xi + (1 - \omega) \xi \\
 &= \omega \sum_{j=1}^r E_j P_j^{q^{(l,j)}} \xi + (1 - \omega) \xi \\
 &= \omega \sum_{j=1}^r E_j P_j^{q^{(l,j)}-1} [F_j^{-1} G_j \xi + F_j^{-1} b] + (1 - \omega) \xi \\
 &= \omega \sum_{j=1}^r E_j P_j^{q^{(l,j)}-2} [F_j^{-1} G_j (F_j^{-1} G_j \xi + F_j^{-1} b) + F_j^{-1} b] + (1 - \omega) \xi \\
 &= \omega \sum_{j=1}^r E_j P_j^{q^{(l,j)}-2} [(F_j^{-1} G_j)^2 \xi + (F_j^{-1} G_j) F_j^{-1} b + F_j^{-1} b] + (1 - \omega) \xi \\
 &= \omega \sum_{j=1}^r E_j P_j^{q^{(l,j)}-2} [(F_j^{-1} G_j)^2 \xi + (F_j^{-1} G_j + I) F_j^{-1} b] + (1 - \omega) \xi \\
 &\dots \\
 &= \left[\omega \sum_{j=1}^r E_j (F_j^{-1} G_j)^{q^{(l,j)}} + (1 - \omega) I \right] \xi \\
 &+ \omega \sum_{j=1}^r E_j \left(\sum_{i=0}^{q^{(l,j)}-1} (F_j^{-1} G_j)^i \right) F_j^{-1} b. \tag{4.12}
 \end{aligned}$$

Considerando la definición de las matrices $H^{(l)}$ en la expresión (4.10) y de los vectores $r^{(l)}$ en (4.11), la igualdad dada en (4.12) puede escribirse como

$$\xi = H^{(l)} \xi + r^{(l)}, \quad l = 0, 1, \dots \tag{4.13}$$

De las expresiones (4.9) y (4.13) se concluye

$$\begin{aligned}
 e^{(l+1)} &= x^{(l+1)} - \xi \\
 &= H^{(l)} x^{(l)} + r^{(l)} - (H^{(l)} \xi + r^{(l)})
 \end{aligned}$$



$$\begin{aligned}
 &= H^{(l)}(\mathbf{x}^{(l)} - \xi) \\
 &= H^{(l)}\mathbf{e}^{(l)},
 \end{aligned}$$

con $H^{(l)}$ definida en (4.10). ■

Observación 1. A partir del Lema 2, podemos expresar el vector error $\mathbf{e}^{(l)}$, para cada iteración l , $l = 0, 1, \dots$, en función del vector error en la iteración inicial, denotado por $\mathbf{e}^{(0)}$.

$$\mathbf{e}^{(l+1)} = H^{(l)}\mathbf{e}^{(l)} = \dots = \prod_{i=0}^l H^{(i)}\mathbf{e}^{(0)}, \quad l = 0, 1, \dots$$

Teniendo en cuenta esta expresión, la sucesión de vectores error $\{\mathbf{e}^{(l)}\}_{l=0}^{\infty}$ converge al vector $\mathbf{0}$, si y sólo si

$$\lim_{l \rightarrow \infty} \prod_{i=0}^l H^{(i)} = \mathbf{0}. \quad (4.14)$$

Siguiendo estas consideraciones y a partir del Lema 3 del Capítulo 3, la expresión (4.14) es cierta si existe una constante real α , tal que

$$\|H^{(l)}\| \leq \alpha < 1, \quad l = 0, 1, \dots, \quad (4.15)$$

para alguna norma matricial compatible $\|\cdot\|$.

En consecuencia el esquema (4.7) es convergente, si se satisface la expresión (4.15).

Una vez construido el esquema caótico síncrono (4.7), vamos a dar el resultado principal de convergencia que se ha obtenido para este modelo.

Si la matriz A es una H -matriz y expresamos $A = D - B$, con $D = \text{diag}(A)$, por el Lema 1 sabemos que $|D|$ es noringular y $\rho(|D|^{-1}|B|) < 1$. Con esta

notación, se prueba la convergencia del esquema iterativo relajado (4.7), cuando la matriz del sistema A es una H -matriz y el factor de relajación pertenece al intervalo $(0, \frac{2}{1+\rho(|D|^{-1}|B|)})$.

Teorema 2. *Sea A una H -matriz. Supongamos que las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$, son particiones H -compatibles de A , con $\text{diag}(|F_j|) \leq |D|$. Entonces, si $q(l, j) \geq 1$, $l = 0, 1, \dots$, $1 \leq j \leq r$, el esquema (4.7) converge a la única solución del sistema $Ax = b$, para $\omega \in (0, \frac{2}{1+\rho})$, con $\rho = \rho(|D|^{-1}|B|)$.*

Demostración. Por la Observación 1, para probar la convergencia del esquema (4.7) es suficiente encontrar una norma matricial compatible $\|\cdot\|$ y una constante real α , de forma que

$$\|H^{(l)}\| \leq \alpha < 1, \quad l = 0, 1, \dots$$

Como las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$, son H -compatibles, se sigue por el apartado (b) del Lema 6 del Capítulo 3, que dichas particiones son H -particiones. Entonces, por el apartado (a) del mismo lema, se obtiene que las matrices F_j , $j = 1, 2, \dots, r$, son H -matrices.

Claramente, por la forma de las matrices de iteración $H^{(l)}$ en (4.10) obtenemos

$$|H^{(l)}| \leq \omega \sum_{j=1}^r E_j |F_j^{-1} G_j|^{q(l,j)} + |1 - \omega|I, \quad l = 0, 1, \dots \quad (4.16)$$

Ahora bien, teniendo en cuenta que las matrices F_j , $1 \leq j \leq r$ son H -matrices, como acabamos de demostrar, y utilizando el apartado (a) del Lema 5 del Capítulo 3, podemos acotar $|F_j^{-1} G_j|$ de la siguiente forma

$$|F_j^{-1} G_j| \leq |F_j^{-1}| |G_j| \leq (F_j)^{-1} |G_j|, \quad j = 1, 2, \dots, r. \quad (4.17)$$

Por la definición de partición H -compatible (ver Definición 14 de la Sección 1.4), sabemos que las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$, satisfacen $\langle A \rangle = \langle F_j \rangle - |G_j|$, $j = 1, 2, \dots, r$, y por tanto, sustituyendo $|G_j| = \langle F_j \rangle - \langle A \rangle$ en la expresión (4.17), se obtiene

$$\begin{aligned} |F_j^{-1}G_j| &\leq \langle F_j \rangle^{-1}|G_j| \\ &= \langle F_j \rangle^{-1}(\langle F_j \rangle - \langle A \rangle) \\ &= I - \langle F_j \rangle^{-1}\langle A \rangle, \quad 1 \leq j \leq r. \end{aligned} \quad (4.18)$$

Por otro lado, puesto que $A = D - B$, con $D = \text{diag}(A)$, y teniendo en cuenta que construir la matriz comparación de A (ver Definición 11 de la Sección 1.4) consiste en considerar los elementos de la diagonal en valor absoluto y los no diagonales, en valor absoluto con signo negativo, la matriz comparación de A se puede escribir como

$$\langle A \rangle = |D| - |B| = |D|(I - |D|^{-1}|B|). \quad (4.19)$$

Denotando $J = |D|^{-1}|B|$ y sustituyendo en la expresión (4.19) se tiene

$$\langle A \rangle = |D|(I - J). \quad (4.20)$$

De las expresiones (4.18) y (4.20) se sigue

$$|F_j^{-1}G_j| \leq I - \langle F_j \rangle^{-1}|D|(I - J), \quad 1 \leq j \leq r. \quad (4.21)$$

Consideremos el vector n -dimensional $\mathbf{e} = (1, 1, \dots, 1)^t$. Como la matriz $J = |D|^{-1}|B|$ es no negativa, para todo $\epsilon \in \mathbb{R}$ mayor que cero, la matriz

$J + \epsilon \mathbf{e} \mathbf{e}^t$ tiene todos los elementos positivos, por tanto, el grafo dirigido asociado a la matriz $J + \epsilon \mathbf{e} \mathbf{e}^t$ es fuertemente conexo (ver Sección 4.1). Por el Teorema 1 se sigue que esta matriz es irreducible. Aplicando el resultado del apartado (iv) del Teorema 2 del Capítulo 2, sobre la matriz $J + \epsilon \mathbf{e} \mathbf{e}^t$, obtenemos que para todo $\epsilon > 0$, existe un vector positivo \mathbf{x}_ϵ , satisfaciendo

$$(J + \epsilon \mathbf{e} \mathbf{e}^t) \mathbf{x}_\epsilon = \rho_\epsilon \mathbf{x}_\epsilon, \quad (4.22)$$

donde $\rho_\epsilon = \rho(J + \epsilon \mathbf{e} \mathbf{e}^t)$.

Puesto que F_j , $j = 1, 2, \dots, r$, son H -matrices, aplicando el apartado (a) del Lema 5 del Capítulo 3, se tiene

$$|F_j^{-1}| \leq \langle F_j \rangle^{-1}, \quad j = 1, 2, \dots, r,$$

y por tanto

$$\langle F_j \rangle^{-1} \geq 0, \quad j = 1, 2, \dots, r.$$

De aquí se sigue que la matriz $\langle F_j \rangle^{-1} |D| \epsilon \mathbf{e} \mathbf{e}^t$ es no negativa, obteniendo las siguientes desigualdades

$$\begin{aligned} I - \langle F_j \rangle^{-1} |D| (I - J) &\leq I - \langle F_j \rangle^{-1} |D| (I - J) + \langle F_j \rangle^{-1} |D| \epsilon \mathbf{e} \mathbf{e}^t \\ &= I - \langle F_j \rangle^{-1} |D| (I - (J + \epsilon \mathbf{e} \mathbf{e}^t)). \end{aligned} \quad (4.23)$$

De las expresiones (4.21) y (4.23) obtenemos

$$|F_j^{-1} G_j| \leq I - \langle F_j \rangle^{-1} |D| (I - (J + \epsilon \mathbf{e} \mathbf{e}^t)).$$

Multiplicando ambos miembros de esta desigualdad por el vector positivo \mathbf{x}_ϵ y utilizando la expresión (4.22) se tiene

$$\begin{aligned} |F_j^{-1} G_j| \mathbf{x}_\epsilon &\leq [I - \langle F_j \rangle^{-1} |D| (I - (J + \epsilon \mathbf{e} \mathbf{e}^t))] \mathbf{x}_\epsilon \\ &= \mathbf{x}_\epsilon - \langle F_j \rangle^{-1} |D| (1 - \rho_\epsilon) \mathbf{x}_\epsilon. \end{aligned} \quad (4.24)$$

Por hipótesis, $\text{diag}(|F_j|) \leq |D|$, $j = 1, 2, \dots, r$, y como todos los elementos no diagonales de $\langle F_j \rangle$ son no positivos, se tiene que

$$\langle F_j \rangle \leq |D|. \quad (4.25)$$

El Lema 1 nos asegura que el radio espectral de la matriz J es menor que 1. Si tenemos en cuenta la continuidad del radio espectral podemos afirmar que, para algún $\epsilon > 0$

$$\rho_\epsilon = \rho(J + \epsilon e e^t) < 1. \quad (4.26)$$

Utilizando la desigualdad (4.25) en la expresión (4.24) y teniendo en cuenta que según la expresión (4.26), $1 - \rho_\epsilon > 0$, obtenemos

$$\begin{aligned} |F_j^{-1}G_j|\mathbf{x}_\epsilon &\leq \mathbf{x}_\epsilon - \langle F_j \rangle^{-1}|D|(1 - \rho_\epsilon)\mathbf{x}_\epsilon \\ &\leq \mathbf{x}_\epsilon - \langle F_j \rangle^{-1}\langle F_j \rangle(1 - \rho_\epsilon)\mathbf{x}_\epsilon \\ &= \mathbf{x}_\epsilon - (1 - \rho_\epsilon)\mathbf{x}_\epsilon \\ &= \rho_\epsilon\mathbf{x}_\epsilon. \end{aligned} \quad (4.27)$$

Multiplicando ambos miembros de la desigualdad (4.16) por el vector positivo \mathbf{x}_ϵ y teniendo en cuenta la acotación dada por la expresión (4.27), se obtienen las siguientes desigualdades

$$\begin{aligned} |H^{(l)}|\mathbf{x}_\epsilon &\leq \omega \sum_{j=1}^r E_j |F_j^{-1}G_j|^{\varrho^{(l,j)}} \mathbf{x}_\epsilon + |1 - \omega|\mathbf{x}_\epsilon \\ &\leq \omega \sum_{j=1}^r E_j |F_j^{-1}G_j|^{\varrho^{(l,j)}-1} \rho_\epsilon \mathbf{x}_\epsilon + |1 - \omega|\mathbf{x}_\epsilon \end{aligned}$$

$$\dots \leq \omega \sum_{j=1}^r E_j \rho_c^{(l,j)} \mathbf{x}_c + |1 - \omega| \mathbf{x}_c, \quad l = 0, 1, \dots \quad (4.28)$$

Puesto que ρ_c es menor que 1, como ya se ha comprobado en (4.26), de la expresión (4.28) se sigue

$$\begin{aligned} |H^{(l)}| \mathbf{x}_c &\leq \omega \sum_{j=1}^r E_j \rho_c^{(l,j)} \mathbf{x}_c + |1 - \omega| \mathbf{x}_c \\ &\leq \omega \sum_{j=1}^r E_j \rho_c \mathbf{x}_c + |1 - \omega| \mathbf{x}_c \\ &\leq (\omega \rho_c + |1 - \omega|) \mathbf{x}_c, \quad l = 0, 1, \dots \end{aligned} \quad (4.29)$$

Distinguimos dos casos atendiendo al valor de ω .

(a) Si $0 < \omega < 1$,

$$\omega \rho_c + |1 - \omega| = \omega \rho_c + 1 - \omega < 1.$$

(b) Si $1 \leq \omega < \frac{2}{1+\rho_c}$

$$\begin{aligned} \omega \rho_c + |1 - \omega| &= \omega \rho_c + \omega - 1 \\ &= \omega(1 + \rho_c) - 1 \\ &< \frac{2}{1 + \rho_c}(1 + \rho_c) - 1 \\ &= 1. \end{aligned}$$

De (a) y (b), junto a la propiedad de continuidad del radio espectral, se tiene que para $\omega \in (0, \frac{2}{1+\rho_c})$

$$\omega \rho_c + |1 - \omega| < 1$$

y por tanto por (4.29)

$$|H^{(l)}| \mathbf{x}_l < \alpha_l \mathbf{x}_l,$$

con $\alpha_l = \omega \rho_l + |1 - \omega| < 1$ y $\mathbf{x}_l > \mathbf{0}$.

De aquí se sigue que la norma matricial compatible y monótona $\|\cdot\|_{\mathbf{x}_l}$, introducida en la Sección 1.1 cumple

$$\|H^{(l)}\|_{\mathbf{x}_l} \leq \alpha_l < 1, \quad l = 0, 1, \dots$$

Queda así, demostrada la convergencia del esquema (4.7). ■

Las condiciones del Teorema 2 relativas a las particiones de A , es decir, $\text{diag}(|F_j|) \leq |D|$ y $\langle A \rangle = \langle F_j \rangle - |G_j|$ se cumplen esencialmente cuando los elementos no diagonales de F_j están entre los correspondientes elementos de A y 0. Esto se pone de manifiesto en el siguiente lema.

Lema 3. Sea A una matriz de tamaño $n \times n$, con diagonal D . Existen matrices $F = [f_{ij}]_{1 \leq i, j \leq n}$ y $G = [g_{ij}]_{1 \leq i, j \leq n}$ de tamaño $n \times n$ satisfaciendo $\text{diag}(|F|) \leq |D|$ y $\langle A \rangle = \langle F \rangle - |G|$, si y sólo si,

$$|f_{ij}| \in [0, |a_{ij}|], \quad i \neq j, 1 \leq i, j \leq n,$$

$$|f_{ii}| = |a_{ii}|, \quad 1 \leq i \leq n.$$

Demostración. (\rightarrow) Si $i \neq j$, $1 \leq i, j \leq n$, por la igualdad $\langle A \rangle = \langle F \rangle - |G|$, se tiene

$$-|a_{ij}| = -|f_{ij}| - |g_{ij}|.$$

Despejando $|f_{ij}|$

$$|f_{ij}| = |a_{ij}| - |g_{ij}| \in [0, |a_{ij}|].$$

Si $i = j$, $1 \leq i \leq n$, como $\langle A \rangle = \langle F \rangle - |G|$ entonces

$$|a_{ii}| = |f_{ii}| - |g_{ii}|,$$

$$|f_{ii}| = |a_{ii}| + |g_{ii}|.$$

Pero al ser, por hipótesis, $|f_{ii}| \leq |a_{ii}|$, necesariamente $|g_{ii}| = 0$. Luego

$$|a_{ii}| = |f_{ii}|, \quad 1 \leq i \leq n.$$

(\leftarrow) En primer lugar construimos los elementos diagonales de F y G como

$$|f_{ii}| = |a_{ii}|, \quad g_{ii} = 0, \quad 1 \leq i \leq n.$$

Para $i \neq j$, $1 \leq i, j \leq n$ sabemos que los elementos $|f_{ij}|$ están en el intervalo $[0, |a_{ij}|]$, luego existen elementos x_{ij} con $0 \leq x_{ij} \leq |a_{ij}|$ tales que

$$|f_{ij}| = |a_{ij}| - x_{ij}.$$

Definimos ahora los elementos no diagonales de G de forma que verifiquen

$$|g_{ij}| = x_{ij}, \quad 1 \leq i, j \leq n, \quad i \neq j.$$

Obviamente se satisface que $\text{diag}(|F|) \leq |D|$. Por otro lado para $i \neq j$

$$-|f_{ij}| - |g_{ij}| = -|f_{ij}| - x_{ij} = -|f_{ij}| - (|a_{ij}| - |f_{ij}|) = -|a_{ij}|.$$

Por tanto

$$\langle A \rangle = \langle F \rangle - |G|.$$

Queda pues, probado el resultado de este lema. ■

Algunos ejemplos de multipartición de A con posible solapamiento, que satisface el lema anterior son los siguientes.

Ejemplo 1. Sean para $k = 1, 2, \dots, r$, los subconjuntos $S_k \subset \{1, 2, \dots, n\}$, de forma que $\cup_{k=1}^r S_k = \{1, 2, \dots, n\}$, no necesariamente disjuntos dos a dos. Construimos una multipartición de A , $\{F_k, G_k, E_k\}_{k=1}^r$, de forma que las matrices $F_k = [f_{ij}^k]_{1 \leq i, j \leq n}$ están definidas como

$$f_{ij}^k = \begin{cases} a_{ij} & \text{si } i = j \text{ o } i, j \in S_k \\ 0 & \text{en otro caso.} \end{cases}$$

Se supone además que las matrices F_k son invertibles. El elemento i -ésimo de las matrices E_k , $k = 1, 2, \dots, r$ se toma cero si $i \notin S_k$ y positivo si $i \in S_k$. Además $\sum_{k=1}^r E_k = I$, y las matrices G_k se eligen de forma que $A = F_k - G_k$.

Ejemplo 2. Sean los subconjuntos $S_k \subset \{1, 2, \dots, n\}$, $k = 1, 2, \dots, r$, elegidos de forma análoga al ejemplo anterior, es decir, $\{S_k\}_{k=1}^r$ es un recubrimiento del conjunto $\{1, 2, \dots, n\}$. Construimos una multipartición de A , $\{F_k, G_k, E_k\}_{k=1}^r$, de forma que las matrices $F_k = [f_{ij}^k]_{1 \leq i, j \leq n}$ están definidas como

$$f_{ij}^k = \begin{cases} a_{ij}, & \text{si } i < j \text{ y } i, j \in S_k \\ a_{ij}, & \text{si } i = j \\ 0 & \text{en otro caso.} \end{cases}$$

Supondremos que las matrices F_k son invertibles. El elemento i -ésimo de las matrices E_k , $k = 1, 2, \dots, r$ se toma cero si $i \notin S_k$ y positivo si $i \in S_k$. Además $\sum_{k=1}^r E_k = I$, y las matrices G_k se eligen de forma que $A = F_k - G_k$.

Cuando los conjuntos S_k , $k = 1, 2, \dots, n$, sean disjuntos dos a dos, es decir, cuando $\{S_k\}_{k=1}^r$ sea una partición del conjunto $\{1, 2, \dots, n\}$, las matrices



E_k , $k = 1, 2, \dots, r$, se eligen de forma que

$$(E_k)_{ii} = 1 \quad \text{si } i \in S_k$$

$$(E_k)_{ii} = 0 \quad \text{si } i \notin S_k.$$

Observación 2. (i) La convergencia del esquema (4.7) puede demostrarse suponiendo r parámetros de relajación distintos, asociados a cada una de las r particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$. Es decir, el esquema caótico síncrono con r parámetros de relajación $\omega_1, \omega_2, \dots, \omega_r$

$$\mathbf{x}^{(l+1)} = \sum_{j=1}^r E_j \left[\omega_j P_j^{q(l,j)} \mathbf{x}^{(l)} + (1 - \omega_j) \mathbf{x}^{(l)} \right], \quad l = 0, 1, 2, \dots,$$

converge para las mismas hipótesis del Teorema 2, si los r parámetros de relajación están en el intervalo $(0, \frac{2}{1+\rho})$ con $\rho = \rho(|D|^{-1}|B|)$.

(ii) Cuando $\omega = 1$, el esquema (4.7) se reduce al esquema iterativo

$$\mathbf{x}^{(l+1)} = \sum_{j=1}^r E_j P_j^{q(l,j)} \mathbf{x}^{(l)}, \quad l = 0, 1, 2, \dots, \quad (4.30)$$

que corresponde al primer modelo caótico dado por Bru, Elsner y Neumann en [10]. Recordamos que en [10] se prueba la convergencia del esquema (4.30) cuando A es una matriz monótona y las particiones de A son débilmente regulares. Hacemos notar que el Teorema 2 da un nuevo criterio de convergencia para el esquema (4.30) que no había sido considerado hasta ahora en la literatura sobre el tema, exigiendo que A sea una H -matriz, matrices que no son necesariamente monótonas y que como se refleja en el Teorema 9 de la Sección 1.4 cubre varios tipos interesantes de matrices.



4.3 Modelo asíncrono

En esta sección planteamos la versión asíncrona del esquema iterativo caótico construido en la expresión (4.7) de la sección anterior. Como ya hemos comentado en el Capítulo 2 la idea de los modelos asíncronos surge con la intención de obtener una mayor utilización de los procesadores en un sistema multiprocesador. En los modelos asíncronos aplicados sobre procedimientos iterativos de resolución de sistemas de ecuaciones lineales, se supone que cada procesador puede actualizar o recuperar algunas componentes del vector aproximación a la solución en cualquier instante. Esto, teóricamente, significa que ningún procesador se mantiene ocioso.

A continuación describimos el esquema caótico asíncrono basado en el esquema caótico síncrono (4.7).

Sean E_i , $i = 1, 2, \dots, r$, matrices diagonales no negativas cuya suma es la matriz identidad. Si suponemos que $\mathcal{F}_0 = \{i_j\}_{j=1}^{\infty}$, $1 \leq i_j \leq r$, es una sucesión regulada (ver Definición 5 del Capítulo 2), a partir de un vector inicial, construimos el siguiente esquema iterativo asíncrono

$$\mathbf{x}^{(j+r)} = (I - E_{i_j})\mathbf{x}^{(j+r-1)} + E_{i_j} \left(\omega \sum_{i=1}^r E_i P_i^{q(i,j)} + (1 - \omega)I \right) \mathbf{x}^{(j)}, \quad (4.31)$$

$$j = 0, 1, 2, \dots,$$

siendo P_i , $i = 1, 2, \dots, r$, los operadores definidos en la expresión (4.4) de la sección anterior

$$P_i \mathbf{x} = F_i^{-1} G_i \mathbf{x} + F_i^{-1} \mathbf{b}, \quad i = 1, 2, \dots, r,$$



4.3 Modelo asíncrono

153

y r_j el menor entero positivo tal que

$$i_j = i_{j+r_j}.$$

Por la elección de r_j , como $\{i_j\}_{j=1}^{\infty}$ es regulada existe un entero positivo K que satisface

$$0 \leq r_j - 1 < K, \quad j = 0, 1, 2, \dots \quad (4.32)$$

Nótese que el valor de $r_j - 1$ es igual al número de veces que la aproximación que reside en un procesador, que denominamos *central*, es actualizada por otros procesadores distintos del i_j -ésimo, durante el intervalo de tiempo en el que dicho procesador ejecuta sus cálculos.

El siguiente teorema da condiciones suficientes de convergencia para el esquema asíncrono definido en (4.31). Las hipótesis que se exigen son análogas a las exigidas para el caso síncrono en el Teorema 2. Al igual que en dicho teorema denotaremos por D a la diagonal de A , con lo que $\mathcal{A} = D - B$; además $\rho = \rho(|D|^{-1}|B|)$.

Teorema 3. *Sea A una H -matriz. Supongamos que las siguientes particiones $A = F_i - G_i$, $i = 1, 2, \dots, r$, son particiones H -compatibles de A , con $\text{diag}(|F_i|) \leq |D|$ y $\{i_j\}_{j=0}^{\infty}$, $1 \leq i_j \leq r$, una sucesión regulada de enteros. Entonces, si $q(j, i) \geq 1$, $j = 0, 1, \dots$, $1 \leq i \leq r$, el esquema (4.31) converge a la única solución del sistema (4.9), para $\omega \in (0, \frac{2}{1+\rho})$.*

Demostración. Análogamente a como se hizo en el Capítulo 3, para analizar la convergencia del esquema asíncrono (4.31) se sumerge este procedimiento iterativo en un procedimiento iterativo en \mathbb{R}^{nK} , siendo K el entero positivo que regula a la sucesión $\{i_j\}_{j=1}^{\infty}$ en la Definición 5 del Capítulo 2.

Utilizamos, para esto, la siguiente notación.

$$\epsilon^{(j)} = z^{(j)} - \xi, \quad \bar{z}_j = \begin{bmatrix} \epsilon^{(j)} \\ \epsilon^{(j-1)} \\ \vdots \\ \epsilon^{(j-K+1)} \end{bmatrix} \in \mathbb{R}^{nK},$$

y

$$\bar{z} = \begin{bmatrix} z_c \\ z_c \\ \vdots \\ z_c \end{bmatrix} \in \mathbb{R}^{nK}, \quad (4.33)$$

donde $z_c > 0$ es el vector de Perron de la matriz $J + \epsilon e e^t$, con $J = |D|^{-1}|B|$, $\epsilon > 0$ y $e = (1, 1, \dots, 1)^t$. En el Teorema 2 se probó que

$$|H^{(l)}| z_c \leq \alpha_c z_c, \quad (4.34)$$

siendo $H^{(l)}$, $l = 0, 1, \dots$, las matrices de iteración definidas en (4.10) y $\alpha_c < 1$.

Con esta notación, y siguiendo los mismos pasos que en la demostración del Teorema 7 del Capítulo 3 podemos escribir

$$\epsilon^{(j)} = S_j \bar{z}_{j+r-1}, \quad (4.35)$$

donde S_j , $j = 0, 1, \dots$, es la matriz de tamaño $n \times nK$ dividida en bloques de tamaño $n \times n$ y que tiene un bloque identidad de orden n en la posición r , y los $K - 1$ bloques restantes son nulos (ver expresión (3.39)).

Analizamos seguidamente la expresión del error

$$\epsilon^{(j+r)} = z^{(j+r)} - \xi.$$

siendo ξ la solución exacta del sistema lineal $Az = b$. Puesto que ξ es un punto fijo de los operadores P_i , tal y como se probó en la demostración del Lema 2, se tiene

$$\begin{aligned} \epsilon^{(j+r)} &= \mathbf{x}^{(j+r)} - \xi \\ &= (I - E_i)\mathbf{x}^{(j+r,-1)} + E_i \left(\omega \sum_{i=1}^r E_i P_i^{q(j,i)} + (1 - \omega)I \right) \mathbf{x}^{(j)} \\ &\quad - (I - E_i)\xi - E_i \left(\omega \sum_{i=1}^r E_i P_i^{q(j,i)} + (1 - \omega)I \right) \xi \\ &= (I - E_i)\epsilon^{(j+r,-1)} + E_i \left(\omega \sum_{i=1}^r E_i P_i^{q(j,i)} + (1 - \omega)I \right) (\mathbf{x}^{(j)} - \xi) \end{aligned}$$

Ahora bien, la expresión $\omega \left(\sum_{i=1}^r E_i P_i^{q(j,i)} + (1 - \omega)I \right) (\mathbf{x}^{(j)} - \xi)$ ha sido analizada en el Lema 2, demostrando que

$$\omega \left(\sum_{i=1}^r E_i P_i^{q(j,i)} + (1 - \omega)I \right) (\mathbf{x}^{(j)} - \xi) = H^{(j)} \epsilon^{(j)}, \quad (4.36)$$

donde $H^{(j)}$ está definida en (4.10) y satisface (4.34).

Por tanto, de las expresiones (4.35) y (4.36), y siguiendo los mismos pasos que el Teorema 7 del Capítulo 3, se comprueba fácilmente que

$$\epsilon^{(j+r)} = (I - E_i)\epsilon^{(j+r,-1)} + E_i H^{(j)} S_j \bar{\epsilon}_{j+r,-1}.$$

Obtenemos por tanto, que

$$\bar{\epsilon}_{j+r} = B_{j+r} \bar{\epsilon}_{j+r,-1}$$

donde B_{j+r} , es la matriz cuadrada de orden nK , dada por

$$B_{j+r} = \begin{bmatrix} I - E_i & O & \dots & O & O \\ I & O & \dots & O & O \\ \vdots & \vdots & & \vdots & \vdots \\ O & O & \dots & I & O \end{bmatrix} + \begin{bmatrix} E_i, H^{(j)} S_j \\ O \\ \vdots \\ O \end{bmatrix}. \quad (4.37)$$

Entonces

$$\bar{z}_{j+2K-1} = B_{j+2K-1} B_{j+2K-2} \cdots B_{j+1} \bar{z}_j.$$

Razonando a partir de aquí, como en el Teorema 7 del Capítulo 3 y utilizando la desigualdad de la expresión (4.34), se prueba que existe una constante real γ , con $0 \leq \gamma < 1$, tal que

$$|B_{j+2K-1} B_{j+2K-2} \cdots B_{j+1}| \bar{z} \leq \gamma \bar{z}, \quad (4.38)$$

donde $\bar{z} \in \mathbb{R}^{nK}$ está definido en (4.33). De aquí, si consideramos la norma matricial inducida $\|\cdot\|_{\bar{z}}$ introducida en la Sección 1.1, se deduce que

$$\|B_{j+2K-1} B_{j+2K-2} \cdots B_{j+1}\|_{\bar{z}} \leq \gamma.$$

y entonces

$$\lim_{j \rightarrow \infty} \bar{z}_j = 0,$$

con lo que queda demostrada la convergencia del esquema (4.31). ■

4.4 Experimentos numéricos

Esta sección está dedicada al estudio, desde el punto de vista experimental, de los métodos caóticos paralelos síncrono y asíncrono estudiados en este

capítulo. Los esquemas teóricos correspondientes han sido planteados en las expresiones (4.7) y (4.31). Este estudio se ha realizado atendiendo a los mismos objetivos que los planteados en la Sección 3.4.

Tal y como se dijo en dicha sección, tomamos como problema modelo la obtención de una solución aproximada de la ecuación de Laplace. Dicho problema, al ser discretizado, nos conduce a la resolución de un gran sistema de ecuaciones lineales $Az = b$. La matriz de coeficientes A viene definida en la expresión (3.50).

Presentamos a continuación, las implementaciones de los algoritmos caóticos tanto síncrono como asíncrono estudiados en este capítulo.

4.4.1 Algoritmos

Se desea resolver el sistema lineal $Az = b$. Supongamos r particiones de la matriz A

$$A = F_j - G_j, \quad j = 1, 2, \dots, r.$$

Los siguientes algoritmos detallan los pasos a seguir para generar la sucesión de vectores que converge a la solución de dicho sistema, a partir del esquema caótico síncrono (4.7) y del esquema caótico asíncrono (4.31).

**ALGORITMO SÍNCRONO****Etapa 0. Lectura de datos:**

Matriz de coeficientes: F
 Término independiente: b
 Factores caóticos: $q(l, j)$, $1 \leq j \leq r$, $l = 0, 1, \dots$
 Factor de relajación: ω

Etapa 1. Elegir $x^{(0)}$ arbitrariamente y $l = 0$.

Etapa 2. (Etapa caótica) En paralelo,

Para $j = 1, 2, \dots, r$

Para $k = 1, 2, \dots, q(l, j)$, con $x_j^{(0)} = x^{(l)}$.

Resolver el sistema $F_j x_j^{(k)} = G_j x_j^{(k-1)} + b_j$.

Etapa 3. (Etapa síncrona) $x^{(l+1)} = \omega \sum_{j=1}^r E_j x_j^{(q(l,j))} + (1 - \omega)x^{(l)}$, siendo E_j , matrices diagonales no negativas cuya suma es la identidad.

Etapa 4. Comprobación del criterio de convergencia

$$\sum_{k=1}^n |x_k^{(l+1)} - x_k^{(l)}| < 5 \cdot 10^{-10}.$$

- Si test de convergencia = VERDADERO, entonces FIN.
- Si test de convergencia = FALSO, entonces

— $l \leftarrow l + 1$.

— Volver a la etapa 3.

ALGORITMO ASÍNCRONO**Etapa 0. Lectura de datos:**

Matriz de coeficientes: A
Término independiente: b
Factores caóticos: $q(l, j), 1 \leq j \leq r, l = 0, 1, \dots$
Factor de relajación: ω

Etapa 1. Elegir $z^{(0)}$ arbitrariamente y $l_j = 0, 1 \leq j \leq r$. En paralelo, para $j = 1, 2, \dots, r$ (etapas 2-5).

Etapa 2. $z^{(1,j)} = \text{comms}$, donde comms es un vector compartido por todos los procesadores.

Etapa 3. Para $k = 1, 2, \dots, q(l_j, j)$, con $z_j^{(0)} = z^{(1,j)}$.

Resolver el sistema $F_j z_j^{(k)} = G_j z_j^{(k-1)} + b_j$.

$$z_j^{(k+1,j)} = \omega z_j^{(k,j)} + (1 - \omega) z^{(1,j)}.$$

Etapa 4. $\text{comms} = (I - E_j) \text{comms} + E_j z_j^{(k+1,j)}$, siendo E_j matrices diagonales no negativas cuya suma es la identidad.

Etapa 5. Comprobación del criterio de convergencia.

• Si $\text{test de convergencia} = \text{VERDADERO}$, entonces FIN.

• Si $\text{test de convergencia} = \text{FALSO}$, entonces

– $l_j \leftarrow l_j + 1$.

– Volver a la etapa 2.

4.4.2 Resultados numéricos

Antes de comentar los resultados numéricos obtenidos, vamos a detallar la forma de las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$, que se han utilizado para desarrollar los distintos algoritmos.

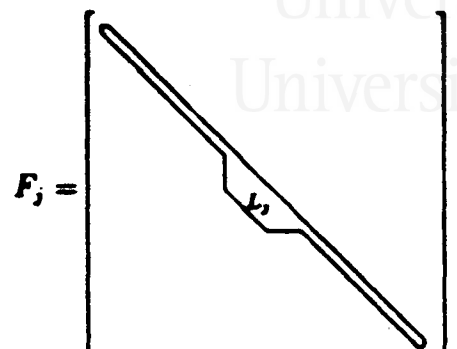
Básicamente, las particiones utilizadas son las que se expusieron en el Ejemplo 2, sin contemplar solapamiento. Es decir si representamos la matriz A de la siguiente forma

$$A = \begin{bmatrix} D_1 & & & & \\ & -I & & & \\ & & D_2 & & \\ & & & -I & \\ & & & & \ddots & \\ & & & & & -I & \\ & & & & & & D_r \end{bmatrix}$$

donde D_j , $j = 1, 2, \dots, r$, están definidas por

$$D_j = \begin{bmatrix} B & -I & & & & \\ -I & B & -I & & & \\ & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & -I \\ & & & & -I & B \end{bmatrix},$$

entonces las matrices F_j , $j = 1, 2, \dots, r$, que consideramos son

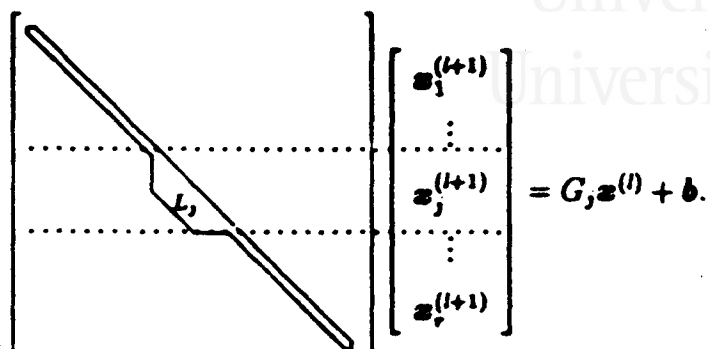


donde cada matriz L_j , $j = 1, 2, \dots, r$, es la parte triangular inferior de D_j . Las matrices G_j , $j = 1, 2, \dots, r$, se eligen de forma que $A = F_j - G_j$.

Cada matriz E_j , $j = 1, 2, \dots, r$, tiene un bloque identidad en la posición correspondiente al bloque triangular inferior de la matriz F_j , y todos los restantes elementos son nulos.

Estas particiones son claramente H -compatibles y además la matriz A es una H -matriz tal y como se explicó en la Sección 3.4. Por tanto, el Teorema 2 asegura la convergencia para el caso síncrono y el Teorema 3 para el caso asíncrono.

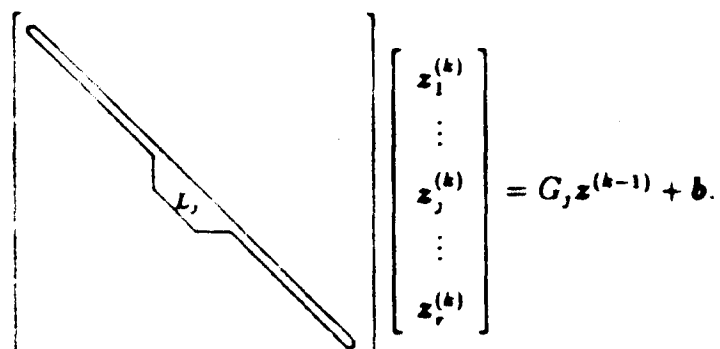
Sea $\mathbf{z}_j^{(i+1)}$ el bloque de vector $\mathbf{z}^{(i+1)}$ asociado al bloque triangular L_j de F_j . En un esquema iterativo de multipartición no caótico ($q(l, j) = 1$), el procesador j sólo tiene que actualizar el bloque $\mathbf{z}_j^{(i+1)}$ resolviendo el sistema triangular representado en la siguiente figura



$$\begin{bmatrix} \vdots \\ \vdots \\ z_j^{(l+1)} \\ \vdots \\ \vdots \end{bmatrix} = G_j z^{(l)} + b.$$

Ahora bien, si al procesador j se le asigna un factor caótico superior a 1, pongamos por ejemplo $q(l, j) = 2$, entonces los cálculos que debe realizar y que vienen indicados en la etapa 2 del algoritmo síncrono son los siguientes.

Para $k = 1, 2$, con $z^{(0)} = z^{(l)}$



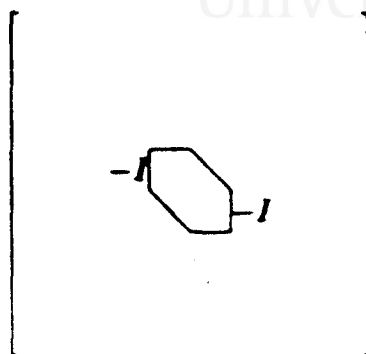
$$\begin{bmatrix} z_1^{(k)} \\ \vdots \\ z_j^{(k)} \\ \vdots \\ z_r^{(k)} \end{bmatrix} = G_j z^{(k-1)} + b.$$

Debido a la forma de la matriz A y de las particiones elegidas, para $k = 1$ el procesador j deberá calcular las variables

$$z_{j-1}^{(1)}, z_j^{(1)}, z_{j+1}^{(1)}, \quad (4.39)$$

donde el tamaño de $z_j^{(1)}$ corresponde al tamaño de L_j , y las demás son de tamaño correspondiente a los bloques B de A . Esto es debido a los dos bloques identidad que nos encontramos en el bloque fila j de A tal y como muestra la

siguiente figura, lo cual nos indica que las nuevas variables indicadas en (4.39) son necesarias para el posterior cálculo de $z_j^{(2)}$.



Para $k = 2$ sólo es necesario calcular $z_j^{(2)}$. Estas componentes son las únicas que se necesitan en esta última etapa local para actualizar la solución global, $z_j^{(i+1)} = z_j^{(2)}$.

Esta situación se puede generalizar para cualquier valor de los factores caóticos, de forma que cuando mayores son estos factores, mayor es el subsistema que se ha de resolver. Nótese que el crecimiento en el subsistema que se ha de resolver, depende también del tamaño de los bloques B , ya que cuanto mayores sean estos bloques mayor será dicho subsistema. Estas consideraciones nos hacen pensar en la idea de escoger factores caóticos no excesivamente grandes y tamaños de la matriz B pequeños, de manera que el aumento en el orden de los subsistemas que cada procesador debe resolver no sea elevado.

La Tabla 1 explica la forma de las matrices con las que realizamos los experimentos correspondientes a este capítulo. Nótese que todas estas matrices corresponden con las utilizadas en el capítulo anterior, excepto la matriz de orden 4096. Los bloques diagonales B de esta matriz son de orden 64 que como se observa es bastante menor que el considerado para las matrices de ordenes 2816

Orden A	Orden B	Bloques por proc.	Filas por proc.
384	32	2,2,2,2,1,1,1,1	64,64,64,64,32,32,32,32
1024	32	4,4,4,4,4,4,4,4	128,128,128,128,128,128,128,128
2816	256	2,2,2,1,1,1,1,1	512,512,512,256,256,256,256,256
4096	64	16,16,8,8,4,4,4,4	1024,1024,512,512,256,256,256,256
5632	512	2,2,2,1,1,1,1,1	1024,1024,1024,512,512,512,512,512

Tabla 1: Matrices de prueba.

y 5632. Esto evitará, tal y como hemos comentado anteriormente, que cuando el factor caótico sea grande el procesador correspondiente tenga que resolver un subsistema de un orden demasiado elevado. En dicha tabla, se presenta también, el número de bloques B que forman parte del bloque triangular de cada matriz F_j , $j = 1, 2, \dots, 8$. Es con este bloque triangular con el que esencialmente cada procesador trabaja. Damos además el número de filas por procesador, para de esta forma hacernos una idea de la cantidad de trabajo que cada procesador debe realizar.

En la Tabla 2 se presentan, para estas matrices, los resultados obtenidos para diferentes modelos caóticos sincrónicos ejecutados en secuencial. Se ha obtenido experimentalmente, para cada uno de estos esquemas, el factor de relajación óptimo, observando en todos los casos que este factor es menor para el correspondiente modelo no caótico, es decir, cuando los factores caóticos son iguales a la unidad. Esta situación se generaliza al rango de convergencia del factor de relajación, dando lugar a que frecuentemente el factor de relajación óptimo,



Orden	$\sigma(l, j), 1 \leq j \leq 8$	steps	Iter.	Tiempo (seg.)
204	1,1,1,1,2,2,2,2	1.525	520	5.820
	2,2,2,2,4,4,4,4	1.725	254	4.820
	2,2,2,2,2,2,2,2	1.725	208	3.780
	2,2,2,2,6,6,6,6	1.800	192	5.877
	1,1,1,1,1,1,1,1	1.200	603	6.132
1024	2,2,2,2,2,2,2,2	1.725	607	25.520
	2,2,2,2,2,2,2,2	1.825	806	27.260
	1,1,1,1,1,1,1,1	1.675	2066	49.366
2048	2,2,2,4,4,4,4,4	1.725	232	34.513
	1,1,1,2,2,2,2,2	1.425	514	40.711
	2,2,2,2,2,2,2,2	1.700	300	27.126
	1,1,1,1,1,1,1,1	1.275	719	45.364
4096	2,2,4,4,6,6,6,6	1.775	2544	446.082
	1,1,2,2,3,3,3,3	1.800	5720	685.197
	1,1,1,1,1,1,1,1	1.675	7406	699.618
1632	2,2,2,4,4,4,4,4	1.725	230	69.117
	1,1,1,2,2,2,2,2	1.425	531	86.243
	2,2,2,2,2,2,2,2	1.700	310	56.886
	1,1,1,1,1,1,1,1	1.250	786	101.803

Tabla 2: Algoritmo secuencial.

correspondiere a un modelo caótico, sea un factor para el que el esquema no caótico no converge.

Los resultados obtenidos para los esquemas caóticos sincronos ejecutados sobre los 8 procesadores de los que dispone el multiprocesador Alliant FX/80, se muestran en la Tabla 3. Al igual que en el Capítulo 3, hemos utilizado dos directivas de compilación, obteniendo el tiempo de CPU para una ejecución concurrente (T_1), y el tiempo de CPU para una ejecución vectorial-concurrente (T_2).

Uno de nuestros objetivos es el comparar los algoritmos caóticos con los no caóticos. Ahora bien, las consideraciones realizadas sobre el distinto rango de convergencia para el factor de relajación, en los esquemas caóticos y no caóticos, no nos permiten calcular un speed-up que compare los dos esquemas utilizando el mismo factor de relajación. Es por esto que comparamos, en la Tabla 3, cada modelo caótico sincrónico con el modelo secuencial no caótico para su factor de relajación óptimo (ver Tabla 2) mediante el cálculo de S_i , $i = 1, 2$.

$$S_i = \frac{T_o}{T_i}, \quad i = 1, 2,$$

donde T_o es el tiempo de CPU del algoritmo secuencial no caótico para su factor de relajación óptimo. Nótese que este speed-up no está acotado superiormente por el número de procesadores ya que no comparamos el mismo algoritmo.

Un speed-up que sí está acotado por el número de procesadores utilizados, en ejecución sólo concurrente, es el que compara un algoritmo caótico paralelo con el correspondiente caótico secuencial. Este speed-up viene dado por S'_i , donde

$$S'_i = \frac{T_{co}}{T_i}, \quad i = 1, 2,$$

siendo T_{co} el tiempo de CPU del modelo caótico que se está evaluando ejecutado en un único procesador.

Se observa, en la Tabla 3, que todos los modelos caóticos sincrónico, estudiados experimentalmente, aceleran al no caótico. Esta situación está especialmente marcada para la matriz de orden 4096, en la que para el esquema caótico sincrónico $2^3 4^3 6^6$ el tiempo de CPU disminuye en prácticamente un 50% respecto del no caótico. Esto es debido a la gran disminución que se produce en el número de iteraciones globales. Esta situación es bastante general, con independencia

Orden	$\phi(i,j), 1 \leq j \leq 8$	wage	Iter.	Tiempo (seg.)		Speed-up			
				T_1	T_2	S_1	S_2	S'_1	S'_2
384	1,1,1,1,2,2,2,2	1.525	529	1.091	0.482	5.62	12.72	5.34	12.09
	2,2,2,2,4,4,4,4	1.725	254	0.991	0.514	6.18	11.83	4.87	9.39
	2,2,2,2,2,2,2,2	1.725	288	0.677	0.346	9.05	17.72	5.58	10.92
	3,3,3,3,6,6,6,6	1.600	192	1.250	0.682	4.91	8.99	4.70	8.62
	1,1,1,1,1,1,1,1	1.380	663	1.155	0.482	5.31	13.56	5.31	13.56
1024	3,3,3,3,3,3,3,3	1.725	667	3.884	1.926	12.70	25.62	6.57	13.25
	2,2,2,2,2,2,2,2	1.825	885	4.577	1.909	10.78	25.85	5.95	14.27
	1,1,1,1,1,1,1,1	1.475	2068	8.561	3.073	5.77	16.06	5.77	16.06
2048	2,2,2,4,4,4,4,4	1.725	232	5.310	2.526	6.73	18.36	6.50	13.06
	1,1,1,2,2,2,2,2	1.425	514	7.476	2.967	6.20	15.63	5.45	13.62
	2,2,2,2,2,2,2,2	1.700	300	5.024	2.299	9.23	20.16	5.40	11.80
	1,1,1,1,1,1,1,1	1.275	719	8.876	3.284	5.22	13.70	5.22	13.70
4096	2,2,4,4,6,6,6,6	1.775	2544	79.553	40.864	8.79	17.12	5.61	10.91
	1,1,2,2,3,3,3,3	1.800	5739	118.119	51.183	5.92	13.67	5.63	12.99
	1,1,1,1,1,1,1,1	1.475	7486	147.828	61.867	4.73	11.29	4.73	11.29
5632	2,2,2,4,4,4,4,4	1.725	239	11.425	6.342	8.91	16.06	6.04	10.90
	1,1,1,2,2,2,2,2	1.425	531	15.789	7.062	6.44	14.42	5.40	12.07
	2,2,2,2,2,2,2,2	1.700	310	10.826	5.452	9.40	15.63	5.16	10.25
	1,1,1,1,1,1,1,1	1.250	756	19.510	7.521	5.21	13.53	5.21	13.53

Tabla 3: Resultados de los esquemas caóticos sincronos para ω óptimo.

del factor de relajación como se refleja en las Figuras 1, 2, 3, 4 y 5. Estas figuras comparan el tiempo de CPU en ejecución concurrente de distintos esquemas caóticos sincronos, para cada una de las matrices de prueba y un amplio rango de factores de relajación. Destacamos los resultados obtenidos para las matrices de ordenes 1024 y 4096 para las que todo modelo caótico estudiado experimentalmente es mejor que el no caótico (vease Figuras 2 y 4).

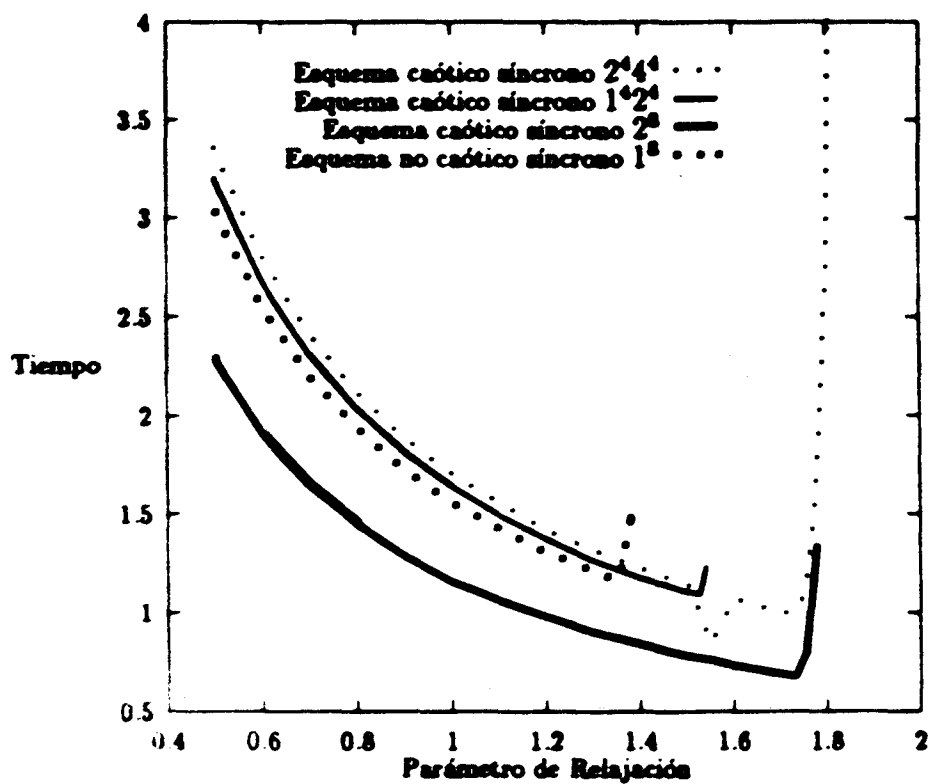


Figura 1: Comparación esquemas paralelos caóticos sincrónicos, $n = 384$.

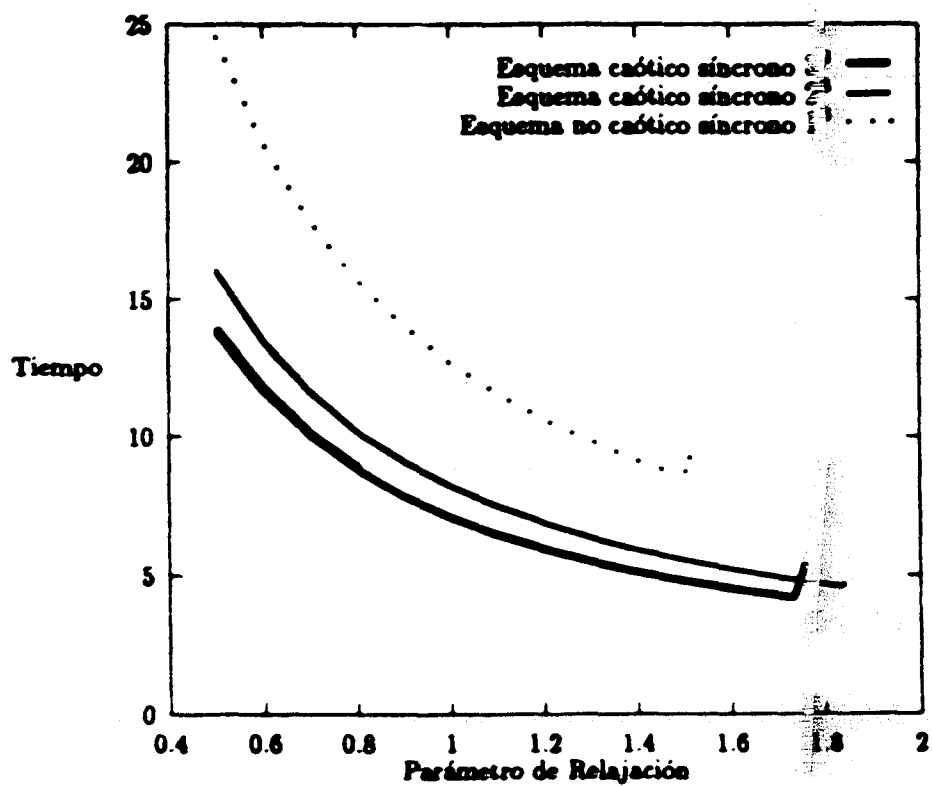


Figura 2: Comparación esquemas paralelos caóticos síncronos, $n = 1024$.

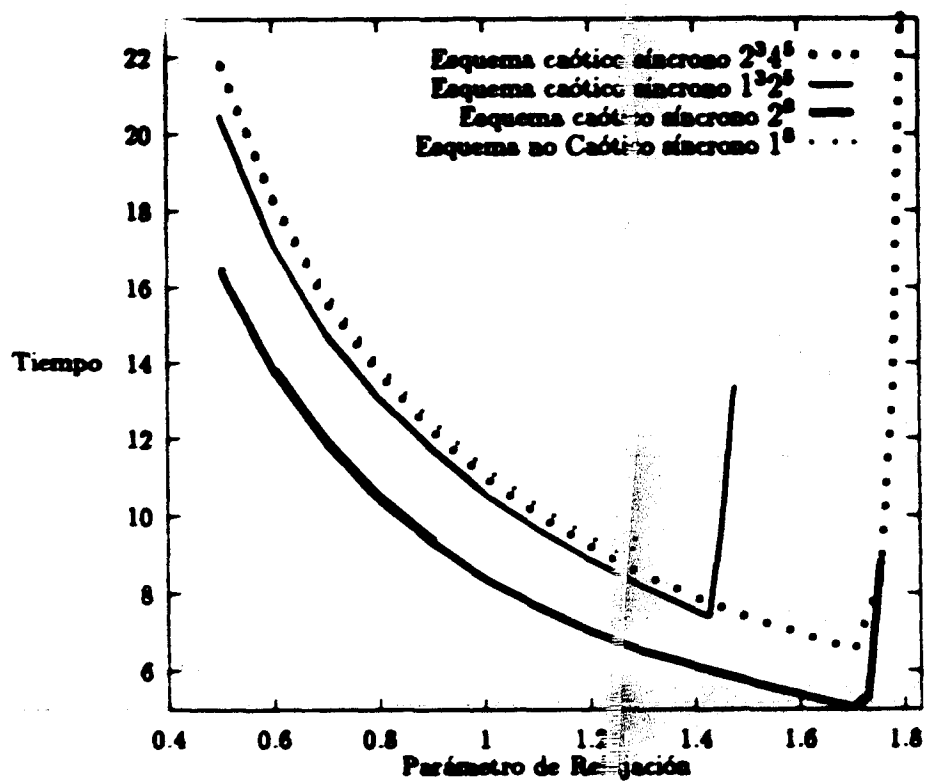


Figura 3: Comparación esquemas paralelos caóticos síncronos, $n = 2816$.

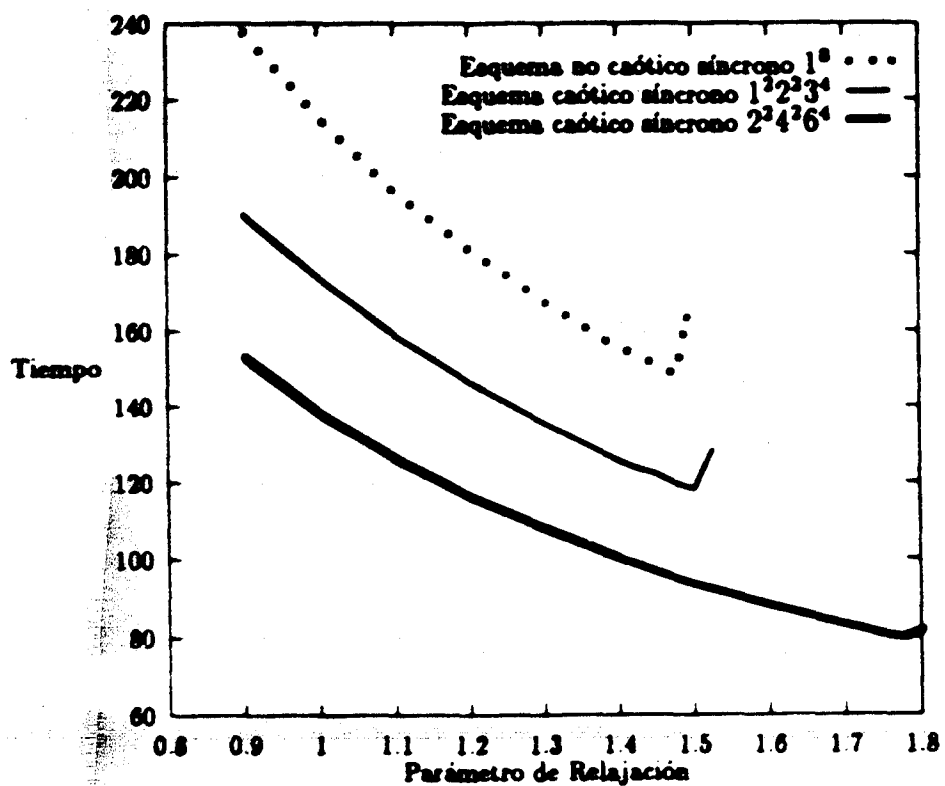


Figura 4: Comparación esquemas paralelos caóticos sincrónicos, $n = 4096$.

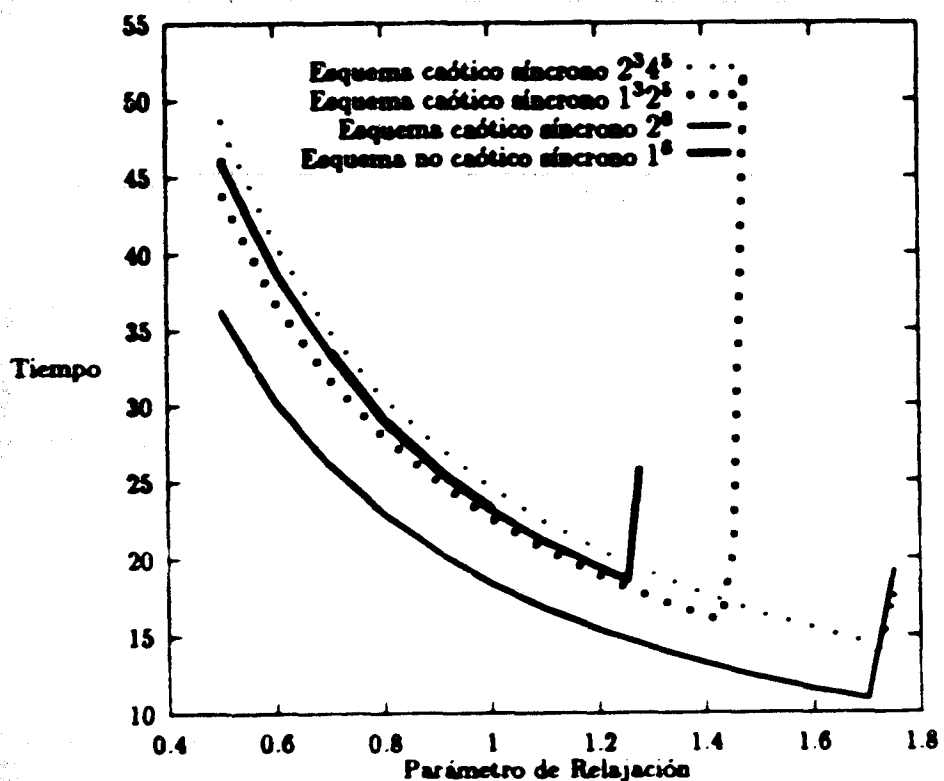


Figura 5: Comparación esquemas paralelos caóticos sincrónicos, $n = 5632$.

Aunque los resultados presentados hasta ahora en este capítulo, muestran que los modelos caóticos sincrónicos son buenos, estos son acelerados de forma clara, y en todos los casos por los correspondientes caóticos asincrónicos. Mostramos esta situación genérica para algunos modelos caóticos de la Tabla 3, en las Figuras 6, 7, 8 y 9.

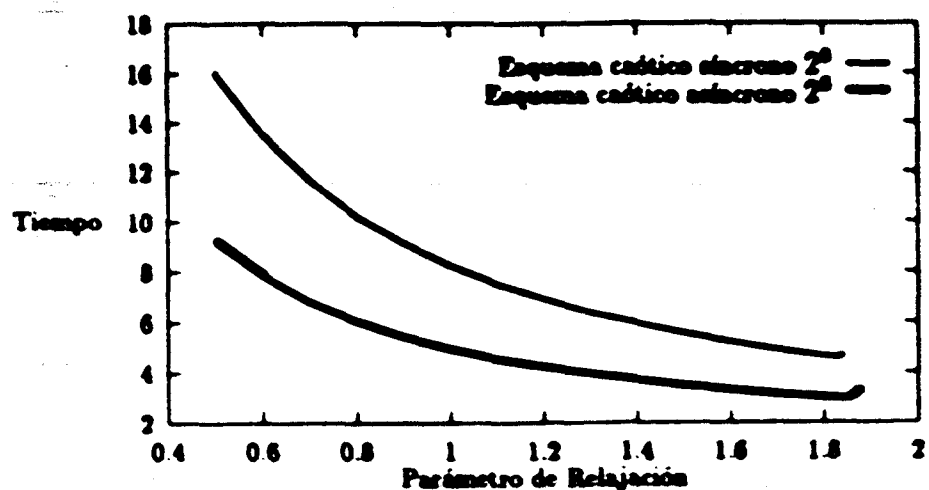


Figura 6: Comparación esquemas paralelos síncrono y asíncrono, $n = 1024$.

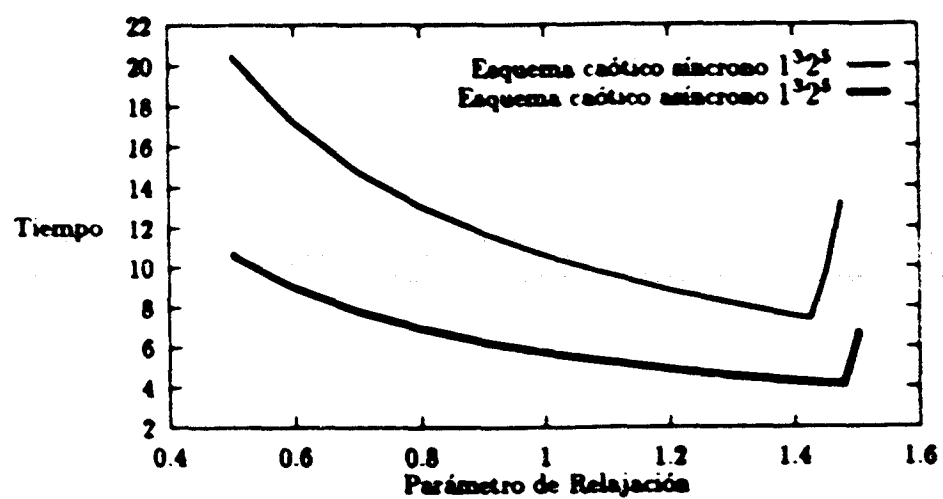


Figura 7: Comparación esquemas paralelos síncrono y asíncrono, $n = 2816$.

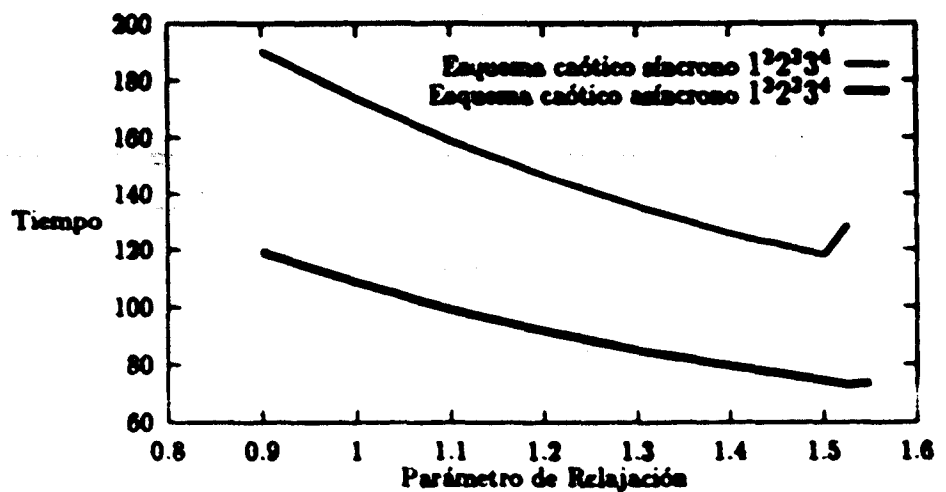


Figura 8: Comparación esquemas paralelos síncrono y asíncrono, $n = 4096$.

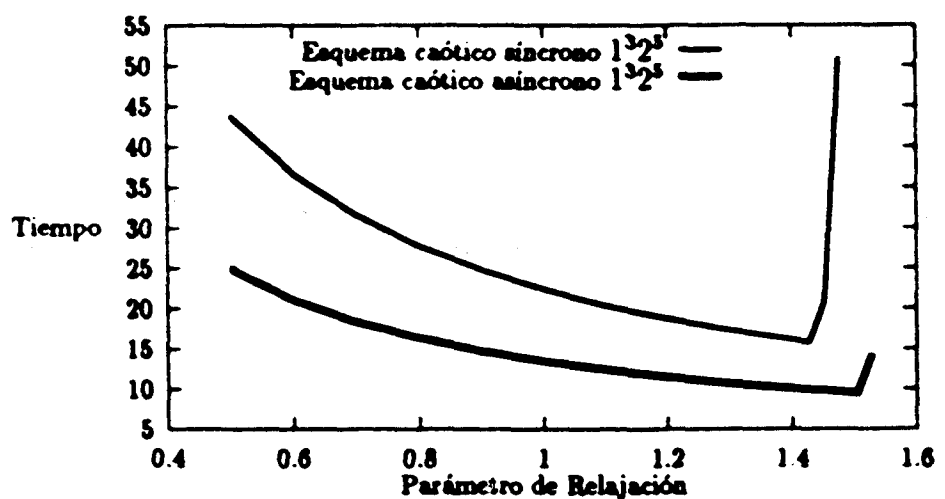


Figura 9: Comparación esquemas paralelos síncrono y asíncrono, $n = 5632$.

Los resultados para los esquemas iterativos caóticos asíncronos pueden verse en la Tabla 4; $\text{iter}(j)$, $j = 1, 2, \dots, 8$, es el número de iteraciones que realiza cada procesador. Esta tabla muestra que para el factor de relajación óptimo siempre encontramos un modelo caótico asíncrono que acelera al modelo asíncrono no caótico. Como en la Tabla 3, hemos calculado el tiempo de *CPU* para los dos tipos de ejecución, concurrente y vectorial-concurrente, denotados por T_1 y T_2 respectivamente.

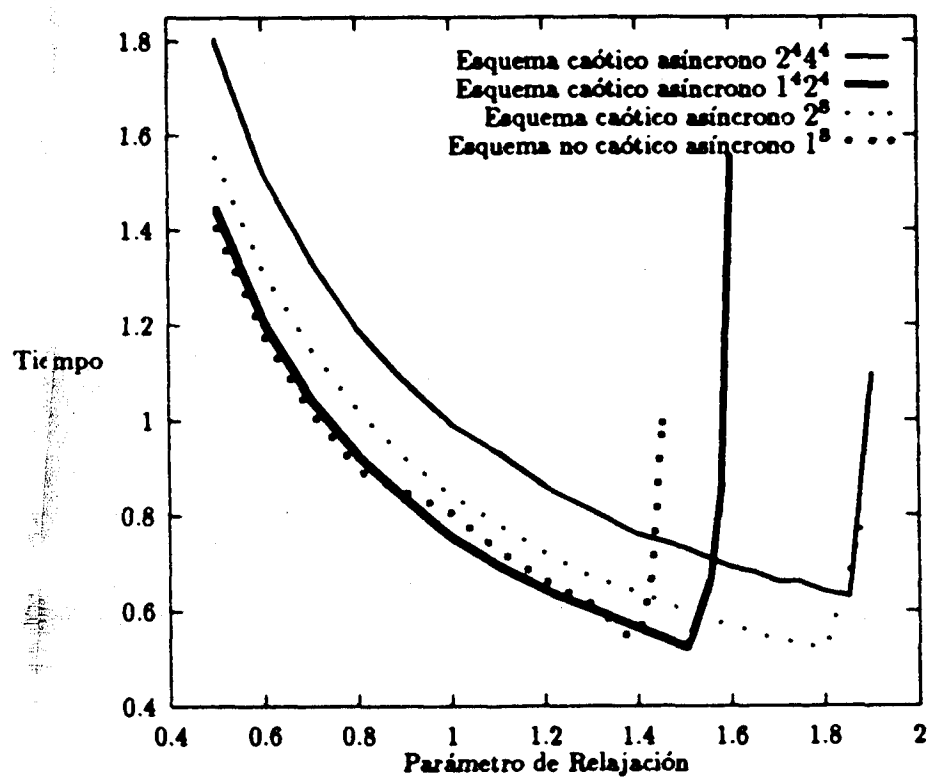
Como ya comentamos en el capítulo anterior, debido a la construcción de los modelos asíncronos, no es factible el planteamiento de una versión secuencial. Es por esto que para dar una medida del incremento de velocidad que se produce con la utilización de los modelos asíncronos se ha tomado como referencia el mejor algoritmo secuencial de los reflejados en la Tabla 2, obteniendo así los respectivos *speed-up's* para los modos de ejecución concurrente y vectorial-concurrente, denotados por S_i , $i = 1, 2$. Como no se está comparando el mismo algoritmo ejecutado en secuencial y en paralelo puede surgir la situación en la que se obtenga un valor del *speed-up* S_1 , superior al número de procesadores, como puede apreciarse en la Tabla 4.

Para finalizar esta sección, comparamos para cada matriz de prueba los distintos esquemas caóticos asíncronos calculando el tiempo de *CPU* en ejecución concurrente, atendiendo al factor de relajación. Este estudio puede analizarse en las Figuras 10, 11, 12, 13 y 14. En ellas se ha observado situaciones análogas a las ya comentadas para el caso síncrono.

Orden	$q(i,j), 1 \leq j \leq 8$	w_{opt}	$har(j), 1 \leq j \leq 8$	Tiempo (seg.)		Speed-up	
				T_1	T_2	S_1	S_2
384	1,1,1,1,2,2,2,2	1.525	632,61 : 631,618,463,446,464,480	0.525		7.20	
		1.55	636,617 : 635,618,515,499,515,508		0.424		8.91
	2,2,2,4,4,4,4,4	1.85	470,35 : 406,389,175,185,212,267	0.634		5.96	
		1.825	400,36 : 371,369,309,213,240,266		0.636		8.66
	2,2,2,2,2,2,2,2	1.80	381,319,327,321,455,445,452,547	0.522		7.24	
		1.85	370,32 : 335,331,485,469,484,520		0.397		9.52
	3,3,3,3,6,6,6,6	1.775	433,325 : 329,326,126,137,159,193	0.895		4.22	
		1.775	384,27 : 301,294,147,154,176,202		0.547		6.91
1,1,1,1,1,1,1,1	1.375	635,635 : 647,634,1006,981,1002,977	0.543		6.96		
	1.40	680,65 : 683,663,984,960,991,965		0.464		8.15	
1024	3,3,3,3,3,3,3,3	1.80	911,76 : 772,761,766,763,772,806	3.278		7.78	
		1.80	867,77 : 788,772,789,771,788,845		2.423		10.53
	2,2,2,2,2,2,2,2	1.85	1115,1007 : 1024, 992,1026,1013,1026,1093	2.889		8.84	
		1.85	1110,1032,1057,1025,1064,1030,1057,1068		2.307		11.06
	1,1,1,1,1,1,1,1	1.525	2103,2057,2106,2057,2106,2056,2104,2056	3.732		6.83	
		1.525	2247,2173,2247,2173,2247,2173,2247,2172		3.080		8.29
2816	1,1,1,2,2,2,2,2	1.475	642,625,636,455,464,443,453,577	3.902		6.80	
		1.50	593,562,617,540,558,536,562,575		3.188		8.51
	2,2,2,4,4,4,4,4	1.80	515,42 : 437,191,194,203,235,283	5.254		5.16	
		1.80	396,35 : 385,223,235,239,267,281		3.372		8.05
	2,2,2,2,2,2,2,2	1.80	362,35 : 305,442,432,426,449,553	3.864		7.02	
		1.825	330,35 : 318,448,468,459,468,491		2.643		10.27
	1,1,1,1,1,1,1,1	1.35	603,5 : 604,857,977,948,977,945	4.094		6.62	
		1.325	669,65 : 655,961,1007,962,1006,888		3.432		7.90
5632	1,1,1,2,2,2,2,2	1.50	638,502,637,483,506,496,471,615	9.229		6.05	
		1.525	621,52 : 615,532,559,544,572,565		7.245		7.71
	2,2,2,4,4,4,4,4	1.75	553,71 : 478,206,214,219,261,323	12.483		4.47	
		1.75	396,62,363,222,223,228,250,289		7.596		7.35
	2,2,2,2,2,2,2,2	1.80	339,199,296,435,429,414,440,473	7.286		7.67	
		1.80	336,313,329,451,476,430,461,461		6.450		8.66
	1,1,1,1,1,1,1,1	1.3	629,65 : 597,1023,1024,970,1024,1005	8.862		6.30	
		1.35	684,64 : 667,976,1008,937,994,895		7.635		7.32

Tabla 4: Resultados de los esquemas asíncronos para ω óptimo.

4096	2,2,4,4,6,6,6,6	1.80	2874,2816,2821,2809,2120,2082,2083,2822	49.911	8.93
		1.80	2749,2724,2790,2683,2780,2580,2654,3120	40.909	10.90
	1,1,2,2,3,3,3,3	1.55	5739,5739,6415,6323,6906,6713,7024,8010	72.192	6.18
		1.55	5850,5475,6471,6108,7273,7149,7273,7196	55.181	8.08
	1,1,1,1,1,1,1,1	1.525	5599,5942,10609,10458,10414,13639,15378,15127	81.724	5.46
		1.525	6004,6005, 9881, 9726,14229,13521,14214,12520	61.413	7.26

Tabla 4. (Continuación) Resultados de los esquemas asíncronos para ω óptimo.Figura 10: Comparación esquemas paralelos caóticos asíncronos, $n = 384$.

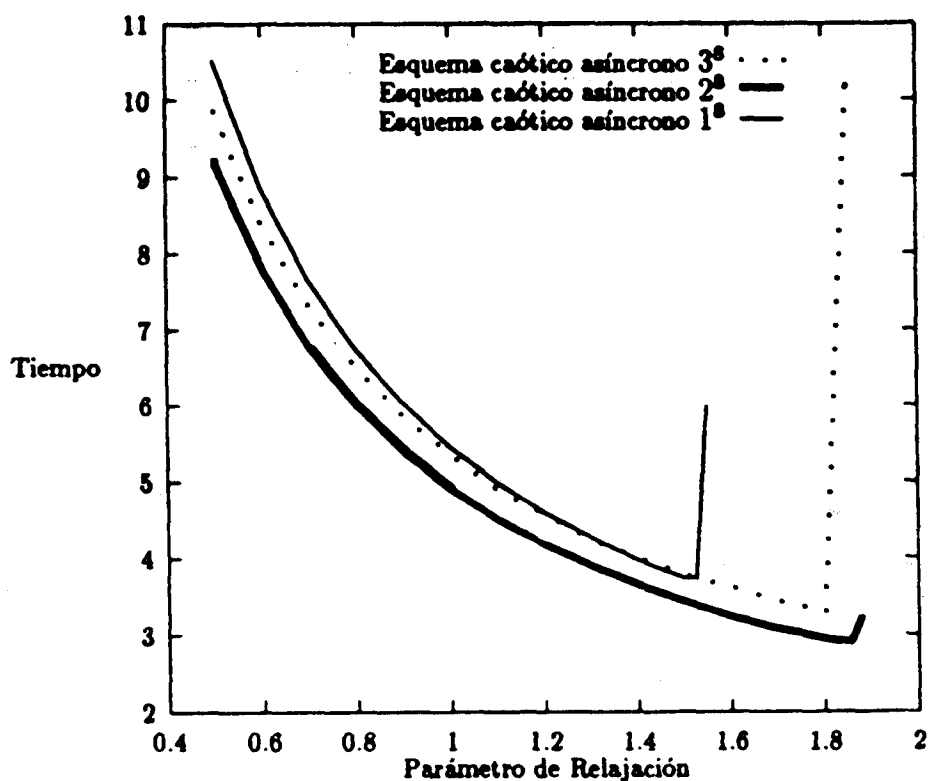
Universitat d'Alacant
Universidad de Alicante

Figura 11: Comparación esquemas paralelos caóticos asíncronos, $n = 1024$.

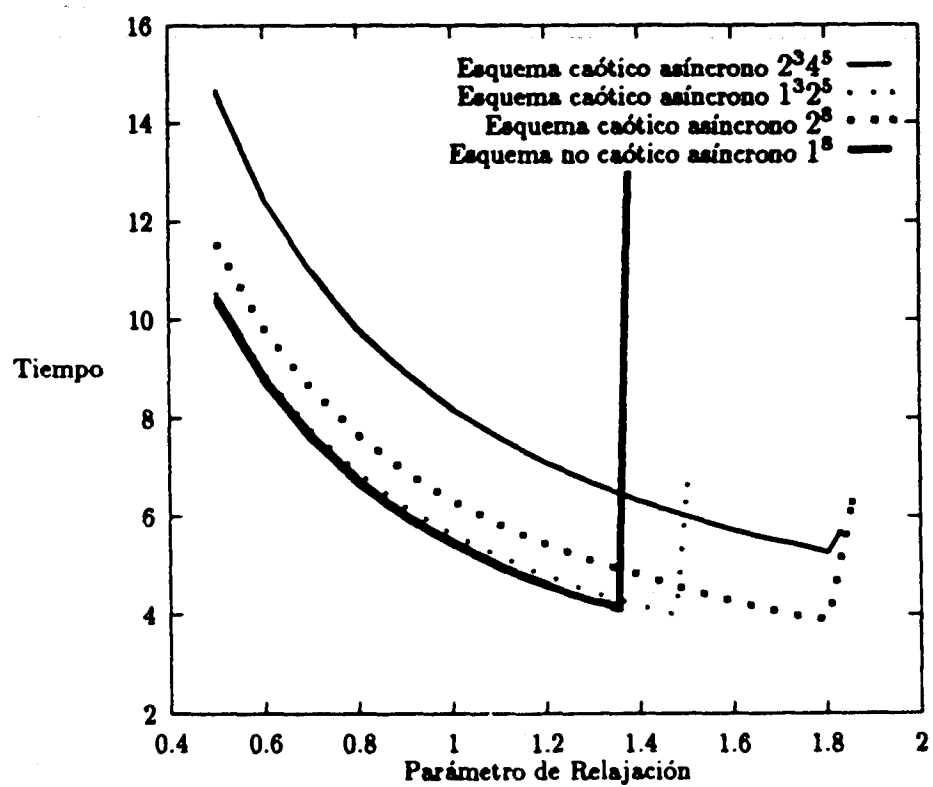


Figura 12: Comparación esquemas paralelos caóticos asíncronos, $n = 2816$.

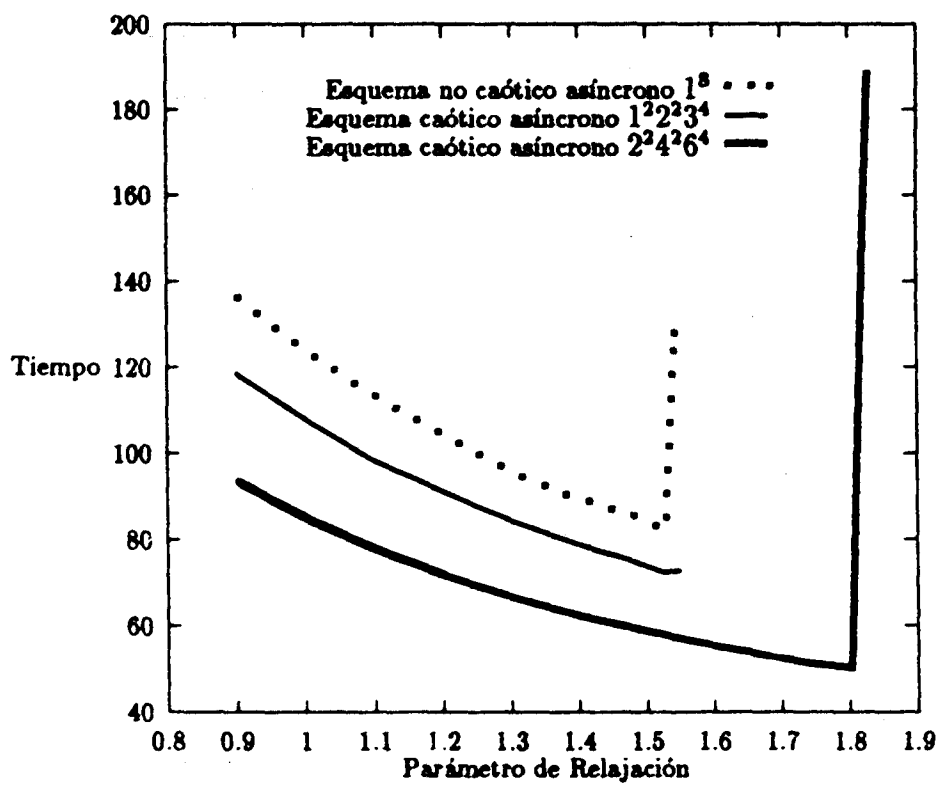


Figura 13: Comparación esquemas paralelos caóticos asíncronos, $n = 4096$.

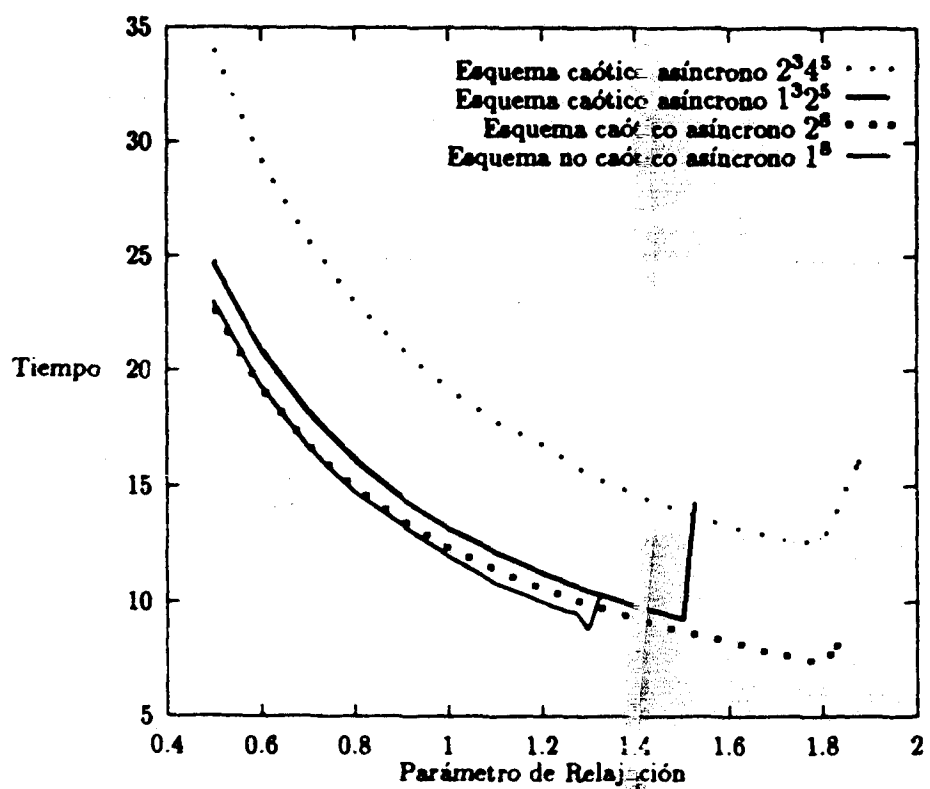


Figura 14: Comparación esquemas paralelos caóticos asíncronos, $n = 5632$.



4.5 Conclusiones

En este capítulo hemos estudiado dos esquemas paralelos caóticos, en sentido de Bru, Elsner y Neumann [10], introduciendo un parámetro de relajación, como lo hacen O'Leary y White [61].

El primero de estos esquemas (ver expresión (4.7)) corresponde a un proceso iterativo caótico síncrono mientras que el segundo es su correspondiente versión asíncrona (ver expresión (4.31)). El estudio teórico de la convergencia se resume en el Teorema 2, para el caso síncrono, y en el Teorema 3 para la versión asíncrona. En ambos casos se supone básicamente que la matriz A es una H -matriz y que las particiones $A = F_j - G_j$, $j = 1, 2, \dots, r$, son H -compatibles.

Los modelos planteados por Bru, Elsner y Neumann en [10] son un caso particular de los estudiados en este capítulo; ellos estudiaron la convergencia de sus modelos para matrices monótonas. En este capítulo se ha obtenido un nuevo criterio de convergencia para dichos modelos que no había sido considerado todavía en la literatura existente sobre este tema; se exige que la matriz del sistema lineal sea una H -matriz, matrices que no son necesariamente monótonas.

Estos esquemas se han evaluado desde un punto de vista experimental para un conjunto de matrices de distintos órdenes obtenidas a partir de la discretización de la ecuación de Laplace. Los resultados obtenidos son análogos a los del capítulo anterior, destacando la mejora obtenida respecto al tiempo de CPU de los modelos caóticos y en particular de los caóticos asíncronos.



Universitat d'Alacant
Universidad de Alicante

Capítulo 5

Método de multipartición basado en el método AOR

5.1 Introducción

En el Capítulo 4 se ha planteado un modelo caótico, a partir del siguiente esquema iterativo

$$\mathbf{x}^{(l+1)} = \omega \sum_{j=1}^r E_j F_j^{-1}(G_j \mathbf{x}^{(l)} + \mathbf{b}) + (1 - \omega) \mathbf{x}^{(l)}. \quad (5.1)$$

El esquema iterativo (5.1), está basado en la técnica de multipartición, y el parámetro de relajación ω se usa de forma análoga a como se utiliza en el método de Jacobi relajado. La convergencia de dicho esquema, como ya se ha comentado en el Capítulo 4, fue estudiada en 1989 por Frommer y Mayer [29] para el caso

de H -matrices. Sin embargo, si todas las matrices F_j , $j = 1, 2, \dots, r$, tienen una forma triangular inferior particular se presenta otra manera de introducir el factor de relajación ω . Para plantear esta variante relajada de multipartición, es conveniente utilizar una notación diferente.

Supongamos que la matriz A está dividida de la forma

$$A = D - L_j - V_j, \quad j = 1, 2, \dots, r, \quad (5.2)$$

donde D es una matriz diagonal, cuyos elementos diagonales coinciden con los de A , L_j , $j = 1, 2, \dots, r$, son matrices estrictamente triangulares inferiores, y V_j , $j = 1, 2, \dots, r$, están construidas de forma que satisfacen la expresión (5.2). Nótese que en general, las matrices V_j , $j = 1, 2, \dots, r$, no son triangulares superiores.

Supongamos que la matriz D es invertible y que los elementos de A , D , L_j y V_j se han denotado respectivamente por a_{ik} , d_{ik} , l_{ik} y v_{ik}^j . A partir de la multipartición de A

$$\{D - L_j, V_j, E_j\}_{j=1}^r, \quad (5.3)$$

el esquema que plantean Frommer y Mayer es el siguiente.

$$x^{(l+1)} = \sum_{j=1}^r E_j y^{jl}, \quad l = 0, 1, \dots,$$

donde los resultados intermedios y^{jl} se calculan resolviendo los siguientes sistemas triangulares

$$(D - L_j) y^{jl} = V_j x^{(l)} + b, \quad j = 1, 2, \dots, r, \quad l = 0, 1, \dots$$

Podemos usar un parámetro de relajación positivo de la misma forma que en el método de Gauss-Seidel, es decir, podemos realizar la iteración

$$x^{(l+1)} = \sum_{j=1}^r E_j y^{jl}, \quad l = 0, 1, \dots, \quad (5.4)$$

donde para $j = 1, 2, \dots, r$, las componentes de $\mathbf{y}^{(j)}$ son calculadas sucesivamente por

$$y_i^{(j)} = \frac{\omega}{\omega_j} \left(\sum_{k=1}^{i-1} p_{ik} y_k^{(j)} + \sum_{k=1}^n v_{ik} x_k^{(l)} + b_i \right) + (1 - \omega) x_i^{(l)}, \quad i = 1, 2, \dots, n.$$

Claramente los cálculos de $\mathbf{y}^{(j)}$ para $j = 1, 2, \dots, r$, son independientes y se pueden realizar en paralelo.

En este caso el esquema iterativo (5.4) se puede escribir como

$$\mathbf{x}^{(l+1)} = H \mathbf{x}^{(l)} + c, \quad l = 0, 1, \dots,$$

con

$$H = \sum_{j=1}^r E_j (D - \omega L_j)^{-1} ((1 - \omega)D + \omega V_j).$$

La convergencia del esquema (5.4) se estudia en [29], como se vio en la Sección 2.2, para H -matrices, probando la convergencia para los mismos parámetros de relajación ω que para el esquema iterativo (5.1) (ver Sección 2.2), y asumiendo que $(A) = |D| - |L_j| - |V_j|$, $j = 1, 2, \dots, r$.

Deren en [18] presenta una clase de algoritmos relajados basados en multiparticiones, llamados algoritmos de multipartición en paralelo AOR (acceleration overrelaxation), para resolver grandes sistemas de ecuaciones.

A partir de la multipartición que aparece de la expresión (5.3), se construye el siguiente esquema iterativo

$$\mathbf{x}^{(l+1)} = \sum_{j=1}^r E_j P_j \mathbf{x}^{(l)}, \quad l = 0, 1, \dots, \quad (5.5)$$

donde para cada $j = 1, 2, \dots, r$, los operadores $P_j : \mathbb{R}^n \rightarrow \mathbb{R}^n$ están definidos como sigue

$$P_j \mathbf{x} = (D - \mu L_j)^{-1} \{ [(1 - \omega)D + (\omega - \mu)L_j + \omega V_j] \mathbf{x} + \omega b \}, \quad (5.6)$$

siendo $\mu \geq 0$ y $\omega > 0$ de \mathbb{R} .

El esquema iterativo (5.5) se puede escribir como

$$z^{(l+1)} = Hz^{(l)} + c, \quad l = 0, 1, \dots,$$

con

$$H = \sum_{j=1}^r E_j (D - \mu L_j)^{-1} [(1 - \omega)D + (\omega - \mu)L_j + \omega V_j]$$

y

$$c = \sum_{j=1}^r E_j (D - \mu L_j)^{-1} \omega b.$$

Obviamente, si $r = 1$, L_1 es la parte estrictamente triangular inferior de A y V_1 la parte estrictamente triangular superior, entonces el esquema iterativo (5.5) se reduce al conocido método clásico *AOR* (ver Sección 1.3). Es por esto que el esquema iterativo (5.5) se denomina método de multipartición paralelo *AOR*. Además, se observa que para valores específicos de los parámetros μ y ω este método se reduce a otros bien conocidos.

- $\mu = 0, \omega = 1$: Método de multipartición paralelo Jacobi.
- $\mu = 1, \omega = 1$: Método de multipartición paralelo Gauss-Seidel.
- $\mu = 0, \omega > 0$: Método de multipartición paralelo Jacobi relajado.

Destacamos además, que el esquema iterativo (5.5) es una generalización del esquema (5.4), ya que eligiendo $\omega = \mu$ en el esquema iterativo (5.5) se obtiene el esquema iterativo (5.4).

Teniendo en cuenta que los cálculos implicados por los operadores P_j en (5.6) pueden ser distintos atendiendo a la partición de la matriz A utilizada, si suponemos que se dispone de r procesadores puede ocurrir que algunos procesadores

hayan finalizado sus cálculos, mientras que otros continúan trabajando. Parece entonces apropiado que los procesadores puedan realizar un número de iteraciones locales variable hasta llegar a una fase de sincronización donde todos los procesadores actualizan la aproximación global a la solución en ese instante. De esta forma se intenta equilibrar la cantidad de trabajo asignada a cada procesador. Esto sugiere el planteamiento de un modelo caótico basado en el esquema iterativo (5.5). En la Sección 5.2 plantearemos dicho esquema y estudiaremos su convergencia cuando la matriz del sistema $Az = b$ es una H -matriz. La Sección 5.3 está dedicada al estudio y planteamiento de la versión asíncrona. Además se presentan los resultados numéricos más significativos obtenidos para estos nuevos modelos teóricos en la Sección 5.4 y por último, en la Sección 5.5 enumeramos las conclusiones más importantes que se obtienen en este capítulo. Previamente introducimos una serie de resultados, que nos serán de utilidad en las restantes secciones de este capítulo.

La demostración del siguiente lema puede verse en Ortega [62].

Lema 1. Sean A y B dos matrices cuadradas reales tales que $|A| \leq B$. Entonces $\rho(A) \leq \rho(B)$.

El siguiente teorema nos da una caracterización de las matrices monótonas. Su demostración puede verse en Varga [78].

Teorema 1. Si $A = [a_{ij}]_{1 \leq i, j \leq n}$, es una matriz real cuadrada de orden n con $a_{ij} \leq 0$ para todo $i \neq j$, entonces las siguientes condiciones son equivalentes.

(i) A es invertible, y $A^{-1} \geq 0$.

(ii) Los elementos diagonales de A son positivos y si $D = \text{diag}(A)$, entonces la matriz $I - D^{-1}A$ es no negativa y convergente.

Para finalizar esta sección introducimos un teorema que caracteriza las M -matrices. Su demostración puede verse en Young [83].

Teorema 2. *Si A es una matriz cuyos elementos diagonales son positivos y los demás no positivos, entonces A es una M matriz, si y sólo si*

$$\rho(B) < 1,$$

donde

$$B = D^{-1}C, \quad D = \text{diag}(A), \quad A = D - C.$$

5.2 Modelo síncrono.

Consideremos el sistema de ecuaciones lineales

$$Ax = b, \quad (5.7)$$

donde A es una matriz cuadrada de orden n invertible.

Supongamos que la matriz A está dividida, tal y como aparece en la expresión (5.2), es decir, $A = D - L_j - V_j$, $j = 1, 2, \dots, r$, siendo $D = \text{diag}(A)$ y L_j , $j = 1, 2, \dots, r$, matrices estrictamente triangulares inferiores.

Para cada $j = 1, 2, \dots, r$, definimos los operadores P_j como sigue

$$P_j: \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$P_j x = J_j x + \omega(D - \mu L_j)^{-1} b, \quad j = 1, 2, \dots, r, \quad (5.8)$$

donde la matriz J_j es

$$J_j = (D - \mu L_j)^{-1} [(1 - \omega)D + (\omega - \mu)L_j + \omega V_j], \quad j = 1, 2, \dots, r,$$

y $\mu > 0$ y $\omega > 0$ son reales.

Si E_j , $j = 1, 2, \dots, r$, son matrices diagonales no negativas cuya suma es la matriz identidad construimos el siguiente esquema caótico

$$\mathbf{x}^{(l+1)} = \sum_{j=1}^r E_j P_j^{q(l,j)} \mathbf{x}^{(l)}, \quad l = 0, 1, \dots \quad (5.9)$$

Los factores $q(l, j)$, denominados factores caóticos, indican el número de veces que el procesador j -ésimo realiza los cálculos indicados por los operadores P_j , definidos en (5.8), en la iteración l . Una vez que el procesador j -ésimo ha finalizado estos cálculos, se premultiplica por E_j el resultado. La actualización global se produce cuando todos estos cálculos han sido completados por todos los procesadores.

Para analizar la convergencia del esquema iterativo (5.9) escribimos dicho esquema de una forma alternativa. Nos basamos en la forma en que se han construido los operadores P_j definidos en la expresión (5.8).

$$\begin{aligned} \mathbf{x}^{(l+1)} &= \sum_{j=1}^r E_j P_j^{q(l,j)} \mathbf{x}^{(l)} \\ &= \sum_{j=1}^r E_j P_j^{q(l,j)-1} [J_j \mathbf{x}^{(l)} + \omega(D - \mu L_j)^{-1} \mathbf{b}] \\ &= \sum_{j=1}^r E_j P_j^{q(l,j)-2} [J_j (J_j \mathbf{x}^{(l)} + \omega(D - \mu L_j)^{-1} \mathbf{b}) + \omega(D - \mu L_j)^{-1} \mathbf{b}] \\ &= \sum_{j=1}^r E_j P_j^{q(l,j)-2} [J_j^2 \mathbf{x}^{(l)} + (J_j + I)\omega(D - \mu L_j)^{-1} \mathbf{b}] \end{aligned}$$

$$\dots = \sum_{j=1}^r E_j \left(J_j^{q^{(l,j)}} \mathbf{x}^{(l)} + \left(\sum_{i=0}^{q^{(l,j)}-1} J_j^i \right) \omega(D - \mu L_j)^{-1} \mathbf{b} \right). \quad (5.10)$$

A partir de la expresión (5.10), podemos escribir el esquema (5.9) como sigue

$$\mathbf{x}^{(l+1)} = H^{(l)} \mathbf{x}^{(l)} + \mathbf{r}^{(l)}, \quad l = 0, 1, \dots, \quad (5.11)$$

donde

$$H^{(l)} = \sum_{j=1}^r E_j J_j^{q^{(l,j)}}, \quad l = 0, 1, \dots, \quad (5.12)$$

y

$$\mathbf{r}^{(l)} = \sum_{j=1}^r E_j \left(\sum_{i=0}^{q^{(l,j)}-1} J_j^i \right) \omega(D - \mu L_j)^{-1} \mathbf{b}, \quad l = 0, 1, \dots \quad (5.13)$$

A las matrices $H^{(l)}$, $l = 0, 1, \dots$, las denominamos *matrices de iteración* del esquema (5.9).

Para probar la convergencia del esquema (5.9), análogamente a como se hizo en los capítulos anteriores vamos a analizar la expresión del error

$$\mathbf{e}^{(l+1)} = \mathbf{x}^{(l+1)} - \boldsymbol{\xi}, \quad l = 0, 1, \dots,$$

siendo $\boldsymbol{\xi}$ la solución exacta del sistema lineal (5.7). El esquema iterativo (5.9) es convergente si la sucesión de vectores error $\{\mathbf{e}^{(l)}\}_{l=0}^{\infty}$ converge al vector cero.

En el siguiente lema demostramos que la solución del sistema lineal (5.7) es un punto fijo del esquema iterativo (5.9). A partir de este resultado se expresarán los vectores error $\mathbf{e}^{(l+1)}$, $l = 0, 1, \dots$, en función de las matrices de iteración $H^{(l)}$.

Lema 2. *La solución del sistema lineal (5.7), $\boldsymbol{\xi}$, es un punto fijo del esquema iterativo (5.9), es decir*

$$\boldsymbol{\xi} = H^{(l)} \boldsymbol{\xi} + \mathbf{r}^{(l)}, \quad l = 0, 1, \dots, \quad (5.14)$$

con $H^{(i)}$ y $r^{(i)}$ definidas en las expresiones (5.12) y (5.13) respectivamente.

Demostración. Teniendo en cuenta las particiones $A = D - L_j - V_j$, podemos escribir para cada $j = 1, 2, \dots, r$

$$\begin{aligned} A &= D - L_j - V_j \\ &= D - \mu L_j + \mu L_j - L_j - V_j \\ &= D - \mu L_j - [(1 - \mu)L_j + V_j]. \end{aligned} \quad (5.15)$$

Ahora bien, si denotamos

$$M_j = D - \mu L_j, \quad \text{y} \quad N_j = (1 - \mu)L_j + V_j, \quad (5.16)$$

a partir de la expresión (5.15) se tiene

$$A = M_j - N_j, \quad j = 1, 2, \dots, r. \quad (5.17)$$

Teniendo en cuenta la expresión (5.17), y que ξ es la solución del sistema lineal (5.7) se tiene

$$\begin{aligned} A\xi &= b \\ \omega(M_j - N_j)\xi &= \omega b, \quad j = 1, 2, \dots, r. \end{aligned} \quad (5.18)$$

Sumando $M_j\xi$, en ambos miembros de la expresión (5.18), para cada $j = 1, 2, \dots, r$, obtenemos

$$\omega(M_j - N_j)\xi + M_j\xi = \omega b + M_j\xi.$$

Despejando se tiene

$$M_j\xi = (1 - \omega)M_j\xi + \omega N_j\xi + \omega b, \quad j = 1, 2, \dots, r. \quad (5.19)$$

De la expresión (5.19) se sigue

$$\xi = M_j^{-1} [(1 - \omega)M_j\xi + \omega N_j\xi + \omega b], \quad j = 1, 2, \dots, r. \quad (5.20)$$

Sustituyendo en (5.20) las matrices M_j y N_j por su expresión en (5.16) se obtiene

$$\begin{aligned} \xi &= (D - \mu L_j)^{-1} [(1 - \omega)(D - \mu L_j)\xi + \omega(1 - \mu)L_j\xi + \omega V_j\xi + \omega b] \\ &= (D - \mu L_j)^{-1} \{[(1 - \omega)D + (\omega - \mu)L_j + \omega V_j]\xi + \omega b\}, \quad (5.21) \\ & \quad j = 1, 2, \dots, r. \end{aligned}$$

De la expresión (5.21) se sigue que ξ es un punto fijo de los operadores P_j , definidos en (5.8) y por tanto del esquema iterativo dado por la expresión (5.9).

■

Observación 1. De las expresiones (5.11) y (5.14) se tiene

$$\begin{aligned} e^{(l+1)} &= x^{(l+1)} - \xi \\ &= H^{(l)}x^{(l)} + r^{(l)} - (H^{(l)}\xi + r^{(l)}) \\ &= H^{(l)}(x^{(l)} - \xi) \\ &= H^{(l)}e^{(l)}, \quad l = 0, 1, \dots, \end{aligned} \quad (5.22)$$

con $H^{(l)}$ definida en (5.12).

Mediante la expresión (5.22) podemos escribir el vector error $e^{(l)}$, para cada iteración l , $l = 0, 1, \dots$, en función del vector error en la iteración inicial, denotado por $e^{(0)}$ de la siguiente forma

$$e^{(l+1)} = H^{(l)}e^{(l)} = \dots = \prod_{i=0}^l H^{(i)}e^{(0)}, \quad l = 0, 1, \dots$$

Por tanto, a partir de esta expresión, la sucesión de vectores error $\{e^{(l)}\}_{l=0}^{\infty}$ converge al vector 0 , si y sólo si

$$\lim_{l \rightarrow \infty} \prod_{i=0}^l H^{(i)} = 0. \quad (5.23)$$

Siguiendo estas consideraciones y a partir del Lema 3 del Capítulo 3, la expresión (5.23) es cierta si existe una constante real α , satisfaciendo

$$\|H^{(l)}\| \leq \alpha < 1, \quad l = 0, 1, \dots, \quad (5.24)$$

para alguna norma matricial compatible $\|\cdot\|$.

En consecuencia el esquema (5.9) es convergente, si se satisface la expresión (5.24).

Una vez construido el esquema caótico síncrono (5.9), vamos a estudiar la convergencia de dicho esquema.

Si la matriz $A = D - L_j - V_j$, $j = 1, 2, \dots, r$, es una H -matriz y $D = \text{diag}(A)$, por el Lema 1 del Capítulo 4 sabemos que $|D|$ es no singular. Si $\langle A \rangle = |D| - |L_j| - |V_j| = |D| - |B|$, denotaremos $\rho = \rho(|D|^{-1}|B|)$. Con esta notación, análoga a la del Capítulo 4, se prueba la convergencia del esquema relajado (5.9). Previamente demostraremos el siguiente lema.

Lema 8. Sea A una H -matriz, y para $j = 1, 2, \dots, r$, sea L_j una matriz estrictamente triángular inferior. Definamos las matrices V_j tal que $A = D - L_j - V_j$, con $D = \text{diag}(A)$. Asumamos que $\langle A \rangle = |D| - |L_j| - |V_j| = |D| - |B|$, para $j = 1, 2, \dots, r$. Entonces, las matrices $D - \mu L_j$, $j = 1, 2, \dots, r$, son H -matrices, para $0 < \mu < \frac{1}{1+\rho}$.

Demostración. Claramente por la construcción de la matriz comparación, los elementos no diagonales de las matrices $(D - \mu L_j)$, $j = 1, 2, \dots, r$, son no positivos. Por tanto será suficiente probar, según la definición de H -matriz, que

$$(D - \mu L_j)^{-1} \geq 0, \quad j = 1, 2, \dots, r. \quad (5.25)$$

Según el Teorema 1, demostrar la expresión (5.25) es equivalente a demostrar que los elementos diagonales de las matrices $(D - \mu L_j)$ son positivos y que las matrices

$$I - |D|^{-1}(D - \mu L_j), \quad j = 1, 2, \dots, r, \quad (5.26)$$

son no negativas y convergentes.

Como $(D - \mu L_j) = |D| - \mu |L_j|$ y $\det(D) \neq 0$, los elementos diagonales de $(D - \mu L_j)$ son positivos. Veamos por tanto, que las matrices que aparecen en la expresión (5.26) son no negativas y convergentes.

$$\begin{aligned} I - |D|^{-1}(D - \mu L_j) &= I - |D|^{-1}(|D| - \mu |L_j|) \\ &= \mu |D|^{-1}|L_j| \geq 0. \end{aligned}$$

Para demostrar que las matrices (5.26) son convergentes, acotaremos dichas matrices de la siguiente forma.

$$\begin{aligned} I - |D|^{-1}(D - \mu L_j) &= \mu |D|^{-1}|L_j| \\ &\leq \mu |D|^{-1}(|L_j| + |V_j|) \\ &= \mu |D|^{-1}|B| \\ &\leq \frac{2}{1 + \rho} |D|^{-1}|B| \end{aligned} \quad (5.27)$$

Teniendo en cuenta el Lema 1 y la acotación (5.27)

$$\rho(I - |D|^{-1}(D - \mu L_j)) \leq \frac{2}{1 + \rho} \rho < 1.$$

Así, las matrices $(D - \mu L_j)$, $j = 1, 2, \dots, r$, son H -matrices. ■

El siguiente teorema prueba la convergencia del esquema iterativo (5.9) cuando la matriz del sistema A es una H -matriz y el factor de relajación pertenece al intervalo $(0, \frac{2}{1+\rho})$.

Teorema 3. *Sea A una H -matriz, y para $j = 1, 2, \dots, r$, sea L_j una matriz estrictamente triangular inferior. Definimos las matrices V_j tal que $A = D - L_j - V_j$, con $D = \text{diag}(A)$. Asumimos que $\langle A \rangle = |D| - |L_j| - |V_j| = |D| - |B|$, para $j = 1, 2, \dots, r$. Entonces si $q(l, j) \geq 1$, $j = 1, 2, \dots, r$, $l = 0, 1, \dots$, el esquema iterativo (5.9) converge a la única solución del sistema $Ax = b$, para $0 < \mu \leq \omega \in (0, \frac{2}{1+\rho})$.*

Demostración. Claramente, por la Observación 1, para probar la convergencia del esquema (5.9) es suficiente encontrar una norma matricial compatible $\|\cdot\|$ y una constante real α , satisfaciendo

$$\|H^{(l)}\| \leq \alpha < 1, \quad l = 0, 1, \dots$$

Según la forma de las matrices de iteración $H^{(l)}$ en la expresión (5.12) podemos acotar $|H^{(l)}|$ de la siguiente forma

$$|H^{(l)}| \leq \sum_{j=1}^r E_j \left| (D - \mu L_j)^{-1} [(1 - \omega)D + (\omega - \mu)L_j + \omega V_j] \right|^{q(l, j)}, \quad (5.28)$$

$$l = 0, 1, \dots$$

Ahora bien, si tenemos en cuenta que las matrices $D - \mu L_j$, $j = 1, 2, \dots, r$, son H -matrices, como se ha demostrado en el Lema 3 y utilizando el apartado (a) del Lema 5 del Capítulo 3, podemos acotar para cada $j = 1, 2, \dots, r$

$$\left| (D - \mu L_j)^{-1} ((1 - \omega)D + (\omega - \mu)L_j + \omega V_j) \right|,$$

de la siguiente forma

$$\begin{aligned} & \left| (D - \mu L_j)^{-1} ((1 - \omega)D + (\omega - \mu)L_j + \omega V_j) \right| \\ & \leq \langle D - \mu L_j \rangle^{-1} |(1 - \omega)D + (\omega - \mu)L_j + \omega V_j|. \end{aligned} \quad (5.29)$$

Si denotamos

$$M_j(\mu) = |D| - \mu|L_j| = \langle D - \mu L_j \rangle, \quad j = 1, 2, \dots, r, \quad (5.30)$$

entonces, a partir de la expresión (5.29) podemos escribir la siguiente desigualdad, para $j = 1, 2, \dots, r$

$$\begin{aligned} & \left| (D - \mu L_j)^{-1} ((1 - \omega)D + (\omega - \mu)L_j + \omega V_j) \right| \\ & \leq M_j(\mu)^{-1} |(1 - \omega)D + (\omega - \mu)L_j + \omega V_j|. \end{aligned} \quad (5.31)$$

Distingamos dos casos atendiendo a los valores de μ y ω .

(a) Si $0 < \mu \leq \omega \leq 1$ se tiene, para $j = 1, 2, \dots, r$

$$|(1 - \omega)D + (\omega - \mu)L_j + \omega V_j| \leq (1 - \omega)|D| + (\omega - \mu)|L_j| + \omega|V_j|.$$

Si denotamos

$$N_j^1(\mu, \omega) = (1 - \omega)|D| + (\omega - \mu)|L_j| + \omega|V_j|,$$

a partir de la desigualdad (5.31) se tiene

$$|(D - \mu L_j)^{-1} ((1 - \omega)D + (\omega - \mu)L_j + \omega V_j)| \leq M_j(\mu)^{-1} N_j^1(\mu, \omega),$$

y teniendo en cuenta la expresión (5.28) obtenemos

$$|H^{(l)}| \leq \sum_{j=1}^r E_j (M_j(\mu)^{-1} N_j^1(\mu, \omega))^{q^{(l,j)}}, \quad l = 0, 1, \dots \quad (5.32)$$

Por otra parte podemos escribir

$$\begin{aligned} M_j(\mu) - N_j^1(\mu, \omega) &= (|D| - \mu|L_j|) - [(1 - \omega)|D| + (\omega - \mu)|L_j| + \omega|V_j|] \\ &= \omega|D| - \mu|L_j| - \omega(|L_j| + |V_j|) + \mu|L_j| \\ &= \omega(|D| - |L_j| - |V_j|), \quad j = 1, 2, \dots, r. \end{aligned} \quad (5.33)$$

Por hipótesis $\langle A \rangle = |D| - |L_j| - |V_j|$, por tanto de la expresión (5.33) se sigue

$$M_j(\mu) - N_j^1(\mu, \omega) = \omega \langle A \rangle \quad (5.34)$$

Puesto que, como se demostró en el Lema 3, las matrices $D - \mu L_j$ son H -matrices para todo $j = 1, 2, \dots, r$, las matrices $\langle D - \mu L_j \rangle = |D| - \mu|L_j|$ son M -matrices. De la definición de $M_j(\mu)$ en la expresión (5.30) se sigue que las matrices $M_j(\mu)$, $j = 1, 2, \dots, r$, son M -matrices.

Nótese además que la matriz $\omega \langle A \rangle$ es una M -matriz puesto que por hipótesis A es H -matriz y $\omega > 0$.

Consideremos e un vector positivo, por ejemplo $e = (1, 1, \dots, 1)^t$, y $x = \omega^{-1} \langle A \rangle^{-1} e$. Como $\omega^{-1} \langle A \rangle^{-1} \geq O$, y ninguna fila de $\omega^{-1} \langle A \rangle^{-1}$ tiene todos sus elementos nulos, entonces $x > o$. Claramente $\omega \langle A \rangle x > o$.

Ahora bien, teniendo en cuenta la expresión (5.34) se sigue

$$M_j(\mu)^{-1}(M_j(\mu) - N_j^1(\mu, \nu)) = M_j(\mu)^{-1}\omega(A).$$

De aquí obtenemos

$$I - M_j(\mu)^{-1}N_j^1(\mu, \omega) = M_j(\mu)^{-1}\omega(A),$$

y entonces, multiplicando por el vector \mathbf{x} , teniendo en cuenta que las matrices $M_j(\mu)$, $j = 1, 2, \dots, r$, son M -matrices y por tanto su inversa es no negativa, y que $\omega(A)\mathbf{x} > \mathbf{o}$, se satisface

$$\mathbf{x} - M_j(\mu)^{-1}N_j^1(\mu, \omega)\mathbf{x} = M_j(\mu)^{-1}\omega(A)\mathbf{x} > \mathbf{o}.$$

Por tanto

$$|M_j(\mu)^{-1}N_j^1(\mu, \omega)|\mathbf{x} < \mathbf{x}. \quad (5.35)$$

A partir de la expresión (5.35) podemos asegurar que existe una constante real $0 \leq \alpha_1 < 1$ que para todo $j = 1, 2, \dots, r$, satisface

$$M_j(\mu)^{-1}N_j^1(\mu, \nu) \leq \alpha_1 \mathbf{x} < \mathbf{x}. \quad (5.36)$$

De las expresiones (5.32) y (5.36) se sigue

$$|H^{(l)}|\mathbf{x} \leq \sum_{j=1}^r E_j (M_j(\mu)^{-1}N_j^1(\mu, \omega))^{q(l,j)} \mathbf{x} \leq \alpha_1 \mathbf{x}, \quad l = 0, 1, \dots, \quad (5.37)$$

con $0 \leq \alpha_1 < 1$ y $\mathbf{x} > \mathbf{o}$.

De aquí se sigue que la norma matricial compatible y monótona $\|\cdot\|_{\mathbf{x}}$, introducida en la Sección 1.1 cumple

$$\|H^{(l)}\|_{\mathbf{x}} \leq \alpha_1 < 1, \quad l = 0, 1, \dots$$

En consecuencia, según la Observación 1, el esquema (5.9) es convergente para $0 < \mu \leq \omega \leq 1$.

(b) Si $0 < \mu \leq \omega$, $1 < \omega < \frac{2}{1+p}$ para $j = 1, 2, \dots, r$, se tiene

$$|(1 - \omega)D + (\omega - \mu)L_j + \omega V_j| \leq (\omega - 1)|D| + (\omega - \mu)|L_j| + \omega|V_j|.$$

Si denotamos

$$N_j^2(\mu, \omega) = (\omega - 1)|D| + (\omega - \mu)|L_j| + \omega|V_j|,$$

a partir de la desigualdad (5.31) y la definición de $M_j(\mu)$ en (5.30) se obtiene para todo $j = 1, 2, \dots, r$

$$|(D - \mu L_j)^{-1} [(1 - \omega)D + (\omega - \mu)L_j + \omega V_j]| \leq M_j(\mu)^{-1} N_j^2(\mu, \omega),$$

y teniendo en cuenta la expresión (5.28) se tiene

$$|H^{(l)}| \leq \sum_{j=1}^r E_j (M_j(\mu)^{-1} N_j^2(\mu, \omega))^{q^{(l,j)}}. \quad (5.38)$$

Por otra parte se puede escribir

$$\begin{aligned} M_j(\mu) - N_j^2(\mu, \omega) &= (|D| - \mu|L_j|) - [(\omega - 1)|D| + (\omega - \mu)|L_j| + \omega|V_j|] \\ &= (2 - \omega)|D| - \omega(|L_j| + |V_j|), \quad j = 1, 2, \dots, r. \end{aligned} \quad (5.39)$$

Por hipótesis se ha denotado $|B| = |L_j| + |V_j|$, por tanto, de la expresión (5.39) se sigue

$$M_j(\mu) - N_j^2(\mu, \omega) = (2 - \omega)|D| - \omega|B|, \quad j = 1, 2, \dots, r. \quad (5.40)$$

Por otra parte

$$\rho\left(\left((2-\omega)|D|\right)^{-1}\omega|B|\right) = \frac{\omega}{2-\omega}\rho(|D|^{-1}|B|),$$

y como $\omega < \frac{2}{1+\rho}$, con $\rho = \rho(|D|^{-1}|B|)$, obtenemos

$$\begin{aligned} \rho\left(\left((2-\omega)|D|\right)^{-1}\omega|B|\right) &= \frac{\omega}{2-\omega}\rho \\ &< \left(\frac{\frac{2}{1+\rho}}{2-\frac{2}{1+\rho}}\right)\rho \\ &= 1. \end{aligned} \tag{5.41}$$

La matriz $(2-\omega)|D| - \omega|B|$ tiene por construcción sus elementos diagonales positivos y los demás no positivos. Entonces teniendo en cuenta la expresión (5.41) y aplicando el Teorema 2 se sigue que $(2-\omega)|D| - \omega|B|$ es una M -matriz.

Consideremos e un vector positivo, por ejemplo $e = (1, 1, \dots, 1)^t$, y el vector $y = \left((2-\omega)|D| - \omega|B|\right)^{-1}e$. Como $\left((2-\omega)|D| - \omega|B|\right)^{-1} \geq O$, y ninguna de sus filas tiene todos sus elementos nulos, entonces $y > o$. Claramente $\left((2-\omega)|D| - \omega|B|\right)y > o$.

Ahora bien, teniendo en cuenta la expresión (5.40) se tiene

$$I - M_j^{-1}(\mu)N_j^2(\mu, \omega) = M_j(\mu)^{-1}\left((2-\omega)|D| - \omega|B|\right).$$

De aquí se sigue, multiplicando por el vector y

$$y - M_j(\mu)^{-1}N_j^2(\mu, \omega)y = M_j(\mu)^{-1}\left((2-\omega)|D| - \omega|B|\right)y > o.$$

Por tanto

$$|M_j(\mu)^{-1}N_j^2(\mu, \omega)|y < y. \tag{5.42}$$

Por la expresión (5.42) podemos asegurar que existe una constante $0 \leq \alpha_2 < 1$ que para todo $j = 1, 2, \dots, r$ satisface

$$M_j(\mu)^{-1} N_j^2(\mu, \omega) \leq \alpha_2 \mathbf{y} < \mathbf{y}. \quad (5.43)$$

De las expresiones (5.38) y (5.43) se sigue

$$|H^{(l)}| \mathbf{y} \leq \sum_{j=1}^r E_j \left(M_j(\mu)^{-1} N_j^2(\mu, \omega) \right)^{q^{(l,j)}} \mathbf{y} \leq \alpha_2 \mathbf{y}, \quad l = 0, 1, \dots, \quad (5.44)$$

con $0 \leq \alpha_2 < 1$ y $\mathbf{y} > \mathbf{o}$.

Por tanto la norma matricial compatible y monótona $\|\cdot\|_{\mathbf{x}}$, introducida en la Sección 1.1 cumple

$$\|H^{(l)}\|_{\mathbf{x}} \leq \alpha_2 < 1, \quad l = 0, 1, \dots$$

En consecuencia, según la Observación 1, obtenemos de forma análoga al apartado (a), que el esquema (5.9) es convergente para $0 < \mu \leq \omega$, y $1 < \omega < \frac{2}{1+\rho}$.

Queda así demostrada la convergencia del esquema (5.9). ■

Observación 2. Hacemos notar que cuando $\omega = \mu$ el esquema iterativo (5.9) corresponde a la versión caótica del esquema iterativo (5.4) dado por Frommer y Mayer en [29].

5.3 Modelo asíncrono

En esta sección presentamos el estudio del modelo asíncrono correspondiente al esquema iterativo caótico (5.9).

Sean E_i , $i = 1, 2, \dots, r$, matrices diagonales no negativas cuya suma es la matriz identidad. Si suponemos que $\mathcal{F}_0 = \{i_j\}_{j=1}^{\infty}$, $1 \leq i_j \leq r$, es una sucesión regulada, para un vector inicial, a partir de dicha sucesión podemos plantear el siguiente modelo asíncrono basado en el esquema caótico (5.9)

$$x^{(j+r_i)} = (I - E_{i_j})x^{(j+r_i-1)} + E_{i_j} \sum_{i=1}^r E_i P_i^{(j,i)} x^{(j)}, \quad j = 0, 1, 2, \dots \quad (5.45)$$

Los operadores P_i , $i = 1, 2, \dots, r$, están definidos en (5.8) de la siguiente forma

$$P_i x = J_i x + \omega(D - \mu L_i)^{-1} b, \quad i = 1, 2, \dots, r,$$

siendo

$$J_i = (D - \mu L_i)^{-1} [(1 - \omega)D + (\omega - \mu)L_i + \omega V_i], \quad i = 1, 2, \dots, r,$$

donde $\mu > 0$ y $\omega > 0$. Además r_j el menor entero positivo tal que

$$i_j = i_{j+r_j}.$$

Por la propia definición, como $\{i_j\}_{j=1}^{\infty}$ es regulada debe existir un entero positivo K que verifique

$$0 \leq r_j - 1 < K.$$

Como ya se ha explicado en los Capítulos 3 y 4, $r_j - 1$ indica el número de veces que la aproximación que reside en un procesador, que denominamos *central*, es actualizada por otros procesadores distintos de i_j -ésimo, durante el intervalo de tiempo en el que dicho procesador calcula su iteración local.

A continuación estudiamos la convergencia de este modelo asíncrono bajo las mismas hipótesis que en el caso síncrono.

Teorema 4. Sea A una H -matriz, y para $i = 1, 2, \dots, r$, sea L_i una matriz estrictamente triangular inferior. Definimos las matrices V_i tal que $A = D - L_i - V_i$ con $D = \text{diag}(A)$. Asumimos que $\langle A \rangle = |D| - |L_i| - |V_i| = |D| - |B|$ para $i = 1, \dots, r$. Entonces si $q(j, i) \geq 1$, $i = 1, 2, \dots, r$, $j = 0, 1, \dots$, dada una sucesión regulada de enteros $\{i_j\}_{j=0}^{\infty}$, $1 \leq i_j \leq r$, el esquema iterativo (5.45) converge a la única solución del sistema $Ax = b$ para $0 < \mu \leq \omega \in (0, \frac{2}{1+\rho})$.

Demostración. Análogamente a como se hizo en los Capítulos 3 y 4 sumergimos el esquema iterativo (5.45) en un procedimiento iterativo sobre \mathbb{R}^{nK} .

Utilizaremos la siguiente notación. Si

$$\epsilon^{(j)} = x^{(j)} - \xi,$$

Llamaremos

$$\bar{\epsilon}_j = \begin{bmatrix} \epsilon^{(j)} \\ \epsilon^{(j-1)} \\ \vdots \\ \epsilon^{(j-K+1)} \end{bmatrix} \in \mathbb{R}^{nK},$$

Con esta notación, podemos escribir

$$\epsilon^{(j)} = S_j \bar{\epsilon}_{j+r_j-1},$$

donde S_j es una matriz de tamaño $n \times nK$, con un bloque identidad de orden n en la posición r_j , y todos sus demás elementos son nulos. Como en el Capítulo 3 se comprueba fácilmente que

$$\begin{aligned} \epsilon^{(j+r_j)} &= x^{(j+r_j)} - \xi \\ &= (I - E_{i_j}) \epsilon^{(j+r_j-1)} + E_{i_j} H^{(j)} S_j \bar{\epsilon}_{j+r_j-1}. \end{aligned}$$

Por tanto, podemos escribir

$$\bar{\epsilon}_{j+r} = B_{j+r} \bar{\epsilon}_{j+r-1},$$

donde B_{j+r} , es la siguiente matriz de orden nK

$$B_{j+r} = \begin{bmatrix} I - E_i & O & \dots & O & O \\ I & O & \dots & O & O \\ \vdots & \vdots & & \vdots & \vdots \\ O & O & \dots & I & O \end{bmatrix} + \begin{bmatrix} E_i, H^{(i)} S_j \\ O \\ \vdots \\ O \end{bmatrix}.$$

Entonces

$$\bar{\epsilon}_{j+2K-1} = B_{j+2K-1} B_{j+2K-2} \dots B_{j+1} \bar{\epsilon}_j. \quad (5.46)$$

Distinguimos dos casos.

(a) Supongamos $0 < \mu \leq \omega \leq 1$. Sea el vector

$$\bar{x} = (x^1, x^2, \dots, x^t)^t,$$

donde $x = \omega^{-1}(A)^{-1}(1, 1, \dots, 1)^t$. Según el Teorema 3, el vector x es positivo y satisface la desigualdad

$$|H^{(l)}|x \leq \alpha_1 x, \quad 0 \leq \alpha_1 < 1, \quad l = 0, 1, 2, \dots$$

Con estas condiciones se prueba, razonando como en el Teorema 7 del Capítulo 3, que

$$|B_{j+2K-1} B_{j+2K-2} \dots B_{j+1}| \bar{x} \leq \alpha_1 \bar{x}.$$

(b) Supongamos $0 < \mu \leq \omega$, $1 < \omega < \frac{2}{1+\rho}$. Consideremos el vector

$$\bar{y} = (y^1, y^2, \dots, y^t)^t,$$

donde $\mathbf{y} = ((2 - \omega)|D| - \omega|B|)^{-1}(1, 1, \dots, 1)^t$. Del Teorema 3, se sigue que el vector \mathbf{y} es positivo y que satisface la siguiente desigualdad

$$|H^{(l)}|\mathbf{y} \leq \alpha_2 \mathbf{y}, \quad 0 \leq \alpha_2 < 1, \quad l = 0, 1, 2, \dots$$

Por tanto, razonando como en el Teorema 7 del Capítulo 3, obtenemos que

$$|B_{j+2K-1}B_{j+2K-2} \cdots B_{j+1}|\bar{\mathbf{y}} \leq \alpha_2 \bar{\mathbf{y}}.$$

De los dos apartados anteriores se sigue que $\lim_{\nu \rightarrow \infty} \bar{\mathbf{e}}_\nu = \mathbf{0}$. ■

5.4 Experimentos numéricos

En esta sección estudiamos experimentalmente los modelos caóticos paralelos síncrono y asíncrono estudiados en este capítulo. El planteamiento de estos esquemas queda reflejado en las expresiones (5.9) y (5.45) respectivamente.

Análogamente a como se ha hecho en los Capítulos 3 y 4, estudiamos el problema de resolver el sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$, que se obtiene de la discretización de la ecuación de Laplace. Las matrices A que utilizaremos son las mismas que las utilizadas en el capítulo anterior y su forma general viene definida en la expresión (3.50).

Las implementaciones de los algoritmos caóticos tanto síncrono como asíncrono estudiados en este capítulo, están detalladas en la Subsección 5.4.1. En la Subsección 5.4.2 analizamos los resultados más relevantes que se han obtenido.

5.4.1 Algoritmos

Se desea resolver el sistema lineal $Ax = b$. Supongamos r particiones de la matriz A

$$A = D - L_j - V_j \quad j = 1, 2, \dots, r, \quad (5.47)$$

donde $D = \text{diag}(A)$, L_j , $j = 1, 2, \dots, r$, son matrices triangulares inferiores y las matrices V_j , $j = 1, 2, \dots, r$, satisfacen (5.47).

Antes de introducir los algoritmos correspondientes a los modelos estudiados en este capítulo hacemos notar que aplicar el operador P_j definido en (5.8), sobre un vector x , equivale a realizar los siguientes cálculos.

(i) Resolver el sistema

$$(D - \mu L_j)y = ((1 - \mu)D + \mu V_j)x + \mu b,$$

(ii) Obtenemos ahora $P_j x$, a partir de los vectores y y x

$$P_j x = \frac{\omega}{\mu} y + \left(1 - \frac{\omega}{\mu}\right) x.$$

Con una sencilla sustitución se obtiene

$$P_j x = (D - \mu L_j)^{-1} \{[(1 - \omega)D + (\omega - \mu)L_j + \omega V_j]x + \omega b\},$$

que corresponde a la definición de P_j en (5.8).

Los siguientes algoritmos detallan los pasos a seguir para resolver el sistema de ecuaciones lineales $Ax = b$ mediante el esquema caótico síncrono (5.9) y el esquema caótico asíncrono (5.45) respectivamente.

ALGORITMO SÍNCRONO

Etapa 0. Lectura de datos:

Matriz de coeficientes: A
 Término independiente: b
 Factores caóticos: $q(l, j)$, $1 \leq j \leq r$, $l = 0, 1, \dots$
 Factores de relajación: ω , μ

Etapa 1. Elegir $x^{(0)}$ arbitrariamente y $l = 0$.

Etapa 2. (Etapa caótica) En paralelo,

Para $j = 1, 2, \dots, r$

Para $k = 1, 2, \dots, q(l, j)$, con $x_j^{(0)} = x^{(l)}$.

Resolver el sistema

$$(D - \mu L_j)x_j^{(k)} = ((1 - \mu)D + \mu V_j)x_j^{(k-1)} + \mu b,$$

$$x_j^{(k)} = \frac{\omega}{\mu} x_j^{(k)} + \left(1 - \frac{\omega}{\mu}\right) x_j^{(k-1)}.$$

Etapa 3. (Etapa síncrona) $x^{(l+1)} = \sum_{j=1}^r E_j x_j^{(q(l, j))}$, siendo E_j , matrices diagonales no negativas cuya suma es la identidad.

Etapa 4. Comprobación del criterio de convergencia

$$\sum_{k=1}^n |x_k^{(l+1)} - x_k^{(l)}| < 5 \cdot 10^{-10}.$$

- Si test de convergencia = VERDADERO, entonces FIN.
- Si test de convergencia = FALSO, entonces
 - $l \leftarrow l + 1$.
 - Volver a la etapa 2.

ALGORITMO ASÍNCRONO

Etapa 0. Lectura de datos:

Matriz de coeficientes: A Término independiente: b Factores caóticos: $q(l, j)$, $1 \leq j \leq r$, $l = 0, 1, \dots$ Factores de relajación: ω, μ

Etapa 1. Elegir $x^{(0)}$ arbitrariamente y $l_j = 0$, $1 \leq j \leq r$. En paralelo, para $j = 1, 2, \dots, r$ (etapas 2-5).

Etapa 2. $x^{(l_j)} = comm$, donde $comm$ es un vector compartido por todos los procesadores.

Etapa 3. Para $k = 1, 2, \dots, q(l_j, j)$, con $\bar{x}_j^{(0)} = x^{(l_j)}$.

Resolver el sistema

$$(D - \mu L_j)z_j^{(k)} = ((1 - \mu)D + \mu V_j)\bar{x}_j^{(k-1)} + \mu b,$$

$$\bar{x}_j^{(k)} = \frac{\omega}{\mu} z_j^{(k)} + \left(1 - \frac{\omega}{\mu}\right) \bar{x}_j^{(k-1)}.$$

Etapa 4. $comm = (I - E_j)comm + E_j \bar{x}_j^{(q(l_j, j))}$, siendo E_j , matrices diagonales no negativas cuya suma es la identidad.

Etapa 5. Comprobación del criterio de convergencia.

- Si test de convergencia = VERDADERO, entonces FIN.

- Si test de convergencia = FALSO, entonces

- $l_j \leftarrow l_j + 1$.

- Volver a la etapa 2.

La Tabla 1 muestra, para las matrices utilizadas, los resultados obtenidos para diferentes modelos caóticos síncronos ejecutados en secuencial. Observamos en estos esquemas, que para el caso particular de ω_{opt} , el modelo no caótico es mejorado rara vez por los caóticos. Esta situación es análoga cuando dichos modelos se ejecutan en paralelo, como puede observarse en la Tabla 2. En dicha tabla se presentan los resultados obtenidos para los esquemas paralelos caóticos síncronos, utilizando los 8 procesadores de los que dispone el multiprocesador Alliant FX/80. Hemos calculado el tiempo de CPU para la ejecución concurrente y vectorial-concurrente, denotándolos por T_1 y T_2 respectivamente. También se han calculado los dos tipos de speed-up correspondientes a cada uno de estos tiempos, explicados en la Sección 4.4. Es decir,

$$S_i = \frac{T_s}{T_i}, \quad i = 1, 2,$$

donde T_s es el tiempo de CPU del algoritmo secuencial no caótico para su factor de relajación óptimo, y

$$S'_i = \frac{T_{cs}}{T_i}, \quad i = 1, 2,$$

siendo T_{cs} el tiempo de CPU del modelo caótico que se está evaluando, ejecutado en un único procesador.

Orden	$q(l, j), 1 \leq j \leq 8$	ω_{opt}	Iter.	Tiempo (seg.)
384	1,1,1,1,2,2,2,2	1.350	480	5.308
	2,1,2,2,4,4,4,4	1.200	345	6.928
	2,1,2,2,2,2,2,2	1.300	352	4.755
	3,3,3,3,6,6,6,6	1.150	258	8.490
	1,1,1,1,1,1,1,1	1.375	536	4.925
1024	3,3,3,3,3,3,3,3	1.525	539	22.135
	2,2,2,2,2,2,2,2	1.625	632	19.869
	1,1,1,1,1,1,1,1	1.650	933	22.616
2816	2,2,2,4,4,4,4,4	1.150	332	52.385
	1,1,1,2,2,2,2,2	1.250	524	43.157
	2,2,2,2,2,2,2,2	1.225	396	37.958
	1,1,1,1,1,1,1,1	1.350	571	36.670
4096	2,2,4,4,6,6,6,6	1.275	2708	486.459
	1,1,2,2,3,3,3,3	1.500	3308	379.421
	1,1,1,1,1,1,1,1	1.680	2676	249.048
5632	2,2,4,4,4,4,4,4	1.150	344	107.301
	1,1,1,2,2,2,2,2	1.250	542	87.202
	1,2,2,2,2,2,2,2	1.225	410	76.339
	1,1,1,1,1,1,1,1	1.350	591	79.430

Tabla 1: Algoritmo secuencial.

Orden	$q(l, j), 1 \leq j \leq 8$	ω_{opt}	Iter.	Tiempo (seg.)		Speed-up			
				T_1	T_2	S_1	S_2	S'_1	S'_2
384	1,1,1,1,2,2,2,2	1.350	480	1.006	0.433	4.89	11.37	5.27	12.25
	2,2,2,2,4,4,4,4	1.200	345	1.422	0.740	3.46	6.65	4.87	9.36
	2,2,2,2,2,2,2,2	1.300	352	0.846	0.426	5.82	11.56	5.62	11.16
	3,3,3,3,6,6,6,6	1.150	258	1.796	0.987	2.74	4.98	4.72	8.60
	1,1,1,1,1,1,1,1	1.375	536	0.923	0.368	5.33	13.38	5.33	13.38
1024	3,3,3,3,3,3,3,3	1.525	539	3.549	1.843	6.37	12.27	6.24	14.21
	2,2,2,2,2,2,2,2	1.625	632	3.261	1.487	6.93	15.20	6.09	13.36
	1,1,1,1,1,1,1,1	1.650	933	3.894	1.383	5.80	16.35	5.80	16.35
2816	2,2,2,4,4,4,4,4	1.150	332	9.868	4.678	3.71	7.83	5.31	11.20
	1,1,1,2,2,2,2,2	1.250	524	7.692	3.102	4.76	11.82	5.61	13.91
	2,2,2,2,2,2,2,2	1.225	396	6.803	3.268	5.39	11.22	5.58	11.62
	1,1,1,1,1,1,1,1	1.350	571	7.030	2.722	5.22	13.47	5.22	13.47
4096	2,2,4,4,6,6,6,6	1.275	2708	87.737	44.870	2.84	5.55	5.54	10.84
	1,1,2,2,3,3,3,3	1.500	3308	68.700	29.590	3.62	8.41	5.52	12.82
	1,1,1,1,1,1,1,1	1.680	2676	52.510	22.390	4.74	11.12	4.74	22.39
5632	2,2,2,4,4,4,4,4	1.150	344	21.087	11.135	3.76	7.13	5.09	9.64
	1,1,1,2,2,2,2,2	1.250	542	16.451	7.166	4.83	11.06	5.30	12.17
	2,2,2,2,2,2,2,2	1.225	410	14.490	7.422	5.48	10.70	5.27	10.28
	1,1,1,1,1,1,1,1	1.350	591	14.330	5.814	5.54	13.66	5.54	13.66

Tabla 2: Resultados de los esquemas paralelos caóticos síncronos para ω óptimo.

Aunque para el factor óptimo ω_{opt} , los modelos caóticos no presentan en general un mejor tiempo de *CPU* que los no caóticos, esta situación no se puede generalizar a todos los valores del parámetro ω . Así, las Figuras 1, 2, 3, 4 y 5 reflejan que siempre existen modelos caóticos que son mejores que el no caótico, en prácticamente todo su rango de convergencia, es decir, en aquellos valores de ω para los que la sucesión de vectores generada por el método iterativo caótico converge a la solución del sistema.

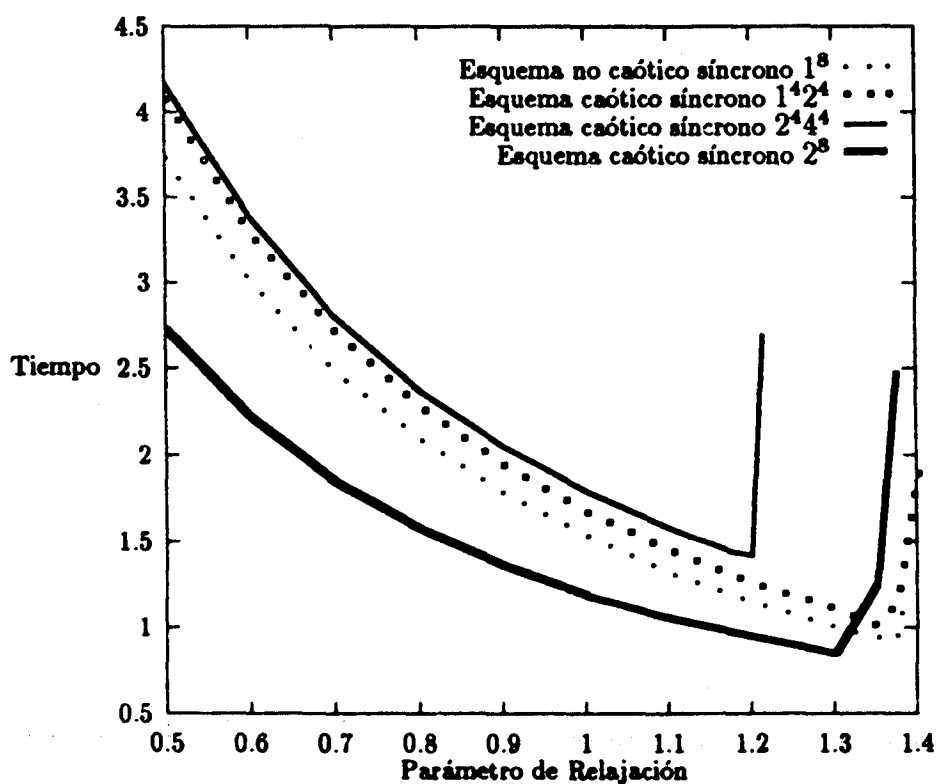


Figura 1: Comparación esquemas paralelos caóticos síncronos, $n = 384$.

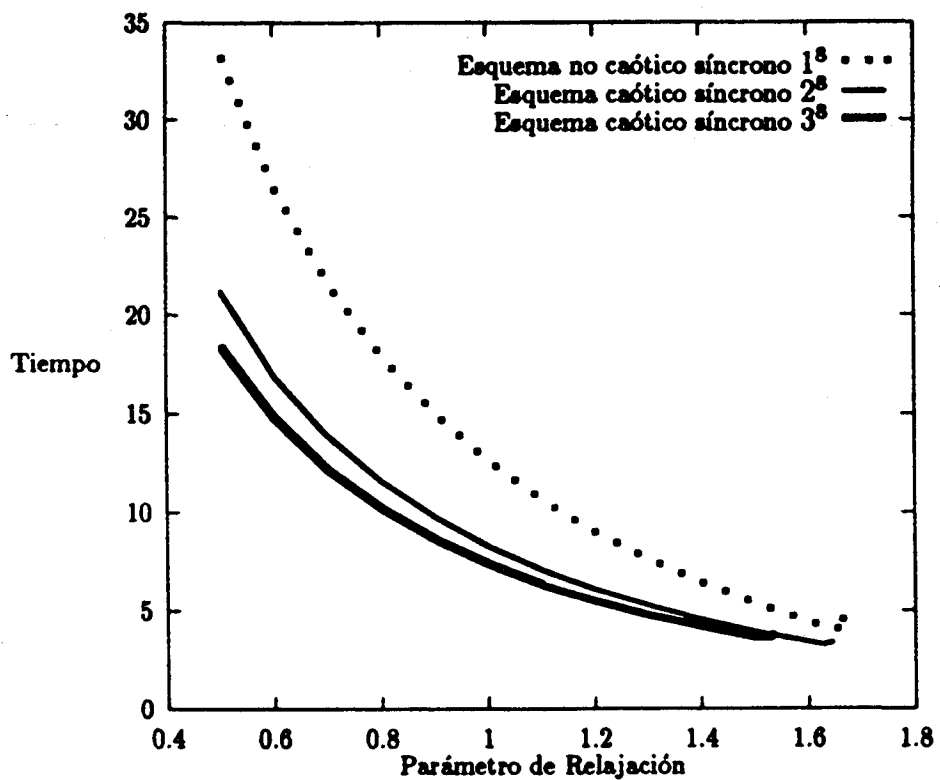


Figura 2: Comparación esquemas paralelos caóticos síncronos, $n = 1024$.

Universitat d'Alacant
 Universidad de Alicante

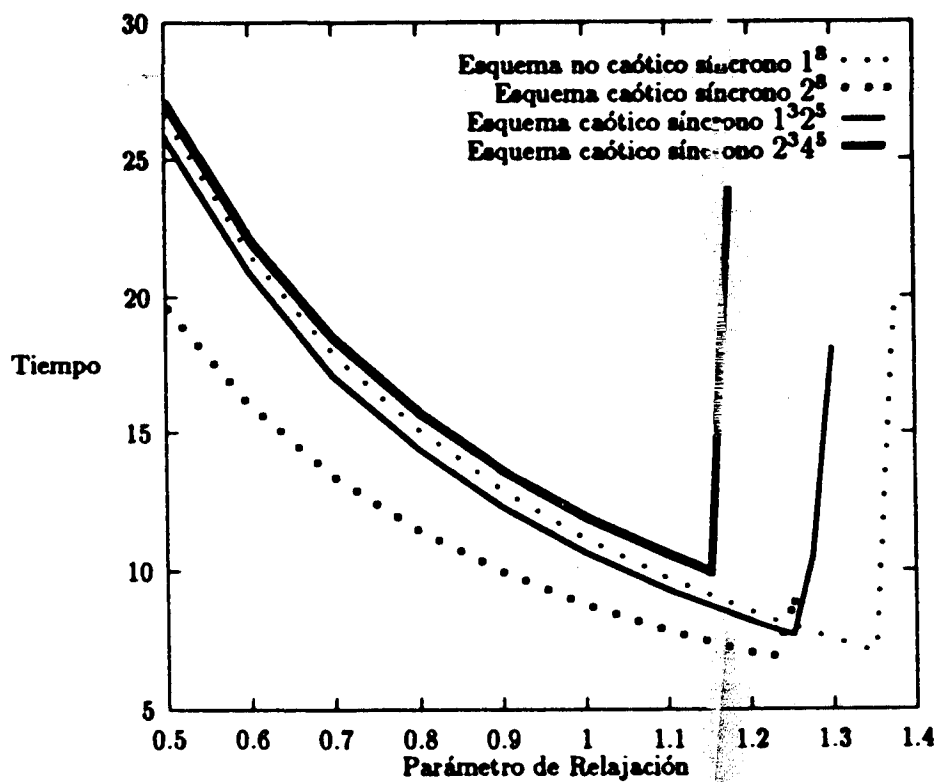


Figura 3: Comparación esquemas paralelos caóticos síncronos, $n = 2816$.

Universitat d'Alacant
Universidad de Alicante

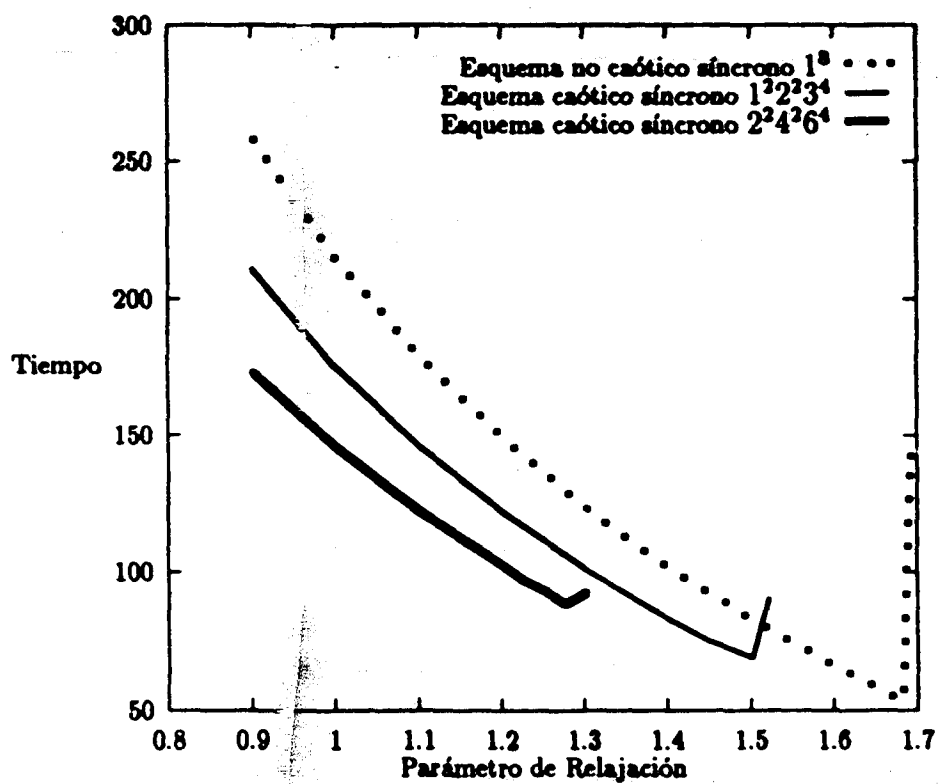


Figura 4: Comparación esquemas paralelos caóticos sincronicos, $n = 4096$.

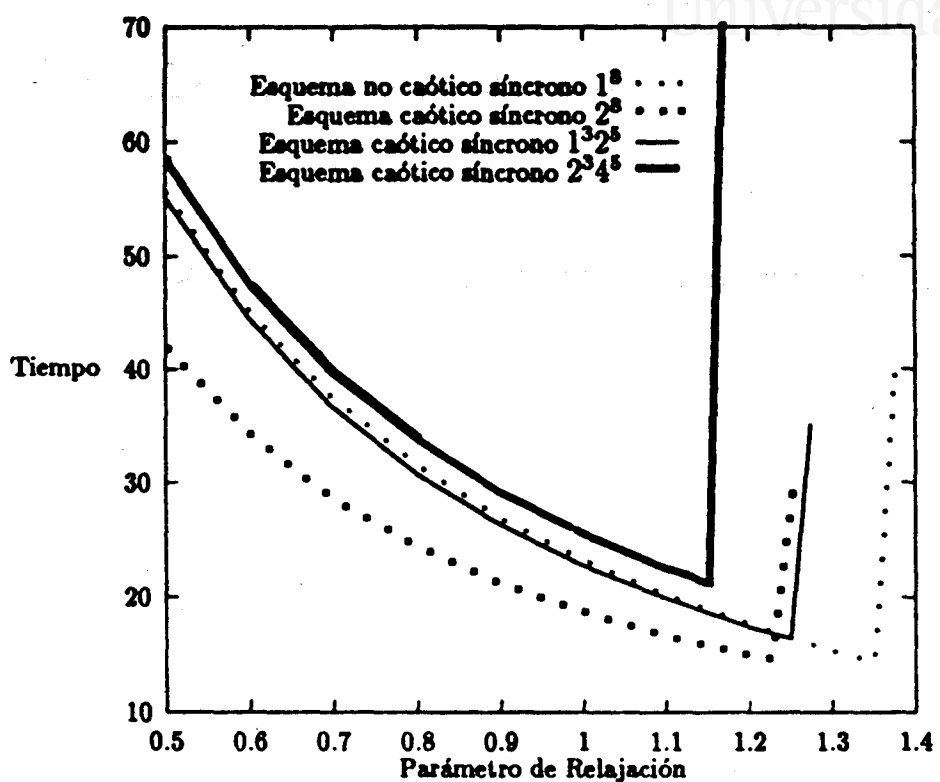


Figura 5: Comparación esquemas paralelos caóticos sincrónicos, $n = 5632$.

Los resultados obtenidos para los modelos paralelos caóticos sincrónicos son mejorados por los correspondientes asíncronos para todo valor del parámetro ω , como se puede observar en las Figuras 6 y 7 en las que se comparan para las matrices de tamaño 4096 y 5632, la versión sincrónica y asíncrona de dos esquemas caóticos. Esta situación ha sido similar para los demás esquemas y matrices.

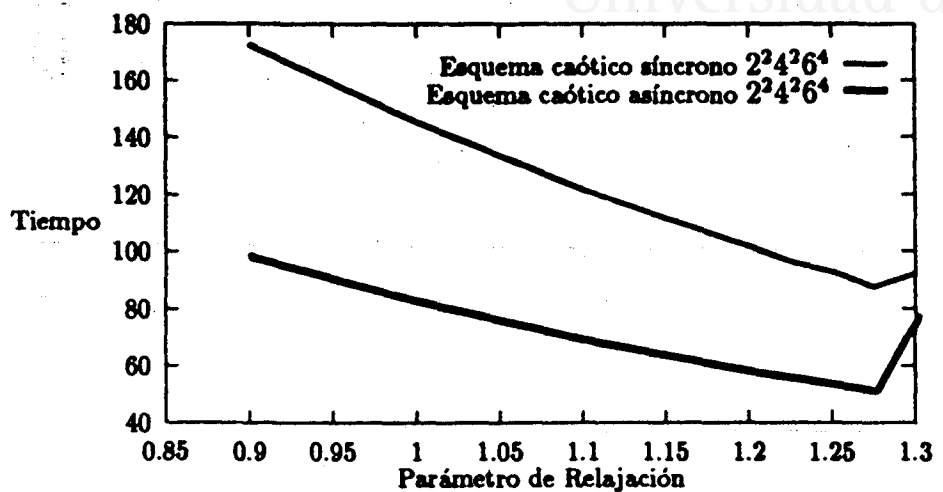


Figura 6: Comparación esquemas paralelos caóticos síncrono y asíncrono, $n = 4096$.

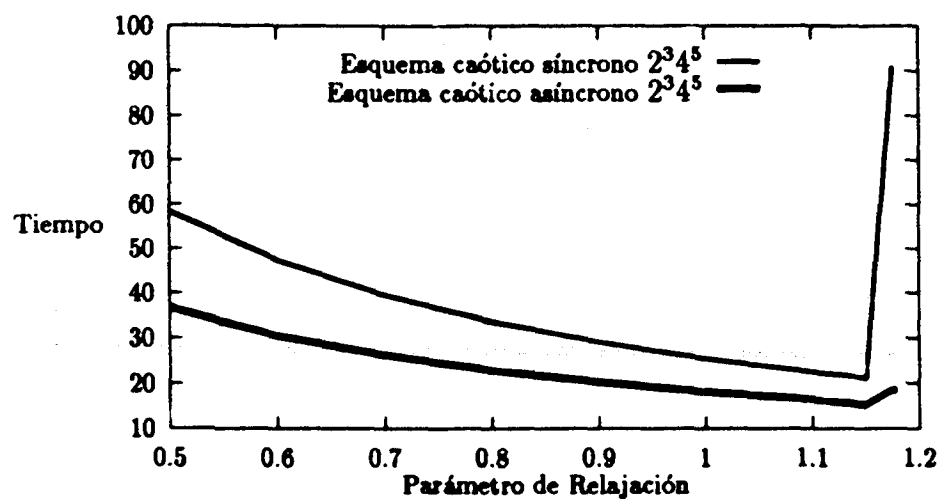


Figura 7: Comparación esquemas paralelos caóticos síncrono y asíncrono, $n = 5632$.

Los resultados obtenidos para los esquemas paralelos caóticos asíncronos, se presentan en la Tabla 3, calculando experimentalmente el mejor valor del parámetro ω . Para estos parámetros, la comparación entre los modelos caóticos y no caóticos nos lleva a las mismas conclusiones que en el caso síncrono. Es decir, para el parámetro ω_{opt} , por lo general estos modelos caóticos no aceleran a los no caóticos.

En esta tabla, $iter(j)$, $j = 1, 2, \dots, 8$, indica el número de iteraciones que realiza cada procesador, que puede variar, como se ha explicado a lo largo de esta memoria, en distintas ejecuciones. Debido a que en los modelos asíncronos cada procesador actualiza la solución global con independencia del resto, no es factible el establecer una versión secuencial única para cada ejecución.

Esto nos ha sugerido, como en los capítulos anteriores, calcular el siguiente speed-up

$$S_i = \frac{T_{mcs}}{T_i}, \quad i = 1, 2,$$

donde T_{mcs} es el tiempo del mejor algoritmo caótico secuencial de los reflejados en la Tabla 1. Puesto que no comparamos el mismo algoritmo ejecutado en secuencial y en paralelo, este speed-up puede ser superior al número de procesadores utilizado, como se observa para algunos resultados de la Tabla 3.

Orden	$q(l, j), 1 \leq j \leq 8$	ω_{opt}	iter(j), $1 \leq j \leq 8$	Tiempo (seg.)		Speed-up	
				T_1	T_2	S_1	S_2
384	1,1,1,1,2,2,2,2	1.350	524,517,525,515,400,367,400,474	0.433		10.98	
		1.350	555,542,555,543,441,437,441,487		0.380		12.51
	2,2,2,4,4,4,4,4	1.175	585,508,514,508,175,185,212,267	0.634		5.96	
		1.175	545,490,501,490,280,285,321,371		0.608		7.82
	2,2,2,2,2,2,2,2	1.300	432,375,380,375,538,551,538,638	0.574		8.28	
		1.325	433,390,397,389,573,567,571,633		0.488		9.74
	3,3,3,3,6,6,6,6	1.150	522,405,411,406,169,167,210,253	1.034		4.59	
		1.150	462,382,391,383,192,201,229,261		0.761		6.25
1,1,1,1,1,1,1,1	1.425	486,474,486,478,755,751,752,734	0.409		11.62		
	1.425	535,521,536,524,791,777,786,766		0.372		12.78	
1024	3,3,3,3,3,3,3,3	1.525	751,642,650,642,650,641,650,742	2.618		7.58	
		1.550	728,651,663,653,661,651,663,709		2.189		9.07
	2,2,2,2,2,2,2,2	1.650	708,650,657,650,659,650,659,696	1.789		11.10	
		1.650	741,691,708,691,704,691,707,723		1.617		12.29
	1,1,1,1,1,1,1,1	1.675	960,933,958,936,958,937,960,936	1.688		11.77	
		1.675	1045,1017,1044,1017,1044,1016,1044,1013		1.536		12.94
2816	1,1,1,2,2,2,2,2	1.250	617,599,636,503,515,503,510,589	4.109		8.92	
		1.250	640,615,661,544,558,501,565,5604		3.644		10.06
	2,2,2,4,4,4,4,4	1.150	634,568,567,273,275,274,320,414	6.424		5.71	
		1.150	593,527,502,286,300,277,376,376		5.257		6.98
	2,2,2,2,2,2,2,2	1.225	427,385,393,524,553,435,572,684	4.591		7.99	
		1.225	430,384,403,586,616,491,613,632		3.662		10.01
	1,1,1,1,1,1,1,1	1.400	471,457,471,722,729,722,666,724	3.206		11.44	
		1.400	524,500,516,786,796,786,796,699		2.790		13.14
4096	2,2,4,4,6,6,6,6	1.300	2998,2912,2756,2722,2417,2311,2413,3136	50.460		4.93	
		1.275	3064,2798,2804,2780,2851,2755,2847,3294		47.724		5.22
	1,1,2,2,3,3,3,3	1.525	3137,2954,3812,3635,4111,3936,4064,4652	40.102		6.21	
		1.500	3496,3281,3588,3583,4111,3824,4018,4019		35.829		6.95
	1,1,1,1,1,1,1,1	1.775	1536,1571,2832,2635,4373,4145,4255,4100	22.555		11.04	
		1.800	1475,1546,2572,2421,3511,3332,3646,3278		16.761		14.86

Tabla 3: Resultados de los esquemas paralelos caóticos asíncronos para ω óptimo.

Orden	$q(l, j), 1 \leq j \leq 8$	ω_{opt}	iter(j), $1 \leq j \leq 8$	Tiempo (seg.)		Speed-up	
				T_1	T_2	S_1	S_2
5632	1,1,1,2,2,2,2,2	1.250	20,620,621,502,497,497,528,604	8.741		8.73	
		1.250	683,604,693,562,589,533,589,604		8.731		8.743
	2,2,2,4,4,4,4,4	1.150	67,589,590,276,280,296,349,419	14.692		5.20	
		1.150	33,504,538,293,329,334,381,381		12.589		6.06
	2,2,2,2,2,2,2,2	1.225	428,390,390,533,552,528,550,690	9.932		7.68	
		1.225	41,384,411,610,610,582,643,656		8.739		8.73
	1,1,1,1,1,1,1,1	1.375	186,484,495,748,780,780,810,780	6.890		11.08	
		1.450	168,468,495,706,738,622,690,647		5.932		12.87

Tabla 3. (Continuación) Resultados de los esquemas paralelos caóticos asíncronos para ω óptimo.

Las Figuras 8, 9, 10, 11 y 12 comparan para cada una de las matrices, los distintos esquemas caóticos asíncronos en un amplio rango del factor ω .

Para las matrices de orden 384, 2816 y 5632, obtenemos como mejor esquema, el esquema no caótico asíncrono, mientras que para la matriz de orden 1024 el mejor esquema es el esquema caótico 2^8 y para la matriz de orden 4096, lo es el esquema $2^2 4^2 6^4$. Nótese que para estas dos últimas matrices el tamaño de los bloques diagonales es relativamente pequeño en comparación con el tamaño de la matriz. Recordamos que, según observamos en la Sección 4.4, cuando mayor es el tamaño de dichos bloques diagonales, mayor es el orden de los subsistemas que cada procesador debe resolver, conforme aumenta el valor de los factores caóticos.

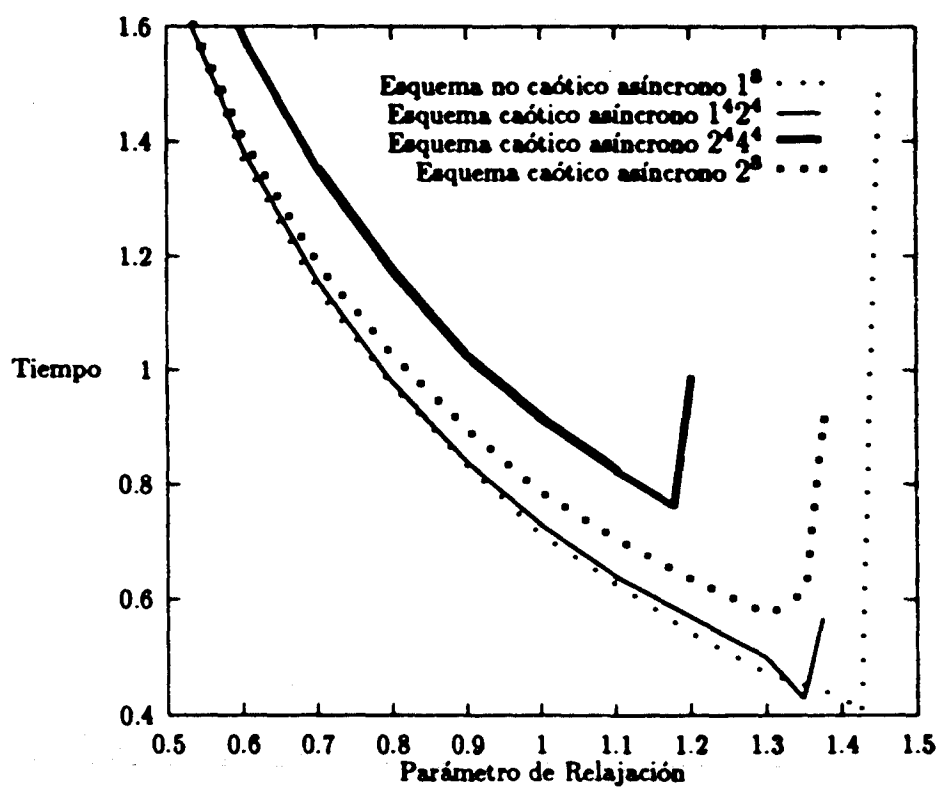
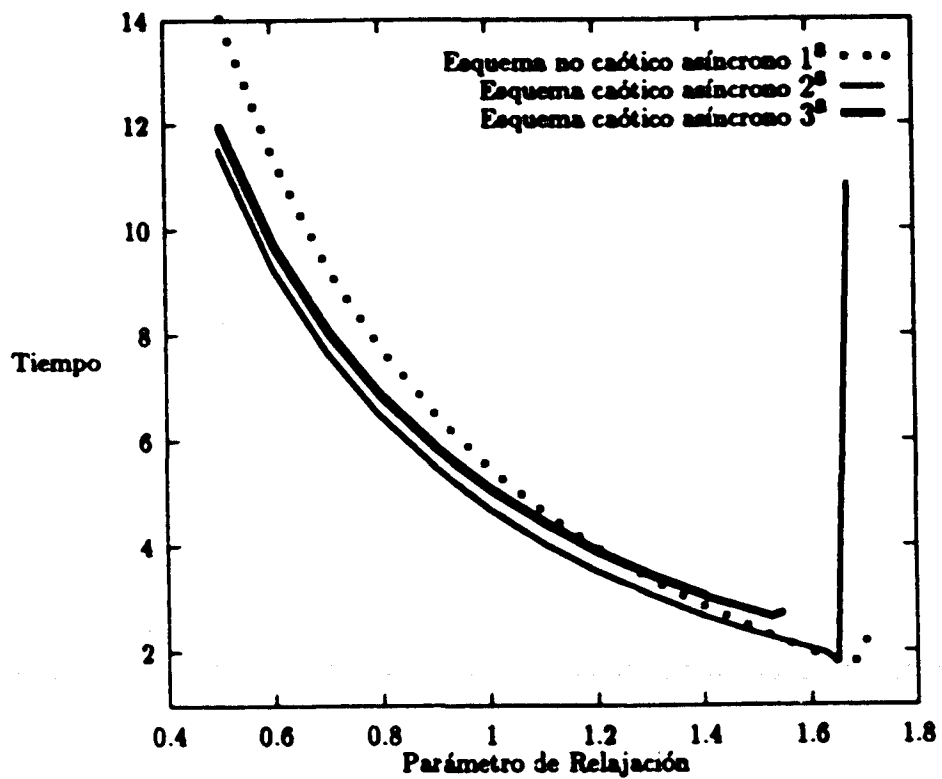


Figura 8: Comparación esquemas paralelos caóticos asíncronos, $n = 384$.

**Figura 9: Comparación esquemas paralelos caóticos asíncronos, $n = 1024$.**

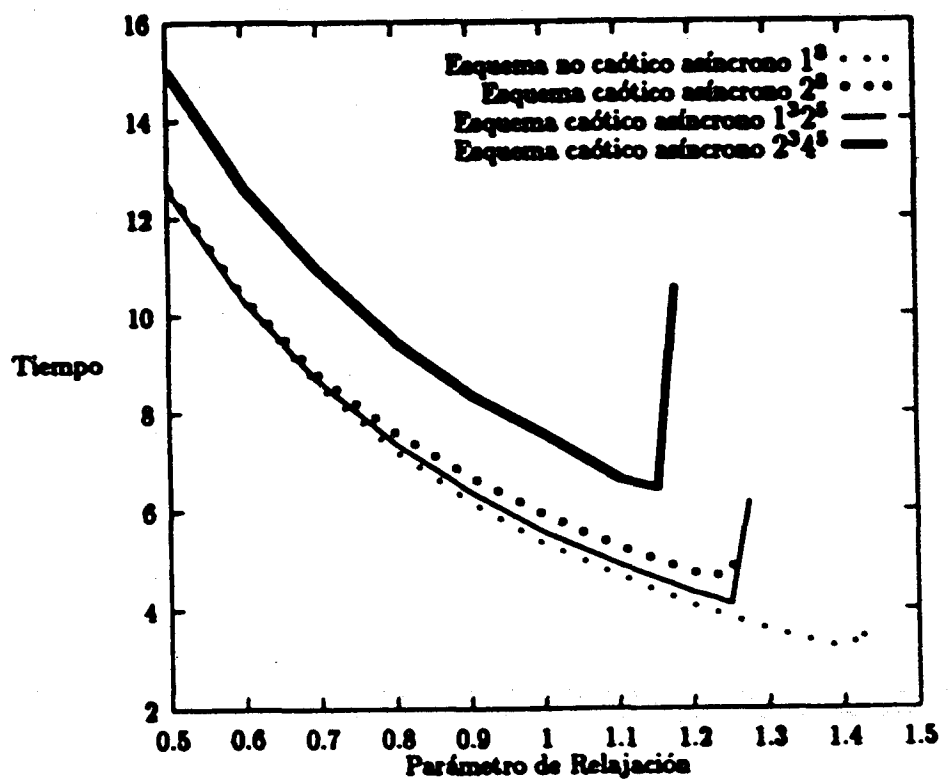


Figura 10: Comparación esquemas paralelos caóticos asíncronos, $n = 2816$.

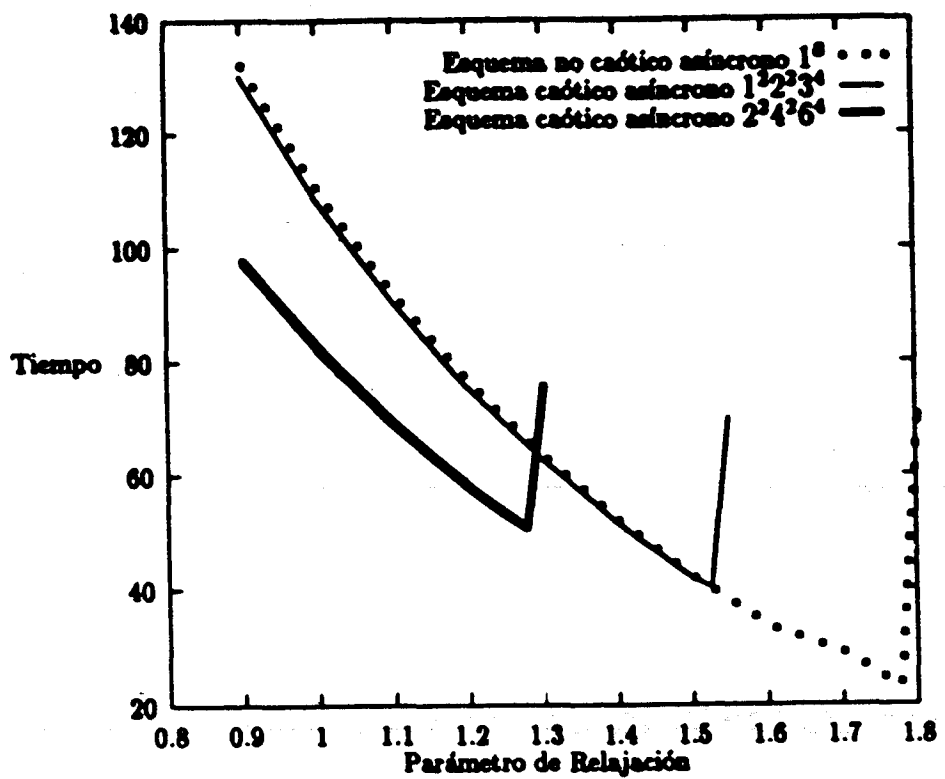


Figura 11: Comparación esquemas paralelos caóticos asíncronos, $n = 4096$.

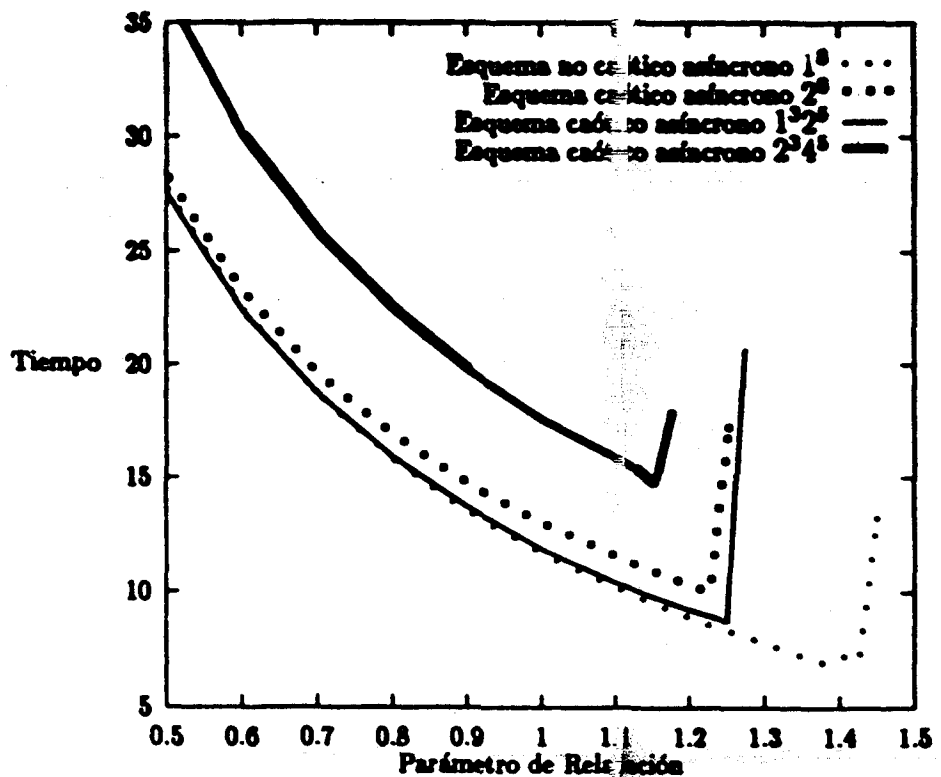


Figura 12: Comparación esquemas paralelos caóticos asíncronos, $n = 5632$.

Para finalizar esta sección, y una vez estudiados experimentalmente todos los modelos que se presentan en esta memoria, parece lógico preguntarse cuál de estos métodos es más eficiente aplicado sobre este tipo de matrices. Para esto se ha comparado, para cada matriz, los mejores esquemas caóticos, tanto síncronos como asíncronos, obtenidos experimentalmente en cada capítulo.

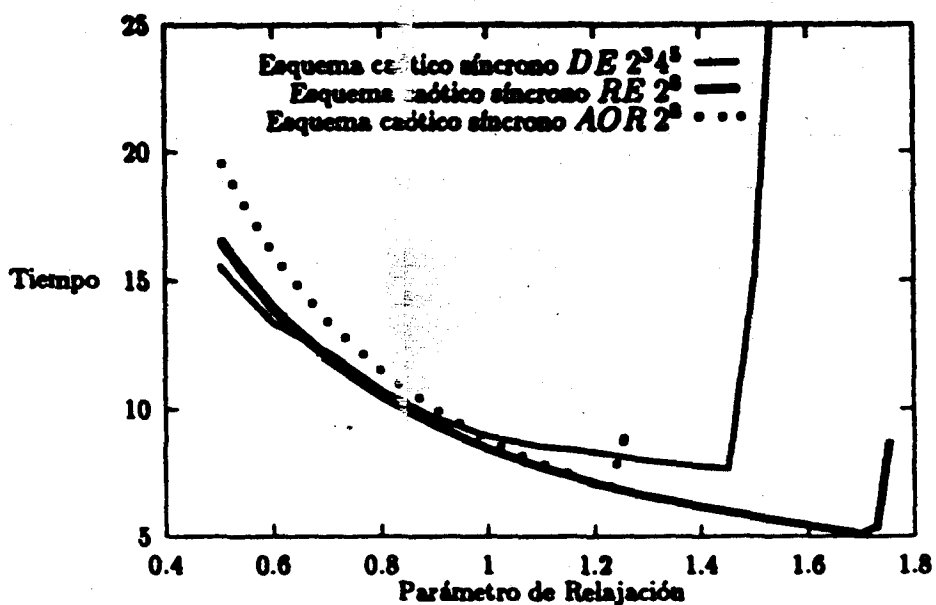


Figura 13: Comparación esquemas paralelos caóticos síncronos *DE*, *RE* y *AOR*, $n = 2816$.

De esta comparación no se deduce que alguno de los métodos sea mejor que el resto, como puede observarse en las Figuras 13, 14 y 15, en el caso síncrono y en las Figuras 16, 17 y 18, en el caso asíncrono. En estas figuras se denota por *DE* a los modelos caóticos en dos etapas estudiados en el Capítulo 1, por *RE* a los modelos relajados planteados en el Capítulo 2 y por *AOR* a los discutidos en este capítulo. Una conclusión que podemos extraer de esta comparación es que todos los modelos caóticos estudiados en esta memoria son perfectamente válidos como alternativa a los modelos paralelos clásicos.

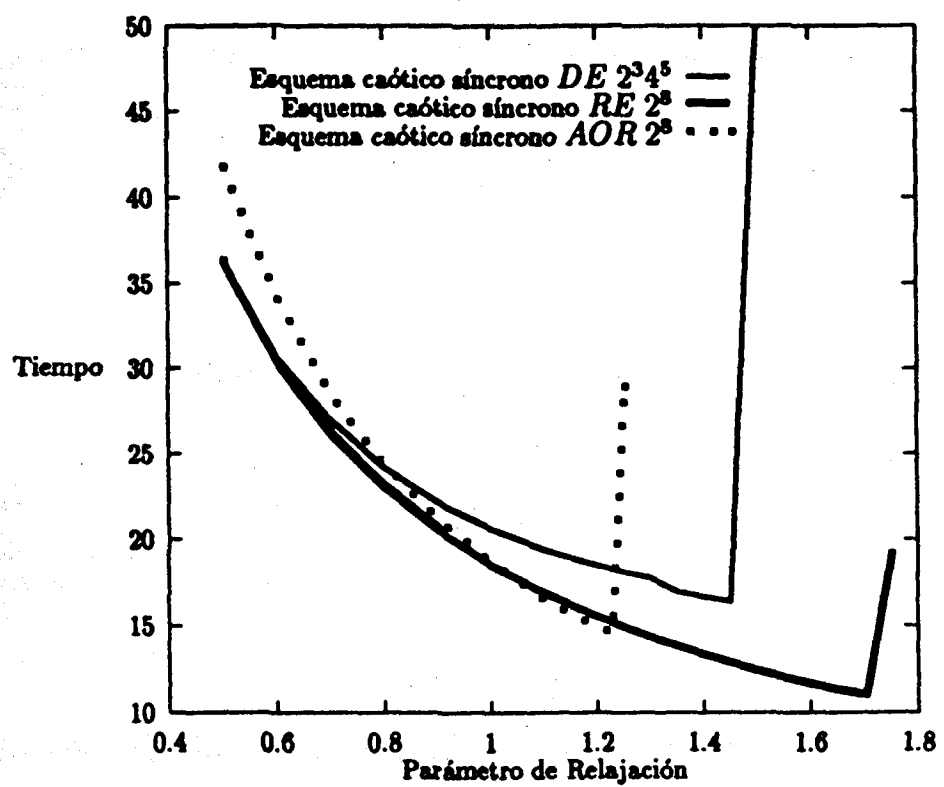


Figura 14: Comparación esquemas paralelos caóticos síncronos DE , RE y AOR , $n = 5632$.

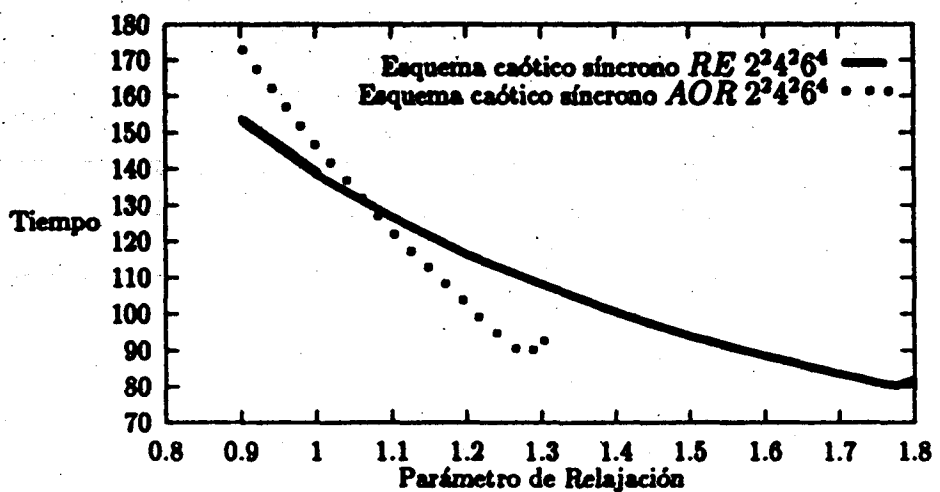


Figura 15: Comparación esquemas paralelos caóticos síncronos RE y AOR, $n = 4096$.

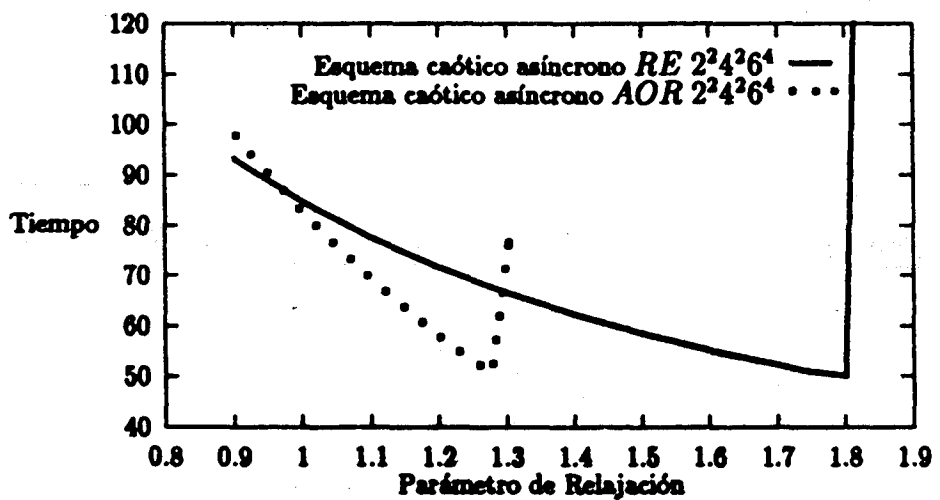


Figura 16: Comparación esquemas paralelos caóticos asíncronos RE y AOR, $n = 4096$.

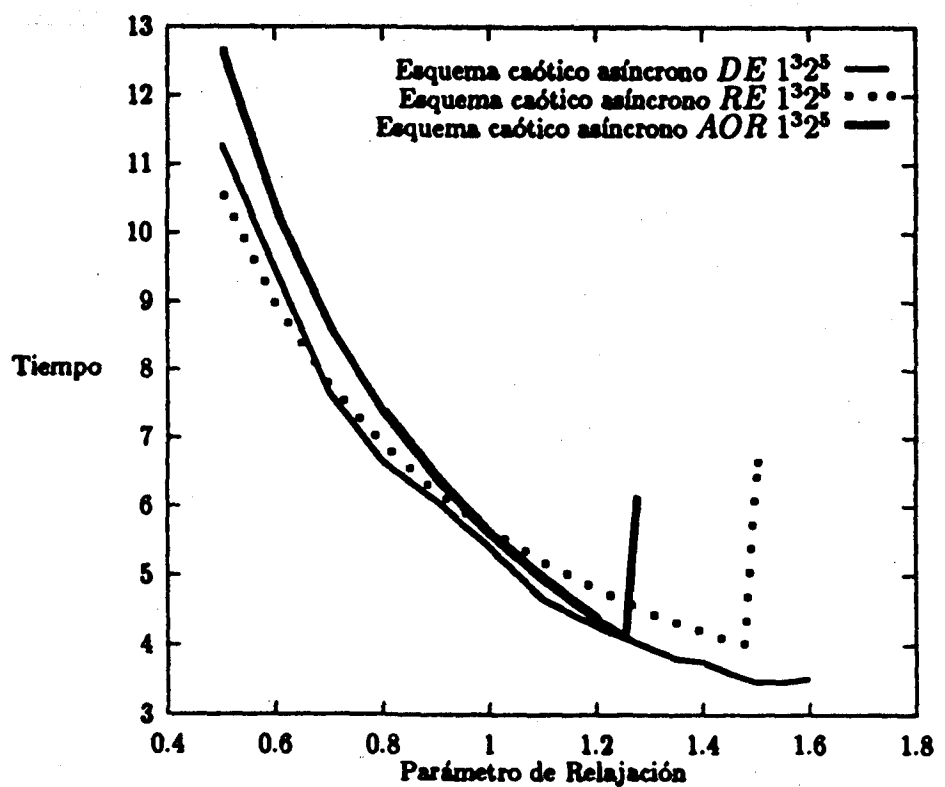


Figura 17: Comparación esquemas paralelos caóticos asíncronos DE , RE y AOR , $n = 2816$.

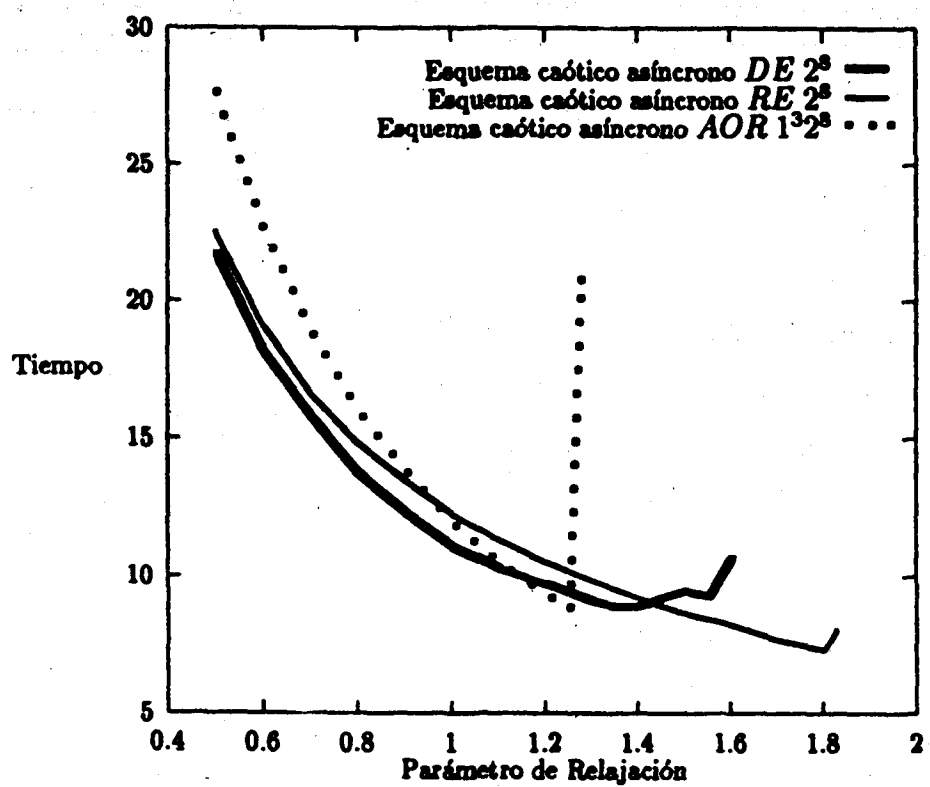


Figura 18: Comparación esquemas paralelos caóticos asíncronos DE , RE y AOR , $n = 5632$.

5.5 Conclusiones

En este Capítulo se ha planteado un método caótico de resolución de sistemas de ecuaciones lineales basado en el método *AOR*. Este método, como los construidos en los Capítulos 3 y 4, está diseñado para resolver el sistema de ecuaciones $Ax = b$ concurrentemente, e intentando equilibrar la carga de trabajo entre los distintos procesadores. Sin embargo, en su formulación se introducen dos parámetros reales, con el fin de acelerar la convergencia, μ y ω , denominados parámetro de aceleración y parámetro de sobrerelajación respectivamente. En la Sección 5.2 se construye y estudia la convergencia de la versión síncrona de dicho método, definida en la expresión (5.9). La Sección 5.3 está dedicada al esquema asíncrono, definido en la expresión (5.45). Hemos demostrado la convergencia de ambas versiones, cuando la matriz del sistema lineal es una H -matriz y los parámetros μ y ω satisfacen $0 < \mu \leq \omega < \frac{2}{1+\rho}$ (ver Teoremas 3 y 4).

En los experimentos numéricos realizados, volvemos a constatar que los modelos caóticos síncronos son mejores cuando se consigue un mejor equilibrio entre la carga de los procesadores. Sin embargo, en el modelo asíncrono, existen casos en los que se obtienen mejores resultados que en la versión no caótica ($q(l, j) = 1$), es decir, en el caso en el que cada procesador, de forma independiente, contribuye a la solución global, cuando ha realizado una única iteración local. No obstante, cabe destacar la superioridad de los modelos caóticos asíncronos respecto a los caóticos síncronos.

En este capítulo se ha realizado experimentalmente, un estudio comparativo

de los métodos contruidos en esta memoria, observando que el comportamiento de los distintos métodos no es análogo para todas las matrices y particiones, dando lugar a que no exista una superioridad clara de ningún método frente a los demás.

Líneas Futuras

En esta memoria se han construido distintos modelos paralelos caóticos, para la resolución de sistemas de ecuaciones lineales, estudiando la convergencia para el caso en que la matriz del sistema lineal $Ax = b$ es una H -matriz, que como se ha visto en el Teorema 9 del Capítulo 1, es una clase importante de matrices. Parece atractivo estudiar la convergencia de los modelos estudiados aquí, para otro tipo de matrices: matrices simétricas y definidas positivas, L -matrices ($A = [a_{ij}]_{1 \leq i, j \leq n}$ es una L -matriz si $a_{ii} > 0$ y $a_{ij} \leq 0$, $i \neq j$, $i, j = 1, 2, \dots, n$),

Los métodos estudiados en los Capítulos 3 y 4 introducen un parámetro de relajación, mientras que el método planteado en el Capítulo 5 introduce un parámetro de aceleración y otro de sobrerrelajación. En esta memoria se ha obtenido, para los esquemas caóticos estudiados, el valor óptimo para estos parámetros de forma experimental. Parece interesante estudiar de forma teórica la elección de estos parámetros.

En el Capítulo 3, se estudia la convergencia del método en dos etapas generalizado, cuando el parámetro de relajación está en el intervalo $(0, 1]$. Se desea extender este resultado demostrando, la convergencia en un intervalo $(0, \omega_0)$,

con $\omega_0 > 1$.

Todos los métodos caóticos estudiados en esta memoria, se basan en un esquema iterativo secuencial de primer orden

$$\mathbf{x}^{(k)} = B\mathbf{x}^{(k-1)} + \mathbf{c}, \quad k = 1, 2, \dots \quad (5.48)$$

Otra posibilidad de conseguir cierta aceleración, no contemplada en esta memoria, consiste en convertir el esquema (5.48) en un proceso iterativo de segundo orden, de la forma

$$\mathbf{x}^{(k)} = B_1\mathbf{x}^{(k-1)} + B_2\mathbf{x}^{(k-2)} + \mathbf{c}, \quad k = 1, 2, \dots$$

Este tipo de esquema, fue estudiado por Golub y Varga [42] en 1961. Un estudio detallado de estos métodos puede verse en Young [83]. Avdelas, de Pillis, Hadjidimos y Neumann [4] estudian bajo qué condiciones se consigue acelerar la convergencia, mediante un proceso de segundo orden. Por otro lado en [33], se estudió una técnica de segundo orden que en determinados casos aceleraba al método de Jacobi caótico extrapolado. Todo esto sugiere la construcción y estudio de la convergencia de esquemas iterativos de segundo orden basados en los construidos en esta memoria.

En los Capítulos 4 y 5, a medida que aumenta el valor de los factores caóticos, aumenta el tamaño de los subsistemas que cada procesador debe resolver. Esta situación, tal y como se explicó en el Capítulo 4, puede dar lugar a que los procesadores realicen demasiados cálculos. Esto no ocurre así en los esquemas planteados en el Capítulo 3, en los que las variables calculadas por cada procesador en cada etapa interna son totalmente independientes del resto. Esto motiva el estudio de esquemas análogos a los planteados en los Capítulos 4 y 5,

pero contruidos con dicha condición de independencia. Aunque la convergencia teórica de este tipo de esquemas no está demostrada, hemos podido constatar su buen comportamiento mediante algunos resultados experimentales.

Todos los métodos estudiados en esta memoria han sido implementados en un multiprocesador de memoria compartida Alliant FX/80. Ya que uno de los objetivos de estos métodos radica en el equilibrio de la carga de trabajo entre los procesadores, es interesante su experimentación en multicomputadores con memoria distribuida, como por ejemplo el PARSYS SN-1000 Supernode de la Universidad Politécnica de Valencia. Por otro lado, teniendo en cuenta el hardware disponible en nuestro Departamento se ha pensado realizar experimentos sobre una red de estaciones de trabajo usando como soporte software *PVM* (parallel virtual machine). *PVM* es un paquete de software que permite que una red de computadores heterogéneos puedan actuar como un sólo recurso computacional concurrente.

Estas son básicamente, las líneas en las que se enmarcarán, nuestros próximos trabajos de investigación.

Bibliografía

- [1] L. M. ADAMS, R. J. LEVEQUE, Y D. M. YOUNG. Analysis of the SOR Iteration for the 9-point Laplacian. *SIAM Journal Numer. Anal.*, **25**: 1156–1180 (1988).
- [2] P. ALBRECHT Y M. P. KLEIN. Extrapolated Iterative Methods for Linear Systems. *SIAM Journal Numerical Analysis*, **21**: 192–201 (1984).
- [3] M. ARIOLI Y F. ROMANI. Stability, Convergence, and Conditioning of Stationary Iterative Methods of the Form $x^{(i+1)} = Px^{(i)} + q$ for the Solution of Linear Systems. *IMA Journal of Numerical Analysis*, **12**: 21–30 (1992).
- [4] G. AVDELAS, J. DE PILLIS, A. HADJIDIMOS, Y M. NEUMAN. A Guide to the Acceleration of iterative Methods whose Iteration Matrix is Nonnegative and Convergent. *SIAM Journal on Matrix Analysis and Applications*, **9**: 329–342 (1988).
- [5] R. E. BANK Y D. J. ROSE. Global Approximate Newton Methods. *Numerische Mathematik*, **37**: 279–295 (1981).

- [6] B. F. BEIDAA Y G. P. PAPAVALASSILOPOULOS. Convergence Analysis of Asynchronous Linear Iterations with Stochastic Delays. *Parallel Computing*, 19: 281–302 (1993).
- [7] A. BERMAN Y R.J. PLEMMONS. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York, 1979.
- [8] D. P. BERTSEKAS Y J. N. TSITSIKLIS. *Parallel and Distributed Computation*. Prentice-Hall Internacional, Inc, New Jersey, 1989.
- [9] A. BHAYA, E. KASZKUREWICZ, Y F. MOTA. Asynchronous Block-Iterative Methods for Almost Linear Equations. *Linear Algebra and its Applications*, 154–156: 487–508 (1991).
- [10] R. BRU, L. ELSNER, Y M. NEUMANN. Models of Parallel Chaotic Iteration Methods. *Linear Algebra and its Applications*, 103: 175–192 (1988).
- [11] R. BRU Y R. FUSTER. Parallel Chaotic Extrapolated Jacobi Method. *Appl. Math, Letters*, 3(4): 65–69 (1990).
- [12] R. BRU, V. MIGALLÓN, Y J. PENADÉS. Chaotic Methods for the Parallel Solution of Linear Systems. Sometido.
- [13] D. CHAZAN Y W. MIRANKER. Chaotic Relaxation. *Linear Algebra and its Applications*, 2: 199–222 (1969).
- [14] B. CODENOTTI Y P. FAVATI. New Techniques for the Solution of Linear Systems by Iterative Methods. *Comput. Math. Applic.*, 14: 425–428 (1987).
- [15] J. DANCIS. The Optimal ω is not best for the SOR Iteration Method. *Linear Algebra and its Applications*, 154: 819–835 (1991).

- [16] V. DEL OLMO Y R. FUSTER. Some Iterative Methods Related to Frobenius Norm Minimization. *Computers Math. Applic*, 22: 121-126 (1991).
- [17] R. S. DEMBO, S. C. EISENSTAT, Y T. STEihaug. Inexact Newton Methods. *SIAM Journal on Numerical Analysis*, 18: 400-408 (1982).
- [18] W. DEREN. On the Convergence of the Parallel Multisplitting AOR Algorithm. *Linear Algebra and its Applications*, 154-156: 473-486 (1991).
- [19] J. J. DONGARRA, I. S. DUFF, D. C. SORENSEN, Y H.A. VAN DER VORST. *Solving Linear Systems on Vector and Shared Memory Computers*. SIAM, Philadelphia, 1991.
- [20] J. DOUGLAS. Alternating Direction Iteration for Mildly non Linear Difference Equations. *Numerische Mathematik*, 3: 92-98 (1961).
- [21] T. DUPONT. Factorization Procedure for the Solution of Elliptic Difference Equations. *Journal on Numerical Analysis*, 5: 753-782 (1968).
- [22] L. ELSNER. Comparisons of Weak Regular Splittings and Multisplittings Methods. *Numerische Mathematik*, 56: 283-290 (1989).
- [23] L. ELSNER Y M. NEUMANN. Monotonic Sequences and Rates of Convergence of Asynchronized Iterative Methods. *Linear Algebra and its Applications*, 180: 17-33 (1993).
- [24] L. ELSNER, M. NEUMANN, Y B. VEMMER. The Effect of the Number of Processors on the Convergence of the Parallel Block Jacobi Method. *Linear Algebra and its Applications*, 154-156: 311-330 (1991).

- [25] D. J. EVANS Y C. LI. Extrapolated Gauss-Seidel and SOR Methods for Least-Squares Problems. *Numerische Mathematik*, 53: 485-498 (1988).
- [26] J. FENG Y Z. RONG-FU. An Error Bound on the AOR Method for Solving Linear Systems. *Linear Algebra and its Applications*, 158: 61-75 (1991).
- [27] A. FROMMER. Parallel Nonlinear Multisplitting Methods. *Numerische Mathematik*, 56: 269-282 (1989).
- [28] A. FROMMER Y KARLSRUHE G. MAYER. Safe Bounds for the Solutions of Nonlinear Problems using a Parallel Multisplitting Method. *Computing*, 42: 171-186 (1989).
- [29] A. FROMMER Y G. MAYER. Convergence of Relaxed Parallel Multisplitting Methods. *Linear Algebra and its Applications*, 119: 141-152 (1989).
- [30] A. FROMMER Y G. MAYER. Parallel Interval Multisplittings. *Numerische Mathematik*, 56: 255-267 (1989).
- [31] A. FROMMER Y D. SZYLD. H -Splittings and Two-Stage Iterative Methods. *Sometido*.
- [32] A. FROMMER, WUPPERTAL, Y KARLSRUHE G. MAYER. On the Theory and Practice of Multisplitting Methods in Parallel Computation. *Computing*, 49: 63-74 (1992).
- [33] R. FUSTER, V. MIGALLÓN, Y J. PENADÉS. Asynchronous Parallel Extrapolated Methods. *Sometido*, 1993.

- [34] R. FUSTER, V. MIGALLÓN, Y J. PENADÉS. Resultados Prácticos del Método de Jacobi Caótico en Paralelo. Actas II Jornadas de Paralelismo, San Lorenzo de El Escorial, 1991.
- [35] R. FUSTER, V. MIGALLÓN, Y J. PENADÉS. Un Método Caótico de Segundo Orden en Paralelo. Actas IV Jornadas de Paralelismo, Santander, 1993.
- [36] P. D. GADER. Necessary and Sufficient Conditions for the Existence of Local Matrix Decompositions. *SIAM Journal on Matrix Analysis and Applications*, 9: 305–313 (1988).
- [37] K. A. GALLIVAN, R. J. PLEMMONS, Y A. H. SAMEH. Parallel Algorithms for Dense Linear Algebra Computations. *SIAM Review*, 1: 54–135 (1990).
- [38] L. GIRAUD Y P. SPITERI. Résolution Parallèle de Problèmes aux Limites non Linéaires. *Modélisation mathématique et Analyse numérique*, 25: 579–606 (1991).
- [39] I. GONBERG, P. LANCASTER, Y L. RODMAN. *Invariant Subspaces of Matrices with Applications*. John Wiley and Sons, New York, 1986.
- [40] G. H. GOLUB Y M. L. OVERTON. Convergence of a Two-Stage Richardson Iterative Procedure for Solving Systems of Linear Equations. In G.A. Watson, editor, *Numerical Analysis (Proceedings of the Ninth Biennial Conference, Dundee, Scotland, 1981)*, *Lecture Notes in Mathematics 912*, pages 128–139, New York, 1982. Springer Verlag.

- [41] G. H. GOLUB Y M. L. OVERTON. The Convergence of Inexact Chebyshev and Richardson Iterative Methods for Solving Linear Systems. *Numerische Mathematik*, **53**: 571–593 (1988).
- [42] G.H. GOLUB Y R.S. VARGA. Chebyshev Semi-Iterative Methods, Successive Over-Relaxation Iterative Methods, and Second-Order Richardson Iterative Methods. *Numerische Mathematik*, **3**: 147–168 (1961).
- [43] A. D. GUNAWARDENA, S. K. JAIN, Y L. SNYDER. Modified Iterative Methods for Consistent Linear Systems. *Linear Algebra and its Applications*, **154–156**: 123–143 (1991).
- [44] M. H. GUTKNECHT, W. NIETHAMMER, Y R. S. VARGA. K -step Iterative Methods for Solving Nonlinear Systems of Equations. *Numerische Mathematik*, **48**: 699–712 (1986).
- [45] A. HADJIDIMOS. Accelerated Overrelaxation Method. *Mathematics of computation*, **32**: 149–157 (1978).
- [46] A. HADJIDIMOS. The Optimal Solution to the Problem of Complex Extrapolation of a First-Order Scheme. *Linear Algebra and its Applications*, **62**: 241–261 (1984).
- [47] M. HANKE Y M. NEUMANN. Preconditionings and Splittings for Rectangular Systems. *Numerische Mathematik*, **57**: 85–95 (1990).
- [48] R. A. HORN Y C. A. JOHNSON. *Matrix Analysis*. Cambridge University Press, 1985.

- [49] S. L. JOHANSSON Y CHING-TIEN HO. Algorithms for Matrix Transposition on Boolean n -Cubo Configured Ensemble Architectures. *SIAM Journal on Matrix Analysis and Applications*, 9: 419–454 (1988).
- [50] E. KASZKUREWICZ, A. BHAYA, Y D. D. SILJAK. On the Convergence of Parallel Asynchronous Block-Iterative Computation. *Linear Algebra and its Applications*, 131: 139–160 (1990).
- [51] I. KOLTRACHT Y P. LANCASTER. Constraining Strategies for Linear Iterative Processes. *IMA Journal of Numerical Analysis*, 10: 555–567 (1990).
- [52] P. LANCASTER Y M. TISMENESKY. *The Theory of Matrices*. Academic Press, Orlando, Florida, Second Ed., 1985.
- [53] P. J. LANZKRON, D. J. ROSE, Y D. B. SZYLD. Convergence of Nested Iterative Methods for Linear Systems. Por aparecer en *Numerische Mathematik*.
- [54] B. LUBACHEVSKY Y D. MITRA. A Chaotic Asynchronous Algorithm for Computing the Fixed Point of a Nonnegative Matrix of Unit Spectral Radius. *Journal of the Association for Computing Machinery*, 33: 130–150 (1987).
- [55] G. MAYER. Comparison Theorems for Iterative Methods Based on Strong Splittings. *SIAM Journal on Numerical Analysis*, 14: 215–227 (1987).
- [56] G. MEURANT. Parallel Algorithms for Supercomputers. *Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms*, 70: 289–318 (1991).

- [57] M. NEUMANN Y R. J. PLEMMONS. Convergence of Parallel Multisplitting Iterative Methods for M -Matrices. *Linear Algebra and its Applications*, 88-89: 559-573 (1987).
- [58] N. K. NICHOLS. On The Convergence of Two-Stage Iterative Processes for Solving Linear Equations. *SIAM Journal on Numerical Analysis*, 10: 460-469 (1973).
- [59] J. NIEPLOCHA, T. Z. MAI, Y C. C. CAFROLL. Asynchronous Algorithms for Solving Large Systems of Linear Equations on Parallel Computers. *So-metido*.
- [60] WILHELM NIETHAMMER. The SOR Method on Parallel Computers. *Nu-merische Mathematik*, 9: 247-254 (1989).
- [61] D. P. O'LEARY Y R. E. WHITE. Multi-Splittings of Matrices and Parallel Solution of Linear Systems. *SIAM Journal on Algebraic Discrete Methods*, 6: 630-640 (1985).
- [62] J. M. ORTEGA. *Numerical Analysis*. Academic Press, New York, 1972.
- [63] J. M. ORTEGA Y W. C. RHEINBOLDT. *Iterative Solution of Nonlinear Equations in Several Variables*. Prentice-Hall Internacional, Inc, New Jer-sey, 1989.
- [64] A. M. OSTROWSKI. Iterative Solution of Linear Systems of Functional Equations. *Journal of Mathematical Analysis and Applications*, 2: 351-369 (1961).

- [65] D. J. PIERCE, A. HADJIDIMOS, Y R. J. PLEMMONS. Optimality Relationships for p -Cyclic SOR. *Numerische Mathematik*, 56: 635–643 (1990).
- [66] SANZHENG QIAO. Recursive Least Squares Algorithm for Linear Prediction Problems. *SIAM Journal on Matrix Analysis and Applications*, 9: 323–328 (1988).
- [67] R. BEVILACQUA, B. CODENOTTI, Y F. ROMANI. Parallel Solution of Block Tridiagonal Linear Systems. *Linear Algebra and its Applications*, 104: 39–57 (1988).
- [68] W. C. RHEINBOLDT Y J. S. VANDERGRAFT. A Simple Approach to the Perron-Frobenius Theory for Positive Operators on General Partially-Ordered Finite-Dimensional Linear Spaces. *Math. Comp.*, 27: 139–145 (1973).
- [69] Y. ROBERT Y D. TRYSTRAM. Comments on Scheduling Parallel Iterative Methods on Multiprocessor Systems. *Parallel Computing*, 7: 253–255 (1988).
- [70] Y. S. RODITIS Y P. D. KIOUSIS. Parallel Multisplitting, Block Jacobi Type Solutions of Linear Systems of Equations. *International Journal for numerical methods in Engineering*, 29: 619–632 (1990).
- [71] A. H. SAMEH Y D. J. KUCK. On Stable Parallel Linear System Solvers. *Journal of the Association for Computing Machinery*, 25: 81–91 (1978).

- [72] Y. G. SARIDAKIS. On the Analysis of the Unsymmetric Successive Overrelaxation Method when Applied to p -Cyclic M -matrices. *Numerische Mathematik*, 49: 461–473 (1986).
- [73] U. SCHENDEL. *Introduction to Numerical Methods for Parallel Computers*. Ellis Horwood Limited, Chichester, 1984.
- [74] J. SMITH. The Coupled Equation Approach to the Numerical Solution of the Biharmonic Equation by Finite Differences. *I and II, Journal on Numerical analysis*, 5: 323–339 (1969).
- [75] H. S. STONE. An Efficient Parallel Algorithm for the Solution of a Tridiagonal Linear System of Equations. *Journal of the Association for Computing Machinery*, 20: 27–38 (1973).
- [76] J. L. STUART. Inflation Matrices and ZME-Matrices that Commute with a Permutation Matrix. *SIAM Journal on Matrix Analysis and Applications*, 9: 408–418 (1988).
- [77] D. B. SZYLD Y M. T. JONES. Two-stage and Multisplitting Methods for the Parallel Solution of Linear Systems. Sometido.
- [78] R. S. VARGA. *Matrix Iterative Analysis*. Prentice Hall, 1962.
- [79] A. M. VIDAL Y V. HERNANDEZ. A Parallel Algorithm for Computing the Singular Value Decomposition on a Distributed Memory Multiprocessor by using Rank-One Modifications. Pont-à-Mousson (France), 1990.
- [80] R. E. WHITE. Multisplitting of a Symmetric Positive Definite Matrix. *SIAM Journal on Matrix Analysis and Applications*, 11: 69–82 (1990).

- [81] H. K. WIMMER. Extremal Problems for Hölder Norms of Matrices and Realizations of Linear systems. *SIAM Journal on Matrix Analysis and Applications*, 9: 314–322 (1988).
- [82] E. G. D' YAKONOV. On the Solution of some Elliptic Difference Equations. *Numerische Mathematik*, 7: 1–20 (1971).
- [83] D. M. YOUNG. *Iterative Solution of Large Linear Systems*. Academic Press, New York, 1972.
- [84] E. L. ZAPATA, F. F. RIVERA, I. BENAVIDES, J. M. CARAZO, Y R. PESKIN. Multidimensional Fast Fourier Transform into SIMD Hypercubes. *IEE proceedings*, 137: 253–260 (1990).
- [85] E. L. ZAPATA, F. F. RIVERA, Y O. G. PLATA. On the Partition of Algorithms into Hypercubes. *Advances in Parallel Computing*, 1: 149–171 (1990).