

# Índice general

<b>1. Introducción</b>	<b>7</b>
1.1. Ecuación de Onda . . . . .	8
1.2. Ecuación de Poisson. . . . .	9
<b>2. Diferencias Finitas</b>	<b>10</b>
2.1. Método Explícito . . . . .	14
2.2. Método Implícito . . . . .	18
2.3. Consistencia . . . . .	21
2.4. Convergencia . . . . .	21
2.5. Estabilidad . . . . .	22
2.6. Teorema de Lax . . . . .	24
<b>3. Métodos Espectrales</b>	<b>26</b>
3.1. Formulación . . . . .	27
3.2. Polinomios de Chebyshev . . . . .	28
3.2.1. Diferenciación . . . . .	32
3.3. Aproximación Galerkin . . . . .	34
3.4. Método de Tau . . . . .	39
3.5. Método de Colocación . . . . .	39

<b>4. Ejemplos</b>	<b>41</b>
4.1. Diferencias Finitas . . . . .	41
4.1.1. Ecuación de Onda . . . . .	41
4.1.2. Ecuación de Poisson . . . . .	44
4.1.3. Método de Jacobi . . . . .	46
4.1.4. Método de Gauss-Seidel . . . . .	47
4.1.5. Método de Sobre Relajación Sucesiva (SOR) . . . . .	47
4.2. Método Espectral de Colocación . . . . .	49
4.2.1. Ecuación de Poisson . . . . .	49
4.3. Método Espectral Tau . . . . .	51
4.3.1. Ecuación de Burgers . . . . .	51
4.3.2. Implementación del método Tau . . . . .	52
<b>5. Resultados Numéricos</b>	<b>58</b>
5.1. Ecuación de Onda . . . . .	58
5.1.1. Resultados Numéricos aplicando Diferencias Finitas . .	58
5.2. Ecuación de Poisson . . . . .	63
5.2.1. Resultados Numéricos aplicando Diferencias Finitas . .	63
5.2.2. Resultados numéricos aplicando el Método de Colocación Espectral . . . . .	75
5.2.3. Comparación entre los Métodos Colocación Espectral y de Diferencias Finitas para la ecuación de Poisson . .	77
5.3. Ecuación de Burgers . . . . .	79
<b>A. Normas de Vectores y Matrices</b>	<b>83</b>
A.1. Normas de Vectores . . . . .	83
A.2. Normas de una Matriz . . . . .	84
A.3. Normas Compatibles . . . . .	84

A.4. Normas de una Matriz Subordinada . . . . .	85
A.5. Método de Runge-Kutta . . . . .	86
A.6. La Expansión Discreta de Fourier . . . . .	87
A.7. Solución de un Sistema Tridiagonal con la primer fila distinta de cero . . . . .	90
A.7.1. Algoritmo: . . . . .	94
A.8. Método de Eliminación de Gauss . . . . .	94

# Agradecimientos

Llegar a esta parte fue muy difícil, pero más difícil es encontrar las palabras correctas para expresar el profundo agradecimiento a DIOS y a todas las personas que estuvieron conmigo durante estos cinco largos años.

Pues bien, primero quiero agradecer a DIOS entre muchas otras cosas por haberme permitido terminar este trabajo. A mi familia: mi mamá Gisela, mi papá Gildardo, mis hermanos Nadia, Thelma, Immanuel, Violeta y Victoria por todas sus enseñanzas, sus consejos, por haber soportado mi mal carácter sobre todo durante los últimos meses, por todo el cariño y apoyo que me brindaron cuando más lo necesitaba, por no permitir que me rindiera en los momentos más difíciles. No tengo, ni tendré la manera de agradecerles todo lo que han hecho por mí.

Como olvidarme de mis mejores amigos Sarai y Antonio a quienes les agradezco por haberme brindado su amistad durante este tiempo, porque ellos también forman parte de este trabajo, por haber compartido conmigo mis frustraciones cuando las cosas no me salían bien, por escucharme cuando necesitaba que lo hicieran, por sopotar mi mal carácter y seguirme demostrando su apoyo y por tantas cosas que compartimos.

Quiero agradecer especialmente al Dr. Ricardo Becerril Barcenás por haberme tenido tanta paciencia, porque sé que trabajar conmigo no es fácil, por haber sido un excelente asesor, pues todas sus enseñanzas y todo el apoyo que me brindó es más de lo que un tesista puede esperar. Pero sobre

todo, gracias por haberme brindado su amistad y convertirse en uno de mis mejores amigos, MHUCHAS GRACIAS.

También agradezco al Dr. Gerardo Tinoco y a la Lic. Patricia Manriquez, porque si ellos no me hubieran dado la oportunidad de ingresar a esta Facultad, no estaría escribiendo estas líneas. Gracias por haberme permitido demostrar que lo único que se necesita para terminar esta carrera es el esfuerzo y el deseo de lograrlo, sin importar ser procedente de un bachillerato de no tan buen prestigio. Gracias Paty por todo el apoyo que siempre me brindaste, gracias por ser una de las pocas personas que se preocupa por los alumnos de esta Facultad aunque no lo reconozcan.

A todos los profesores que formaron parte de mi aprendizaje: Lic. Patricia, Dr. Gerardo Tinoco, Dr. Joaquin, Dr. Luis Valero, Dr. Mario Cesar, Dr. Ricardo por mencionar algunos. Al Secretario Académico Lic. Celerino, por todo el apoyo durante la carrera y durante el proceso de titulación.

Agradezco también al Dr. Francisco Siddartha Gúzman, por haber aceptado ser sinodal de esta tesis, así como por las acertadas observaciones que me permitieron mejorar el contenido de este trabajo.

Por ultimo, pero no por eso menos importante, agradezco al Instituto de Física y Matemáticas en especial a la Lic. Claudia Barriga, por haberme proporcionado un espacio para realizar mi trabajo.

A todos muchas, pero muchas gracias.

## *Resumen*

En esta tesis se hace una exposición de las generalidades de dos de los métodos numéricos más utilizados para resolver ecuaciones diferenciales parciales: el de *Diferencias Finitas* (MDF) y los *Métodos Espectrales* (ME). Los principios básicos expuestos, se ilustran con ejemplos de ecuaciones importantes en la física: la ecuación de calor, la ecuación de onda  $1 + 1$ , la de Burgers, y la de Poisson. Esta última se resuelve con ambas aproximaciones (MDF y ME) y se realiza una comparación de la velocidad de convergencia de ambos métodos.

# Capítulo 1

## Introducción

El hombre desde sus inicios, ha tratado de encontrar una explicación a todo fenómeno que se presenta a su alrededor. Primero, dieron explicaciones empíricas, pero con el paso del tiempo éstas fueron insuficientes, por lo que el hombre se vio en la necesidad de buscar otros métodos que le permitieran obtener explicaciones más acertadas.

De esta manera surgió la formulación matemática, que en la mayoría de los problemas que describe, involucra cambios con respecto a una o más variables independientes, lo cual nos conduce a ecuaciones diferenciales, ya sean ordinarias o parciales. Nosotros trabajaremos con ecuaciones diferenciales parciales de segundo orden, las cuales, en dos dimensiones se escriben en la forma

$$a \frac{\partial^2 \phi}{\partial x^2} + b \frac{\partial^2 \phi}{\partial x \partial y} + c \frac{\partial^2 \phi}{\partial y^2} + d \frac{\partial \phi}{\partial x} + e \frac{\partial \phi}{\partial y} + f \phi + g = 0 \quad (1.1)$$

donde  $a, b, c, e$ , y  $g$  pueden ser funciones de las variables independientes  $x$  e  $y$ , y de la variable dependiente  $\phi(x, y)$ . Dicha ecuación se clasifica de la siguiente manera: si  $b^2 - 4ac < 0$  a la ecuación se le llama elíptica, si  $b^2 - 4ac > 0$  se le llama hiperbólica y si  $b^2 - 4ac = 0$  se le llama parabólica.

Encontrar una solución a una ecuación o a un sistema de ecuaciones diferencial parciales, no es una tarea sencilla, y más cuando las ecuaciones son no-lineales. Es cierto, que en algunas situaciones, es posible encontrar soluciones exactas por métodos analíticos generalmente sofisticados, pero en muchos otros casos, es prácticamente imposible resolverlas analíticamente. Y si las condiciones de frontera son complicadas, la situación puede hacerse aún más compleja.

Es por esto, que el hombre se vio en la necesidad de desarrollar herramientas más eficaces, que le permitieran encontrar soluciones a dichas ecuaciones. Así surgieron varios métodos numéricos. Tres de los más conocidos y más ampliamente utilizados son: métodos de diferencias finitas, de elemento finito y los espectrales.

El propósito de este trabajo, es exponer brevemente los elementos básicos de dos métodos numéricos: el *Método de Diferencias Finitas* y los *Métodos Espectrales*, aplicados como ya se mencionó anteriormente a ecuaciones diferenciales parciales. Utilizando la ecuación de Poisson para hacer una comparación de la velocidad convergencia entre los dos métodos.

## 1.1. Ecuación de Onda

La ecuación 1.2 es la forma general de la ecuación que describe a las ondas en dimensión 1+1: "la segunda derivada del desplazamiento de la onda  $y$  respecto a la coordenada  $x$  es igual al inverso de la velocidad de propagación de la onda al cuadrado por la segunda derivada del desplazamiento de esta con respecto al tiempo."

$$\frac{\partial^2 y}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 y}{\partial t^2} = 0 \quad (1.2)$$



Sabemos que la forma general de la solución de ondas viajeras para esta ecuación es  $y(x, t) = f(x \pm vt)$  pero la forma específica depende de las condiciones iniciales y de frontera. Esta es la ecuación con la que lidaremos en esta tesis.

Pero la ecuación de onda no solo surge en la mecánica sino también en otras situaciones, tal es el caso de electromagnetismo, si usamos las ecuaciones electromagnéticas, obtendremos una ecuación de la misma forma que la que describe a las ondas mecánicas, solo que el desplazamiento  $y$  se sustituye por la intensidad del campo eléctrico y magnético. En este caso la velocidad de propagación de las ondas electromagnéticas que viajan en el vacío es la velocidad de la luz  $c$ . [1]

## 1.2. Ecuación de Poisson.

Una de las utilidades de la ecuación de Poisson se aprecia tanto en Electrodinámica, como en Gravitación, pues ésta describe la variación de potencial eléctrico o gravitacional en el interior de distribuciones eléctricas o de masa (ec.1.3).

$$\nabla^2 \varphi = -k\rho \quad (1.3)$$

ésta es la llamada *Ecuación de Poisson*, donde  $k = \frac{1}{\epsilon_0}$  para el caso de la electrodinámica y  $k = 4\pi G$  para el caso de Gravitación.

La solución para la ecuación 1.3 para  $\rho$  en cierta region  $V$  está dada por [2]

$$\varphi(r) = \frac{1}{4\pi\epsilon_0} \int_V \frac{\rho(r')}{|r - r'|} dv' \quad (1.4)$$

## Capítulo 2

# Diferencias Finitas

En este capítulo describiremos a grandes rasgos el Método de Diferencias Finitas, así como varios conceptos que éste involucra, tales como: *Método Explícito*, *Método Implícito*, *Consistencia*, *Estabilidad* y *Convergencia* (*Teorema de Lax*).

En *Diferencias Finitas*, las ecuaciones diferenciales se resuelven en cierta región  $\mathcal{R}$  acotada por una frontera, donde  $\mathcal{R}$  pertenece al espacio en el cual se describen las ecuaciones diferenciales, (es decir, si la ecuación depende de dos variables, entonces estaremos en un espacio de dos dimensiones, para nosotros dicho espacio será  $\mathcal{R}^2$ ).

Este método consiste, tal como su nombre lo dice, en aproximar por medio del teorema de Taylor las derivadas involucradas en la ecuación diferencial a la que se quiere dar solución. De esta manera obtendremos una expresión en términos de dichas aproximaciones, y así nos sea más sencillo encontrar la solución en cada punto de la región  $\mathcal{R}$ . Lo anterior se describe de una forma más amplia a continuación.

Sea  $F$ , una función tal que ésta y sus derivadas sean univaluadas, finitas y

continuas con respecto a cualquier variable  $x$ , entonces por el Teorema de Taylor tenemos las siguientes aproximaciones.

$$F(x + h) = F(x) + hF'(x) + \frac{1}{2}h^2F''(x) + \frac{1}{6}h^3F'''(x) + \dots \quad (2.1)$$

y

$$F(x - h) = F(x) - hF'(x) + \frac{1}{2}h^2F''(x) - \frac{1}{6}h^3F'''(x) + \dots \quad (2.2)$$

ahora sumemos 2.1 y 2.2, obteniendo como resultado

$$F(x + h) + F(x - h) = 2F(x) + h^2F''(x) + O(h^4) \quad (2.3)$$

en la ecuación 2.3,  $O(h^4)$  representa los términos que contienen cuartas potencias y superiores de  $h$ . Considerando que dichos términos son insignificantes comparados con los términos que tienen potencias de  $h$  inferiores, podemos llegar a

$$F''(x) = \left( \frac{d^2F}{dx^2} \right)_{x=x} \simeq \frac{1}{h^2} [F(x + h) - 2F(x) + F(x - h)] \quad (2.4)$$

con un error de orden  $h^2$ .

Si ahora hacemos la diferencia entre las ecuaciones 2.1 y 2.2, despreciando los términos de orden  $h^3$  nos dará como resultado una expresión para  $F'(x)$

$$F'(x) = \left( \frac{dF}{dx} \right)_{x=x} \simeq \frac{1}{2h} [F(x + h) - F(x - h)] \quad (2.5)$$

con un error de orden  $h^2$ .

Podemos observar en nuestra figura 2.1, que la ecuación 2.5 aproxima la pendiente de la tangente en P por la pendiente de la curva AB, esta ecuación es llamada *aproximación en diferencias centrales*. Pero también podemos

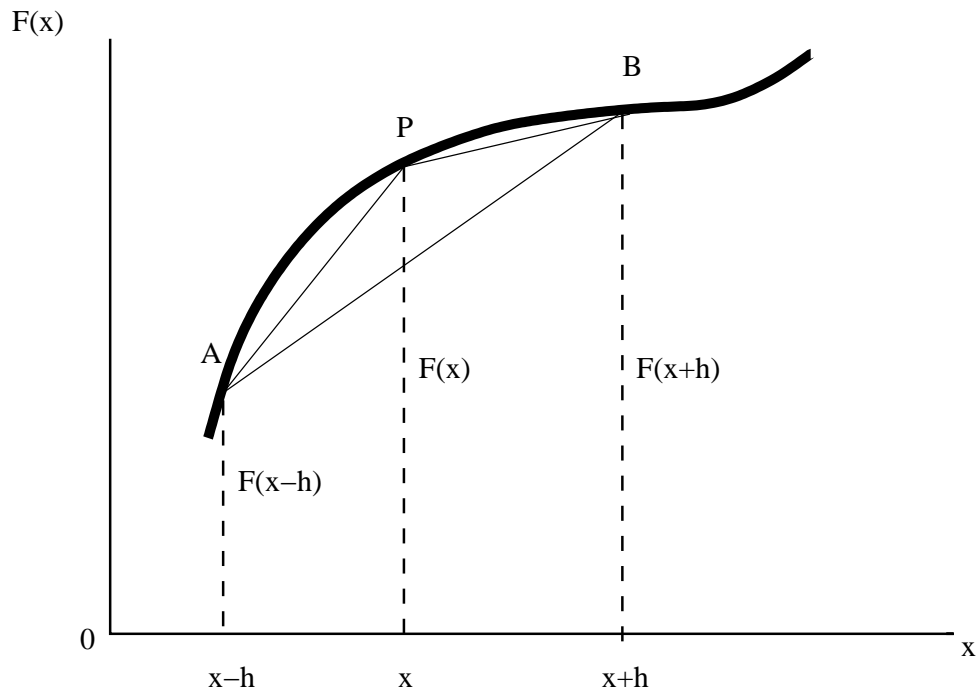


Figura 2.1: Aproximación de la pendiente de la tangente en  $P$ , por: la pendiente de la curva  $AB$ , la pendiente de la curva  $AP$  o la pendiente de la curva  $PB$

aproximar la pendiente de la tangente en  $P$ , ya sea por la pendiente de la curva  $PB$ , dada por la ecuación (*diferencias hacia delante*)

$$F'(x) \simeq \frac{1}{h} F(x+h) - F(x) \quad (2.6)$$

o por la pendiente de la curva  $AP$ , dada por la ecuación (*diferencias hacia atrás*).

$$F'(x) \simeq \frac{1}{h} [F(x) - F(x-h)] \quad (2.7)$$

Supongamos ahora que la función  $F$  depende de las variables  $x$  y  $y$ . Entonces la región  $\mathcal{R}$  estará contenida en el plano  $xy$ . Subdividiremos el plano, de tal manera que el eje horizontal quedará dividido en segmentos de tamaño  $\delta x = h$ , de esta forma la coordenada  $x$  estará representada por  $x_i = ih$  con

$i = 0, \pm 1, \pm 2, \dots$ , y el eje vertical lo dividiremos en segmentos de tamaño  $\delta y = k$ , así la coordenada  $y$  estará representada por  $y_j = jk$  con  $j = 0, \pm 1, \pm 2, \dots$

El valor de  $F$  en cada punto  $P(x_i, y_j) = P(ih, jk)$  de la red está representado por  $F_P = F(ih, jk)$  y para simplificar la notación, en adelante lo denotaremos como  $F_{i,j}$ .

De acuerdo con la ecuación (2.4)

$$\left(\frac{\partial^2 F}{\partial x^2}\right)_P = \left(\frac{\partial^2 F}{\partial x^2}\right)_{i,j} \simeq \frac{F\{(i+1)h, jk\} - 2F\{ih, jk\} + F\{(i-1)h, jk\}}{h^2} \quad (2.8)$$

en términos de la nueva notación

$$\left(\frac{\partial^2 F}{\partial x^2}\right)_{i,j} \simeq \frac{F_{i+1,j} - 2F_{i,j} + F_{i-1,j}}{h^2} \quad (2.9)$$

con un error de orden  $h^2$ . De igual forma para

$$\left(\frac{\partial^2 F}{\partial t^2}\right)_{i,j} \simeq \frac{F_{i,j+1} - 2F_{i,j} + F_{i,j-1}}{k^2} \quad (2.10)$$

con un error de orden  $k^2$ . De acuerdo con la notación de aproximación de diferencias hacia adelante para la expresión  $\frac{\partial F}{\partial x}$  en P está dada por

$$\frac{\partial F}{\partial x} \simeq \frac{F_{i,j+1} - F_{i,j}}{h} \quad (2.11)$$

con un error de orden de  $O(h)$ , similarmente para  $\frac{\partial F}{\partial t}$ , pero con un error de orden de  $O(k)$ . Hemos obtenido, las derivadas en términos de diferencias, tal como se mencionó arriba.

Para la explicación de las secciones del método Explícito e Implícito, nos auxiliaremos con la ecuación de Calor

$$\frac{\partial U}{\partial T} = \kappa \frac{\partial^2 U}{\partial X^2} \quad (2.12)$$

con  $\kappa$  constante.

Obteniendo una ecuación adimensional de la ecuación 2.12, de la forma

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (2.13)$$

con la ayuda de las substituciones  $u = \frac{U}{U_0}$ ,  $x = \frac{X}{L}$  y  $t = \kappa \frac{T}{L^2}$ , donde  $U$  es una temperatura a cierta distancia  $X$  a un tiempo  $T$  y  $U_0$  es la temperatura inicial a tiempo cero.

## 2.1. Método Explícito

Un método para encontrar la solución de ecuaciones con diferencias finitas, es utilizando el *Método Explícito*, el cual ilustraremos con ayuda de la ecuación de calor adimensional.

Una aproximación en diferencias finitas para la ecuación

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}. \quad (2.14)$$

Se obtiene aplicando la ecuación 2.11 al lado derecho y la ecuación 2.9 (para la variable  $t$ ) al lado izquierdo de 2.14,

$$\frac{u_{i,j+1} - u_{i,j}}{k} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \quad (2.15)$$

donde  $u$  es la solución exacta de las ecuaciones de diferencias aproximadas, con

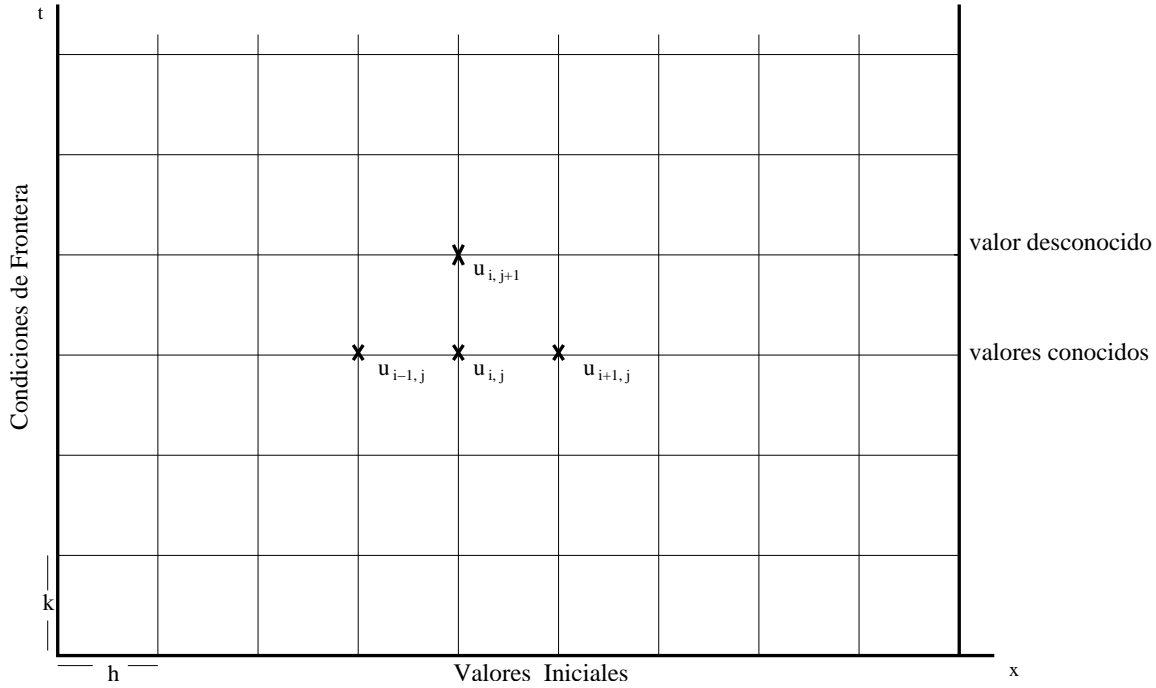


Figura 2.2: Representación gráfica de los valores conocidos  $u_{i-1,j}$ ,  $u_{i,j}$ ,  $u_{i+1,j}$  y del valor desconocido  $u_{i,j+1}$ .

$$x_i = ih \quad y \quad t_j = jk \quad (i, j = 0, 1, 2, \dots)$$

Si tomamos  $r = \frac{\delta t}{(\delta x)^2} = \frac{k}{h^2}$ , podemos escribir la ecuación 2.15 como

$$u_{i,j+1} = ru_{i-1,j} + (1 - 2r)u_{i,j} + ru_{i+1,j} \quad (2.16)$$

La ecuación 2.16, nos será útil para encontrar la temperatura desconocida  $u_{i,j+1}$ , en el punto  $(i, j + 1)$  de la red, en términos de las temperaturas conocidas a lo largo de la  $j$ -ésima fila del tiempo en la red, fig. 2.2.

Entonces vemos que podemos calcular cada una de las temperaturas desconocidas en el tiempo  $t = k$ , (es decir, en la primer fila con  $j = 1$ ), con ayuda de las condiciones iniciales ( fila cero con  $j = 0$ ) y de frontera (consideradas la columna 0 y última, con  $i = 0$  y  $i = N$ , si el eje espacial se dividiera en  $N$

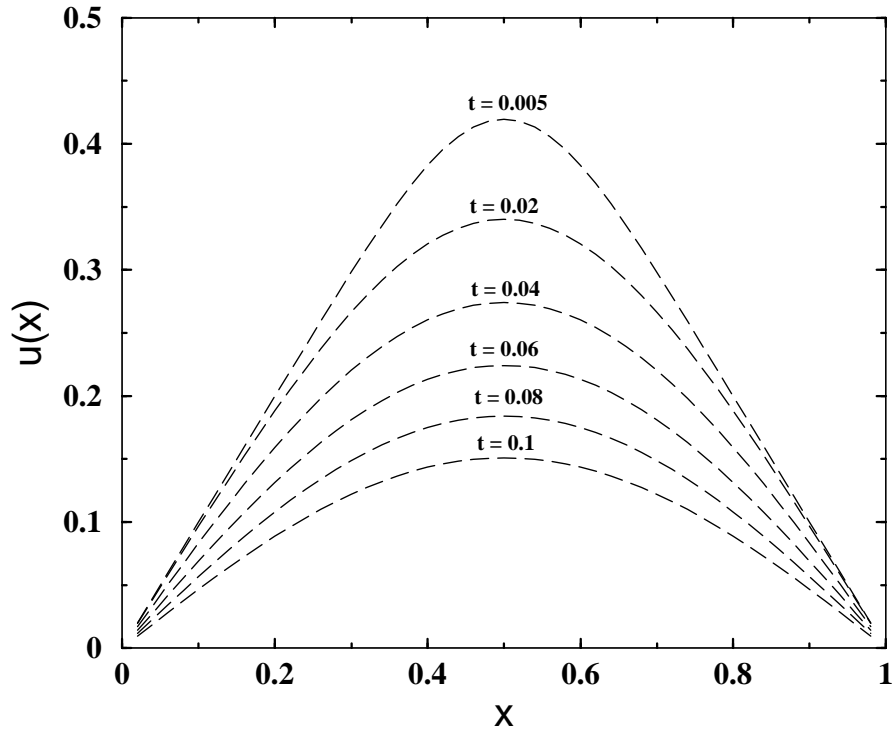


Figura 2.3:  $u(x,t)$  con el método *Explícito* para varios tiempos. Condiciones iniciales:  $u(x_i) = x_i$  para  $x_i < 0.5$  y  $u(x_i) = 1 - x_i$  para  $0.5 \leq x_i$ , con  $N = 50$ ,  $h = 0,02$ ,  $k \approx 1e-3$  y  $r = 0.25$ .

particiones) a lo largo de  $t = 0$ . De la misma manera podemos encontrar cada una de las temperaturas en el tiempo  $t = 2k$ , con ayuda de las ya conocidas temperaturas a lo largo de  $t = k$ , de  $k = 1, \dots, N-1$ .

En general, podemos encontrar los valores de las temperaturas desconocidas en cierto tiempo  $t_n$ , en función de los valores de temperaturas conocidas en un tiempo anterior a  $t_n$ , es decir,  $t_{n-1}$ . A esta forma de encontrar la solución de una ecuación diferencial en cada punto de la red se le conoce como *Método Explícito*.



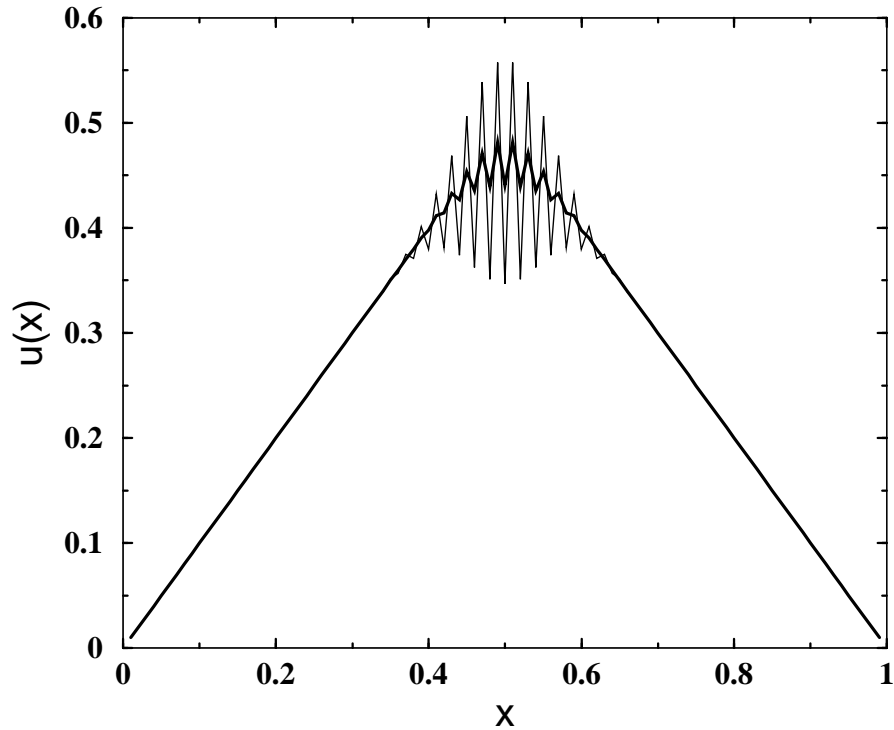


Figura 2.4: Gráfica de  $u(x, t)$  con el Método Explícito para los tiempos  $t = 0,001$  (línea gruesa) y  $t = 0,0016$  (línea delgada). Con las condiciones iniciales  $u(x_i) = x_i$  para  $x_i < 0.5$  y  $u(x_i) = 1 - x_i$  para  $0.5 \leq x_i$ , con  $N = 100$ ,  $h = 0.01$ ,  $k \approx 5.5e-05$  y  $r = 0.55$ .

A pesar de que es más sencillo utilizar el Método Explícito, éste tiene una seria desventaja, pues es inestable (fig. 2.4) a partir de que  $r$  toma valores mayores que  $\frac{1}{2}$ , es decir, el método solo es válido para  $0 < \frac{k}{h^2} \leq \frac{1}{2}$  (fig. 2.3).

Es por eso que Crank y Nicolson [3] proponen un método que fuera válido para cualquier valor finito de  $r$  (es decir, que el método sea estable), este es el Método Implícito.

## 2.2. Método Implícito

Crank y Nicolson consideraron que la ecuación diferencial parcial  $\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}$  se satisface en el punto medio  $ih, (j + \frac{1}{2})k$  y reemplazando el promedio de las aproximaciones de diferencias finitas en el  $j$  y  $(j+1)$  niveles del tiempo para la parte derecha de la ecuación parcial. Entonces la aproximación de la ecuación

$$\left(\frac{\partial U}{\partial t}\right)_{i,j+\frac{1}{2}} = \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j+\frac{1}{2}} \quad (2.17)$$

está dada por,

$$\frac{u_{i,j+1} - u_{i,j}}{k} = \frac{1}{2} \left\{ \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{h^2} + \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \right\}$$

sustituyendo  $r = \frac{k}{h^2}$  obtenemos la siguiente ecuación

$$-ru_{i-1,j+1} + (2 + 2r)u_{i,j+1} - ru_{i+1,j+1} = ru_{i-1,j} + (2 - 2r)u_{i,j} + u_{i-1,j} \quad (2.18)$$

Podemos observar que en la ecuación 2.18, el lado izquierdo contiene tres términos de  $u$  desconocidos en  $t = (j + 1)k$ , y el lado derecho tres términos de  $u$  conocidos, fig. 2.5

Si para cada tiempo  $t_j$ , el número de puntos internos de la malla es igual a  $N$ , entonces para  $j = 0$  e  $i = 1, 2, 3, \dots, N$ , la ecuación 2.18 nos dará  $N$  ecuaciones simultaneas para los  $N$  valores desconocidos a lo largo del primer tiempo, en función de los valores iniciales y de frontera ya conocidos. De igual manera, para  $j = 1$  de la ecuación 2.18 se obtendrán  $N$  valores desconocidos para  $u$  a lo largo de la segunda fila del tiempo, en términos de los valores de  $u$  ya calculados a lo largo de la primera fila del tiempo, y así sucesivamente.

Entonces, el Método Implícito consiste en resolver un sistema de ecuaciones para cada tiempo  $t_j$  con  $j = 0, 1, 2, \dots$ , para encontrar los valores de  $u$  en

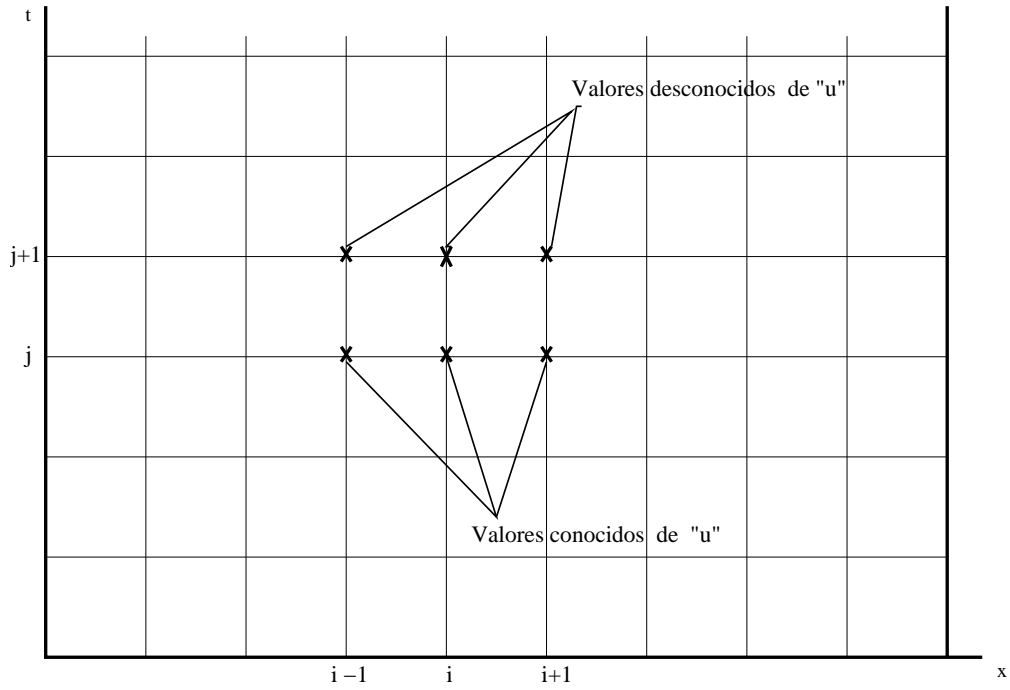


Figura 2.5: Tres valores de  $u$  conocidos en las posiciones  $x_{i-1}$ ,  $x_i$  y  $x_{i+1}$  en el tiempo  $t_j$ , y tres desconocidos en el tiempo  $t_{j+1}$  para las mismas posiciones.

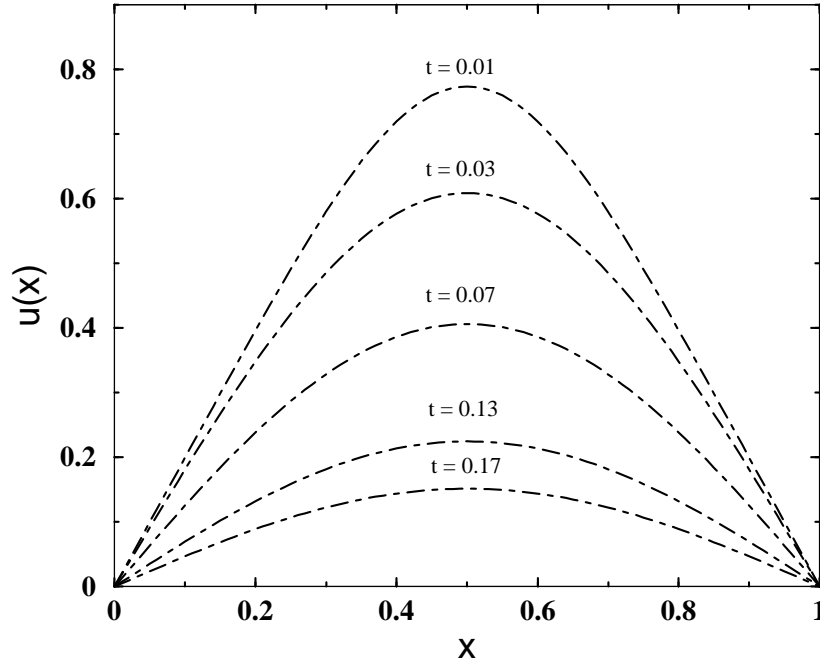


Figura 2.6: Gráfica de  $u(x)$  para varios tiempos utilizando el Método Implícito. Condiciones iniciales  $u(x_i) = x_i$  para  $x_i < 0.5$  y  $u(x_i) = 1 - x_i$  para  $0.5 \leq x_i$  con  $N=100$ ,  $h=0.01$ ,  $k \approx 1e-3$  y  $r=1$ .

cada punto de la red.

Existen varios métodos para resolver el sistema de ecuaciones que se genera en cada tiempo  $t_j$  de acuerdo con la ecuación 2.18, uno de éstos es el *Método de Eliminación de Gauss* (apéndice A) y es el que se utilizará en este trabajo.

Para saber si el método que utilizamos para encontrar la solución de la ecuación diferencial es el apropiado tenemos que verificar que dicho método cumple con las siguientes condiciones: *Consistencia*, *Convergencia* y *Estabilidad*

## 2.3. Consistencia

Sea  $L(U) = 0$  que represente la ecuación diferencial parcial dependiente de  $x$  y  $t$ , con solución exacta  $U$ , y  $F(u) = 0$  representa la ecuación de aproximación de diferencias finitas dependiente de  $x$  y  $t$ , con solución exacta  $u$ .

Definiendo el error de truncamiento  $T_{i,j}(v)$  de la siguiente forma:

$$T_{i,j}(v) = F_{i,j}(v) - L_{i,j}(v) \quad (2.19)$$

donde  $v$  es una función continua en  $x$  y  $t$ , la cual tiene un número suficiente de derivadas continuas para permitir a  $L(v)$  ser evaluada en los puntos  $(ih, jk)$ .

Si  $T_{i,j}(v) \rightarrow 0$ , cuando  $k \rightarrow 0$ ,  $h \rightarrow 0$ , se dice que la ecuación de diferencias finitas es consistente o comparable con la ecuación diferencial parcial.

La mayoría de los autores ponen  $v = U$ , de esta forma la ecuación 2.19 se reduce a

$$T_{i,j}(U) = F_{i,j}(U) \quad (2.20)$$

entonces, el error de truncamiento coincide con el error local de truncamiento  $f_{i,j}(U)$ . Así pues, la ecuación de diferencias es consistente si el valor del límite del error local de truncamiento es cero cuando  $h \rightarrow 0$ ,  $k \rightarrow 0$ .

## 2.4. Convergencia

Sea  $U$  la solución exacta de la ecuación diferencial parcial respecto a las variables independientes  $x$  y  $t$ , y sea  $u$  la solución exacta de la ecuación de diferencias finitas. Se dice que la ecuación de diferencias finitas es conver-

gente cuando  $u \rightarrow U$  en un punto fijo a lo largo de un nivel de  $t$  fijo cuando  $\delta x, \delta t \rightarrow 0$

El teorema de Lax dice que basta con que se cumpla la estabilidad para asegurar que se cumple la convergencia.

## 2.5. Estabilidad

Para un problema lineal con valores iniciales y de frontera, Lax y Richtmyer relacionan la estabilidad y convergencia vía el *El Teorema Equivalente de Lax* [3], definiendo estabilidad, en términos de las soluciones acotadas de las ecuaciones de diferencias finitas para un nivel de tiempo  $t$  fijo cuando  $k \rightarrow 0$ , es decir, cuando  $j \rightarrow \infty$ , esto asumiendo que  $\delta x = h$  está relacionado con  $k$  de tal manera que  $h \rightarrow 0$  cuando  $k \rightarrow 0$ .

Asumiendo que el vector solución  $u_{j+1} = [u_{1,j+1}, u_{2,j+1}, \dots, u_{N-1,j+1}]$  de la ecuación de diferencias finitas en el nivel de tiempo  $(j + 1)$  está relacionado con el vector de solución  $u_j$  en el nivel de tiempo  $j$ -ésimo por medio de la ecuación

$$\mathbf{u}_{j+1} = \mathbf{A}\mathbf{u}_j + \mathbf{b}_j \quad (2.21)$$

donde  $\mathbf{b}_j$  es el vector columna de los valores de frontera conocidos, y  $\mathbf{A}$  la matriz de  $(N - 1) \times (N - 1)$  de coeficientes conocidos. Entonces, mostraremos que una consecuencia de esta definición de estabilidad es que una norma de la matriz  $\mathbf{A}$  compatible con una norma de un vector  $\mathbf{u}$  satisface

$$\|\mathbf{A}\| \leq 1$$

cuando la solución de la ecuación de diferencias no diverge cuando  $t$  aumenta,  
o

$$\|\mathbf{A}\| \leq 1 + O(k) \quad (2.22)$$

cuando la solución de la ecuación de diferencias diverge cuando  $t$  aumenta.

El método matricial de análisis muestra que las ecuaciones son estables si el módulo más grande de los valores propios de la matriz  $\mathbf{A}$ , es decir, el radio espectral  $\rho(\mathbf{A})$  de  $\mathbf{A}$  satisface

$$\rho(\mathbf{A}) \leq 1,$$

cuando la solución de la ecuación de diferencias no diverge cuando  $t$  aumenta,  
o

$$\rho(\mathbf{A}) \leq 1 + O(k)$$

cuando la solución de la ecuación de diferencias diverge cuando  $t$  aumenta.

Aunque esta condición asegura que la solución está acotada, esto no garantiza la convergencia a menos que los valores propios de  $\mathbf{A}$  estén restringidos a satisfacer  $\rho(\mathbf{A}) \leq \|\mathbf{A}\| \leq 1$  cuando  $N \rightarrow \infty$ .

Sea  $\lambda_i$  un valor propio de la matriz  $\mathbf{A}$  de  $(n \times n)$  y  $\vec{x}_i$  el correspondiente vector propio. Entonces

$$\mathbf{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i$$

Obteniendo la norma de la ecuación anterior tenemos que,

$$\|\mathbf{A}\mathbf{x}_i\| = \|\lambda_i\mathbf{x}_i\| = |\lambda_i|\|\mathbf{x}_i\|.$$

Para una matriz compatible, y normas de vectores (apéndice A) llegamos a

$$|\lambda_i| \|\mathbf{x}_i\| = \|\mathbf{A}\mathbf{x}_i\| \leq \|\mathbf{A}\| \|\mathbf{x}_i\|.$$

Por lo tanto,

$$|\lambda_i| \leq \|\mathbf{A}\|, \quad i = 1, \dots, n \quad (2.23)$$

Entonces, de la ecuación 2.23 se tiene que

$$\rho(\mathbf{A}) \leq \|\mathbf{A}\|. \quad (2.24)$$

Es por eso que,

$$\|\mathbf{A}\| \leq 1$$

es condición necesaria y suficiente para que se cumpla la estabilidad.

## 2.6. Teorema de Lax

Dado un problema lineal de valores iniciales apropiado y una aproximación lineal de diferencias finitas la cual satisfaga la condición de **Consistencia**, **Estabilidad** es la condición necesaria y suficiente para que se cumpla la **Convergencia** [3].

Este capítulo se resume en la figura 2.7, donde al tener una ecuación diferencial parcial, se hace una discretización para llegar a un sistema de ecuaciones algebraicas, dicho sistema debe ser consistente con la ecuación diferencial parcial, y además de ser estable con la solución numérica, la cual converge a la solución analítica.



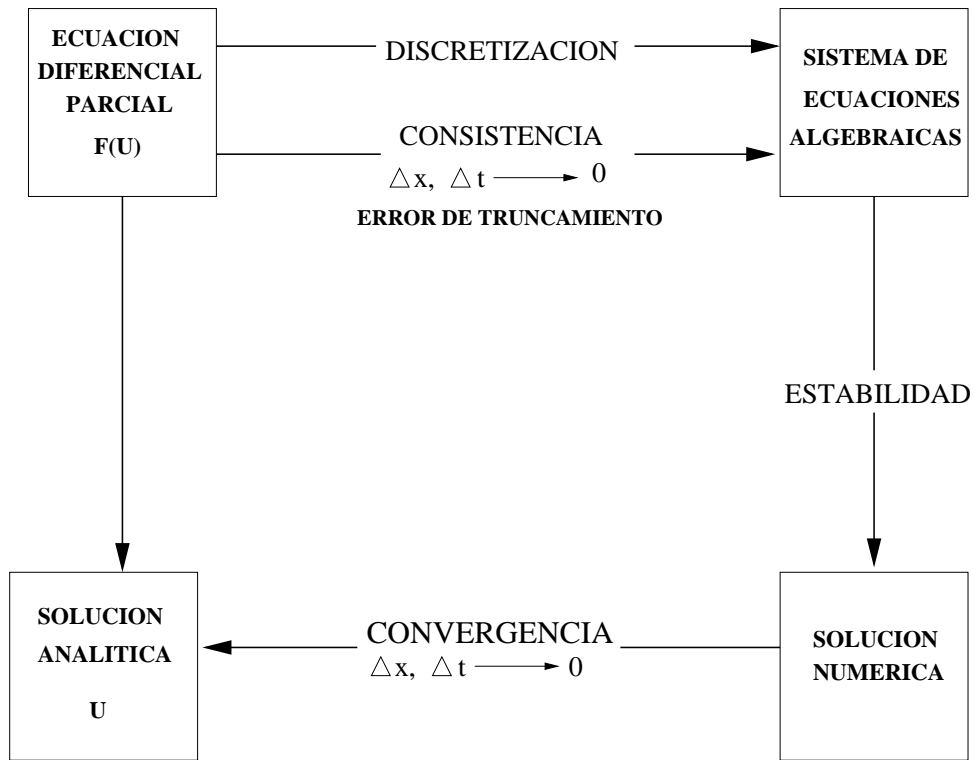


Figura 2.7:

# Capítulo 3

## Métodos Espectrales

Los métodos espectrales involucran la búsqueda de soluciones de una ecuación diferencial o de un sistema de ecuaciones diferenciales, en términos de funciones suaves. Estos métodos se empezaron a utilizar ampliamente desde hace unos 30 años, y son actualmente, una alternativa al método de diferencias finitas, y del elemento finito. Los métodos espectrales han probado ser en extremo útiles en dinámica de fluidos, donde ahora se usan códigos espectrales hidrodinámicos para el estudio de la turbulencia, inestabilidades hidrodinámicas, predicción del clima, y en el movimiento de las corrientes marítimas. Pero su uso no se ha limitado a la dinámica de los fluidos. En años recientes, se han implementado códigos para realizar simulaciones numéricas en conducción de calor, mecánica cuántica, y en el terreno de la astrofísica y la relatividad general, en particular en el estudio de los hoyos negros.

En este capítulo, vamos a dar una exposición general de la formulación de estos métodos. No pretendemos dar una presentación formal y rigurosa de los mismos, pero sí una visión panorámica de su formulación, ilustrando su uso con ejemplos simples primero, y después en los siguientes capítulos, con ejemplos más complejos.

### 3.1. Formulación

Los métodos espectrales se usan para encontrar soluciones a problemas que tienen la forma

$$\frac{\partial u(x, t)}{\partial t} = \mathcal{O}(x, t)u(x, t) \quad x \in \mathcal{D} \quad (3.1)$$

$$B(x)u(x, t) = 0 \quad x \in \partial\mathcal{D} \quad (3.2)$$

$$u(x, t = 0) = g(x) \quad x \in \mathcal{D} \quad (3.3)$$

siendo  $\mathcal{D}$  el dominio espacial del problema, y  $\partial\mathcal{D}$  su frontera,  $\mathcal{O}(x, t)$  es un operador diferencial, y  $B$  representa las condiciones de frontera. La idea básica de los métodos espectrales es la búsqueda de una solución aproximada al problema 3.1-3.3 de la forma

$$u_N(x, t) = \sum_{n=0}^N a_n(t)\Phi_n(x) \quad (3.4)$$

donde  $\Phi_n(x)$  son un conjunto dado de funciones ortogonales conocidas. Así que aquí, las incógnitas son los coeficientes  $a_n$ . Una vez obtenidos estos coeficientes, la solución ya estaría completa, y no se tendría ninguna discretización espacial, es decir, la solución se tendría en *todo el dominio espacial*. Esta característica contrasta con el método de diferencias finitas, donde la solución se da sólo en los puntos de la malla.

Cómo elegir el conjunto ortogonal utilizado en la expansión 3.4, depende mucho de las condiciones de frontera. Para funciones de frontera periódicas, la elección natural de la base de funciones es  $\Phi_k(x) = \exp(ikx)$ . Para condiciones de frontera no periódicas, hay una gran variedad de posibles elecciones, aunque en la práctica, las más usadas son las funciones, seno, coseno o los polinomios de Chebyshev. Cuando hay simetría par o impar, se usan las

funciones coseno o seno respectivamente. Pero en situaciones mas generales, los polinomios de Chebyshev son una mejor opción. En su enciclopédico libro “Chebyshev and Fourier Spectral Methods” Boyd escribió un dicho proverbial sobre la elección de la base de funciones:

- (i) Cuando esté en duda, use los polinomios de Chebyshev, a menos que la solución sea espacialmente periódica, en cuyo caso es mejor una serie de Fourier ordinaria.
- (ii) A menos que, esté usted realmente seguro que otro conjunto de funciones base es mejor, use polinomios de Chebyshev.
- (iii) A menos que esté usted realmente, pero realmente seguro que otro conjunto de funciones base es mejor, use polinomios de Chebyshev.

Los polinomios de Chebyshev, no solamente se amoldan muy bien a condiciones de frontera arbitrarios, que en diferencias finitas sería difícil implementar, sino que también, por ser una serie de Fourier “disfrazada”, se pueden usar los poderosos algoritmos de las transformadas de Fourier rápidas. Con otras bases no existe esta posibilidad, pues no cuentan con dichas transformadas.

El cómo determinar los coeficientes de la expansión 3.4 da lugar a los tres métodos espectrales más utilizados: Galerkin, Colocación, y Tau, que veremos más adelante, pero primero haremos una breve exposición de los polinomios de Chebyshev y sus propiedades mas relevantes para este trabajo.

## 3.2. Polinomios de Chebyshev

El polinomio de Chebyshev  $T_k(x)$  de grado  $k$  está definido (para  $x \in [-1, 1]$ ) por

$$T_k(x) = \cos(k \cos^{-1} x), \quad k = 0, 1, 2, \dots, \quad (3.5)$$

por lo tanto,  $-1 \leq T_k \leq 1$ . Fijando  $x = \cos z$ , tenemos

$$T_k = \cos kz \quad (3.6)$$

por lo que es fácil deducir los primeros polinomios de Chebyshev

$$T_0 = 1, \quad T_1 = \cos z = x, \quad T_2 = \cos 2z = 2 \cos^2 z - 1 = 2x^2 - 1, \dots$$

Más generalmente, para la fórmula de Moivre, tenemos

$$\cos kz = \mathcal{R}e \left\{ (\cos z + i \sin z)^k \right\}$$

y entonces, aplicando la fórmula binomial, podemos expresar  $T_k$  como

$$T_k = \frac{k}{2} \sum_{m=0}^{\lfloor k/2 \rfloor} (-1)^m \frac{(k-m-1)!}{m!(k-2m)!} (2x)^{k-2m} \quad (3.7)$$

donde  $[\Omega]$  denota la parte entera de  $\Omega$ .

De la identidad trigonométrica

$$\cos(k+1)z + \cos(k-1)z = 2 \cos z \cos(kz)$$

se deduce la relación de recurrencia

$$T_{k+1} - 2xT_k + T_{k-1} = 0, \quad k \geq 1, \quad (3.8)$$

la cual nos permite, en particular, deducir la expresión de los polinomios  $T_k$ ,  $k \geq 2$ , del conocimiento de  $T_0$  y  $T_1$ . La gráfica de los primeros polinomios se muestra en la figura 3.1.

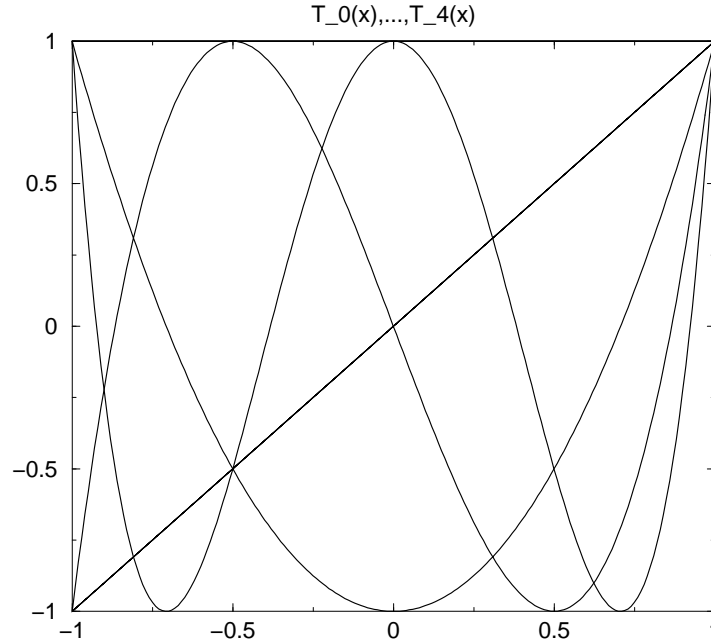


Figura 3.1: Esta figura muestra los primeros cuatro polinomios de Chebyshev

La expresión 3.7 puede ser útil en circunstancias especiales pero la representación 3.6 es generalmente usada en computación así como en estudios teóricos.

Ahora enumeramos algunas características para la comprensión y aplicación de los polinomios de Chebyshev a la solución de ecuaciones diferenciales ordinarias o parciales.

Los valores de  $T_k$  y sus primeras derivadas  $T'_k$  en  $x = \pm 1$  están dadas por

$$T_k(\pm 1) = (\pm 1)^k, \quad T'_k(\pm 1) = (\pm 1)^{k+1} k^2. \quad (3.9)$$

El conocimiento de estos valores puede ser de interés cuando se escriben las condiciones de frontera. Es importante notar que

$$T_k(-x) = (-1)^k T_k(x) \quad (3.10)$$

ésta es la paridad del polinomio que es igual a su grado  $k$ . El polinomio  $T_k$  se anulan en los puntos  $x_i$  (puntos de Gauss) definidos por

$$x_i = \cos \left( i + \frac{1}{2} \right) \frac{\pi}{k}, \quad i = 0, \dots, k-1 \quad (3.11)$$

y alcanza sus valores extremos  $\pm 1$  en los puntos  $\bar{x}_i$  (puntos de Gauss-Lobatto) definidos por

$$\bar{x}_i = \cos \frac{\pi i}{k}, \quad i = 0, \dots, k. \quad (3.12)$$

Una relación de recurrencia de la derivada  $T_k$  pueden obtenerse fácilmente. Primero, la diferenciación de  $T_k$  nos da

$$T_k' = \frac{d}{dz}(\cos kz) \frac{dz}{dx} = k \frac{\sin kz}{\sin z},$$

donde usamos la representación (3.6). Entonces, aplicando fórmulas trigonométricas, obtenemos la relación

$$\frac{T_{k+1}'}{k+1} - \frac{T_{k-1}'}{k-1} = 2T_k \quad (3.13)$$

válida para  $k > 1$ . Una fórmula similar para la  $p$ -ésima derivada es obtenida diferenciando sucesivamente a (3.13).

Los polinomios de Chebyshev son ortogonales en  $[-1, 1]$  con el peso

$$w = (1 - x^2)^{-1/2} \quad (3.14)$$

y con el producto escalar definido por,

$$(u, v)_w = \int_{-1}^1 uvw dx. \quad (3.15)$$

De este modo la condición de ortogonalidad para los polinomios de Chebyshev es

$$(T_k, T_l)_w = \int_{-1}^1 T_k T_l w dx = \frac{\pi}{2} c_k \delta_{k,l}. \quad (3.16)$$

donde  $\delta_{k,l}$  es la delta de Kronecker y  $c_k$  es definido por

$$c_k = \begin{cases} 2 & \text{si } k = 0, \\ 1 & \text{si } k \geq 1. \end{cases}$$

### 3.2.1. Diferenciación

La expresión de derivadas en la base de Chebyshev es más complicada que en la de Fourier. De hecho, la expresión de la derivada de  $T_k(x)$  implica todos los polinomios de paridad opuesta y grado más bajo mientras la derivada de  $e^{ikx}$  es simplemente  $ike^{ikx}$ .

De la relación de recurrencia 3.13, se obtiene

$$T'_k = 2k \sum_{n=0}^K \frac{1}{c_{k-1-2n}} T_{k-1-2n}(x), \quad (3.17)$$

donde  $K = [(k-1)/2]$ . Por lo tanto, considerando la primera derivada de la función  $u_N(x)$

$$u'_N(x) = \sum_{k=0}^N \hat{u}_k T'_k(x) = \sum_{k=0}^N \hat{u}_k^{(1)} T_k(x). \quad (3.18)$$

Utilizando la ec. 3.13 en la ecuación anterior, llegamos a la siguiente expresión

$$k \hat{u}_k = c_{k-1} \hat{u}_{k-1}^{(1)} - \hat{u}_{k+1}^{(1)} \quad k \geq 1; \quad (3.19)$$

de la ecuación anterior se deducen los coeficientes  $\hat{u}_k^{(1)}$



$$\hat{u}_k^{(1)} = \frac{2}{c_k} \sum_{\substack{p=k+1 \\ (p+k)\text{ impar}}}^N p \hat{u}_p, \quad k = 0, \dots, N-1 \quad (3.20)$$

y  $\hat{u}_N^{(1)} = 0$ . Esto se puede escribir en forma matricial como

$$\hat{U}^{(1)} = \hat{\mathcal{D}}\hat{U}, \quad (3.21)$$

donde  $\hat{U} = (\hat{u}_0, \dots, \hat{u}_N)^T$ ,  $\hat{U}^{(1)} = (\hat{u}_0^{(1)}, \dots, \hat{u}_N^{(1)})^T$  y  $\hat{\mathcal{D}}$  es una matriz triangular superior y sus entradas se deducen de la ecuación 3.20.

La expansión de la segunda derivada de la función  $u_N$  es

$$u_N'' = \sum_{k=0}^N \hat{u}_k^{(2)} T_k(x) \quad (3.22)$$

con

$$\hat{u}_k^{(2)} = \frac{1}{c_k} \sum_{\substack{p=k+2 \\ (p+k)\text{ par}}}^N p(p^2 - k^2) \hat{u}_p, \quad k = 0, \dots, N-2 \quad (3.23)$$

y  $\hat{u}_{N-1}^{(2)} = \hat{u}_N^{(2)} = 0$ . Expresando 3.23 en forma matricial

$$\hat{U}^{(2)} = \hat{\mathcal{D}}^2 \hat{U}, \quad (3.24)$$

donde  $\hat{U}^{(2)} = (\hat{u}_0^{(2)}, \dots, \hat{u}_N^{(2)})^T$ .

Las expresiones analíticas 3.20 y 3.22 son de interés cada vez que los coeficientes de la expansión de las derivadas se involucran en cálculos algebraicos. Por otra parte, si solo se necesitan los valores numéricos de los coeficientes, se pueden calcular del producto de matrices y vectores 3.21 y 3.24 o de las

fórmulas de recurrencia que se deducen de 3.13. En general la fórmula de recurrencia para los coeficientes  $\hat{u}_k^{(p)}$  para la  $p$ -ésima derivada es obtenida por la diferenciación sucesiva,

$$c_{k-1}\hat{u}_{k-1}^{(p)} = \hat{u}_{k+1}^{(p)} + 2k\hat{u}_k^{(p-1)}, \quad k \geq 1, \quad (3.25)$$

para la primera derivada

$$\hat{u}_N^{(1)} = 0, \quad \hat{u}_{N-1}^{(1)} = 2N\hat{u}_N, \quad (3.26)$$

y, para la segunda derivada

$$\hat{u}_N^{(2)} = \hat{u}_{N-1}^{(2)} = 0, \quad \hat{u}_{N-2}^{(2)} = 2(N-1)\hat{u}_{N-1}^{(1)} = 2N(N-1)\hat{u}_N. \quad (3.27)$$

### 3.3. Aproximación Galerkin

Cuando desarrollamos una función bien comportada (continua y derivable en su dominio de definición) en términos de una serie truncada de Fourier o de Chebyshev, los coeficientes de la expansión decrecen muy rápidamente. Para ilustrar este hecho, desarrollamos algunas funciones  $f(x)$ , como en (3.4), en términos de polinomios de Chebyshev, y mostramos cómo decaen los coeficientes y cómo se aproxima la expansión a la función  $f(x)$ . Este comportamiento es importante en el método espectral Galerkin, con el que se pretende hallar soluciones del sistema (3.1-3.3). Cuando sustituimos la expansión (3.4) en (3.1), se tiene que el así llamado *Residuo*  $R(x, t, a_n)$ , evidentemente no se anula

$$R = \frac{\partial u_N(x, t)}{\partial t} - \mathcal{O}(x, t)u_N(x, t) \neq 0 \quad (3.28)$$

Cuadro 3.1: En la tabla se muestran los valores de los coeficientes de la expansión en polinomios de Chebyshev, para la función  $\cos(2x^2 + 2x^3)$ , desde  $n = 1, \dots, N$  con  $N = 21$ .

N	Coeficiente	N	Coeficiente
0	4.1343e-01	11	1.5699e-16
1	1.5000e+00	12	2.2626e-05
2	7.4057e-01	13	1.5699e-16
3	5.0000e-01	14	2.5283e-06
4	1.2416e-01	15	3.2678e-16
5	1.7566e-16	16	1.0181e-07
6	3.2923e-12	17	4.1228e-16
7	2.3961e-16	18	8.8341e-09
8	2.6762e-03	19	1.2469e-16
9	2.5708e-16	20	2.8426e-10
10	4.2032e-04	21	13.3719e-16

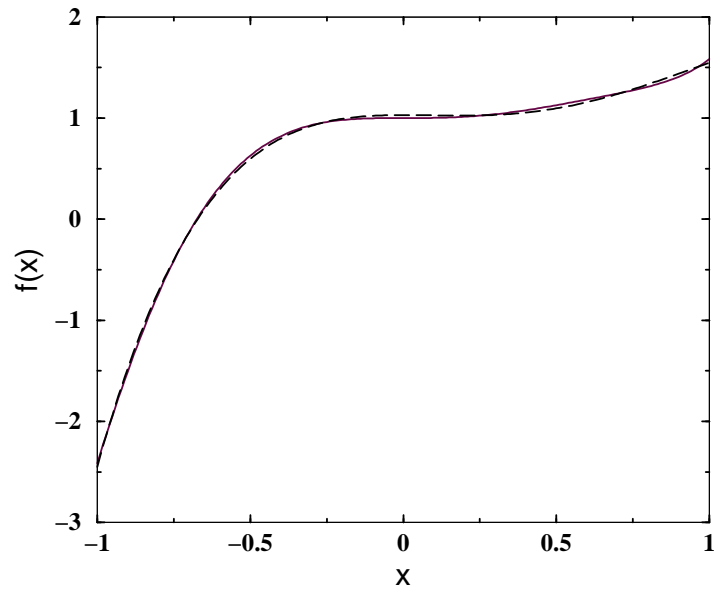


Figura 3.2: Comparación de la función  $\cos(2x^2) + 2x^3$  (línea sólida) y la función aproximada con una expansión en términos de polinomios de Chebyshev (línea punteada) con  $N=4$ . Se requieren realmente muy pocos términos para representar a la función apropiadamente. Al ir aumentando el valor de  $N$ , la solución numérica se va aproximando cada vez más a la solución exacta.

Esta función residual puede expandirse, como cualquier función, en términos de un conjunto ortogonal,

$$R(x, t, a_0, \dots, a_N) = \sum_{k=0}^{\infty} r_k(a_0, \dots, a_N) \Phi_k(x) \quad (3.29)$$

donde, debido a la ortogonalidad de la base de funciones  $\Phi_k(x)$ , los coeficientes  $r_k$  se determinan por el producto interno

$$r_k = (R, \Phi_k) = \int R \Phi_k(x) dx. \quad (3.30)$$

En el método espectral Galerkin, se pide que el residuo sea pequeño, en el sentido de que los primeros  $(N + 1)$  términos de la serie espectral (3.29) se anulen. Las series de Fourier y de Chebyshev de funciones suaves decrecen exponencialmente con  $k$ , de este modo, todos los  $r_k$  para  $k > N$  presumiblemente serán muy pequeños para  $N$  grandes. Así que, forzar que los primeros  $r_k$  sean cero, minimiza a  $R(x)$ . En el límite en el que  $N \rightarrow \infty$   $R(x)$  debe  $\rightarrow \rightarrow 0$ , y por lo tanto, la aproximación, debe converger muy rápido a la solución exacta. Todo esto entonces se traduce a:

*para minimizar el residuo se pide que  $(R, \Phi_n) = 0$  para  $n = 0, 1, 2, \dots, N$ .*

En este método, se pide que la base (o un arreglo de ésta) cumplan de entrada las condiciones de frontera. El ejemplo más sencillo, cuando se tienen condiciones de frontera periódicas, es elegir una base de Fourier.

Daremos a continuación un ejemplo muy simple. Considere la siguiente ecuación hiperbólica y lineal (ecuación de advección)

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \quad (3.31)$$

con dominio en el intervalo  $x \in (0, 2\pi]$ , condiciones periódicas, y condición inicial  $u(x, 0) = g(x)$ . La condición de minimización del residuo es en este

caso

$$\frac{1}{2\pi} \int_0^{2\pi} \left[ \left( \frac{\partial}{\partial t} - \frac{\partial}{\partial x} \right) \sum_{k=0}^N a_k(t) e^{ikx} \right] e^{-inx} dx = 0. \quad (3.32)$$

Haciendo la derivada espacial, esta relación se convierte en

$$\frac{1}{2\pi} \int_0^{2\pi} \left[ \sum_{k=0}^N \left( \frac{da_k}{dt} - ik a_k \right) e^{ikx} \right] e^{-inx} dx = 0. \quad (3.33)$$

Como la base es ortogonal, la integral en  $x$  es proporcional a una delta de Kronecker, dando por resultado un conjunto de ecuaciones diferenciales para los coeficientes  $a_k$ , a saber

$$\frac{da_k}{dt} - ik a_k = 0 \quad (3.34)$$

La condición inicial  $u(x, 0) = g(x)$ , nos da las condiciones iniciales para este conjunto de ecuaciones, pues al integrar la ecuación (3.4) previamente evaluada en cero, se tiene

$$a_k(0) = \int_0^{2\pi} g(x) \Phi_k(x) dx \quad (3.35)$$

El problema consiste en evaluar  $a_k(0)$  y resolver las ecuaciones diferenciales (3.34) para encontrar las  $a_k(t)$  a todo tiempo. Para este ejemplo, podemos verificar la convergencia de la aproximación, usando por ejemplo, la condición inicial  $u(x, 0) = \sin(\pi \cos x)$ . La solución exacta  $u(x, t) = \sin(\pi \cos(x + t))$ , tiene expansión de Fourier cuyos coeficientes están dados por

$$a_k(t) = \sin\left(\frac{k\pi}{2}\right) J_k(\pi) e^{ikt} \quad (3.36)$$

donde  $J_k(t)$  son las funciones de Bessel de orden  $k$ , cuyas propiedades asintóticas implican  $k^m a_k(t) \rightarrow 0$  cuando  $k \rightarrow \infty$  para todo natural  $m$ . Esto

nos dice que la serie truncada de Fourier converge mas rápido que cualquier potencia finita de  $1/N$ , propiedad conocida como convergencia exponencial.

### 3.4. Método de Tau

El método de Tau es una versión modificada del método de Galerkin, que resuelve un sistema un poco diferente (al de Galerkin) para poder incorporar eficientemente las condiciones de frontera en la solución. Nuevamente, se toma una base ortogonal de funciones  $\Phi_n(x)$ , en este método se utiliza una solución aproximada de la forma

$$u_N(x, t) = \sum_{n=0}^{N+K} a_n(t) \Phi_n(x) \quad (3.37)$$

donde  $N$  es el número de términos de la expansión y  $K$  es el número de condiciones de frontera del problema. En la sección de ejemplos, utilizaremos este método para resolver la ecuación de Burgers, que es un modelo reducido de la dinámica de fluidos, y representa ondas que chocan.

### 3.5. Método de Colocación

En este método se asume que la solución aproximada  $u_N$  satisface la ecuación diferencial (3.1) en algunos puntos  $x_1, x_2, \dots, x_N$ , llamados *puntos de colocación* del dominio respectivo. Es decir, se sustituye la expansión (3.4) en el sistema  $\mathcal{O}(u_N(x_i)) = f(x_i)$ , que junto con las condiciones de frontera, forman un sistema lineal de ecuaciones  $(N+1) \times (N+1)$  para los coeficientes  $a_n$ .

Estos son los métodos espectrales más ampliamente usados. Las expansiones se usan en el dominio espacial. Para la evolución temporal, generalmente se usan las diferencias finitas, y entonces se tienen también los conceptos de consistencia y estabilidad, que implican como ya se vió, la convergencia.

La presentación formal de estos conceptos va mas allá del propósito de este trabajo. Esto puede verse en el capítulo 6 del recién editado libro ([7]).

En el capítulo de *Ejemplos* vamos a ilustrar este método al resolver la ecuación de Poisson, y lo compararemos a con las soluciones obtenidas con diferencias finitas.



# Capítulo 4

## Ejemplos

### 4.1. Diferencias Finitas

#### 4.1.1. Ecuación de Onda

A continuación daremos una pequeña explicación de cómo podemos expresar la Ecuación de Onda (dependiente del segundo orden del tiempo y la posición  $x$ ) en ecuaciones más sencillas (dependientes del primer orden del tiempo y  $x$ ) para encontrar su solución. Esto será posible utilizando Diferencias Finitas en las ecuaciones de primer orden. El propósito de realizar este cambio, es debido al segundo orden en el tiempo, pues es más común resolver ecuaciones que evolucionan en primer orden del tiempo.

La Ecuación de Onda está dada por

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = 0 \quad (4.1)$$

Si aplicamos una transformación de coordenadas de la forma  $d\tilde{x} = dx - \beta dt$  y  $d\tilde{t} = dt$ , la nueva expresión para la ecuación de onda será

$$\frac{\partial^2 \phi}{\partial \tilde{x}^2} - \frac{\beta^2}{c^2} \frac{\partial^2 \phi}{\partial \tilde{x}^2} + 2 \frac{\beta}{c^2} \frac{\partial^2 \phi}{\partial \tilde{x} \partial \tilde{t}} - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial \tilde{t}^2} = 0 \quad (4.2)$$

donde  $\tilde{x}$  y  $\tilde{t}$  son las nuevas coordenadas y  $\beta$  un cambio distinto de cero, el cual representa la velocidad con la que el sistema  $(x, t)$  verá que se mueve el nuevo sistema coordenado.

Desarrollando la ec. (4.2) de la forma que sigue:

$$\frac{\partial}{\partial \tilde{x}} \left( \left(1 - \frac{\beta^2}{c^2}\right) \frac{\partial \phi}{\partial \tilde{x}} + \frac{\beta}{c^2} \frac{\partial \phi}{\partial \tilde{t}} \right) + \frac{\partial}{\partial \tilde{t}} \left( \frac{\beta}{c^2} \frac{\partial \phi}{\partial \tilde{x}} - \frac{1}{c^2} \frac{\partial \phi}{\partial \tilde{t}} \right) = 0 \quad (4.3)$$

tomando  $c = 1$  y omitiendo las tildes, de la ec (4.3) podemos introducir estas variables  $\pi = \partial_t \phi - \beta \partial_x \phi$  y  $\psi = \partial_x \phi$

El propósito es expresar la ecuación de onda 4.2 en dos ecuaciones de evolución de primer orden en el tiempo para las nuevas variables. La primera de éstas es inmediata de la ec. (4.3)

$$\partial_t \pi = \partial_x (\psi + \beta \pi) \quad (4.4)$$

Asumiendo que la Ecuación de Onda es  $C^2$ , podemos obtener una expresión para a evolución de  $\psi$  de la siguiente forma

$$\partial_t \psi = \partial_x (\partial_t \phi) \quad (4.5)$$

De la definición de  $\pi$  obtenemos

$$\partial_t \phi = \pi + \beta \psi \quad (4.6)$$

sustituyendo la ec. (4.6) en la ec.(4.5) finalmente obtenemos

$$\partial_t \psi = \partial_x (\pi + \beta \psi) \quad (4.7)$$

De esta manera obtenemos las dos ecuaciones de evolución (4.4) y (4.7) las cuales constituyen la ecuación de onda. Así pues, al integrar la ecuación (4.6) junto con las dos anteriores, encontraremos el valor buscado de la variable  $\phi$ .

Las ecuaciones (4.4), (4.6) y (4.7) se pueden generalizar de la forma:

$$\partial_t y = f(y) \quad (4.8)$$

donde  $f(y)$  es un operador espacial aplicado a la función que queremos integrar.

Aproximaremos las ecuaciones (4.4), (4.6) y (4.7) usando *Diferencias Finitas* y trataremos de encontrar los valores de  $\phi$ ,  $\psi$  y  $\pi$  en un conjunto de puntos en el espacio-tiempo. Discretizaremos el espacio-tiempo de tal forma que la coordenada temporal esté representada por  $t^n = n\Delta t$  y la coordenada espacial por  $x_i = i\Delta x$ , con  $n$  e  $i$  enteros,  $i = 1, \dots, J$  donde  $J$  es el número de los puntos de la red en la dirección espacial. Para simplificar la notación denotaremos  $f(t^n, x_i)$  por  $f_i^n$

Aplicando diferencias finitas a la ecuación (4.6) obtenemos:

$$\frac{\phi_i^n - \phi_i^{n-1}}{\Delta t} \approx \pi + \beta\psi$$

Despejando  $\phi_i^n$  de la ecuación anterior,

$$\phi_i^n = \phi_i^{n-1} + \Delta t(\pi + \beta\psi) \quad (4.9)$$

De la ecuación (4.9) podemos ver que para obtener los valores de  $\phi$  en cada punto de la red, debemos antes conocer los valores de  $\pi$  y  $\psi$  en cada punto, para encontrar dichos valores tenemos que resolver las ecuaciones (4.4) y (4.7).

Ahora apliquemos diferencias finitas al lado derecho de las ecuaciones que queremos resolver (4.49 y (4.7).

$$\partial_t \pi = \frac{\psi_{i+1}^{n-1} - \psi_{i-1}^{n-1} + \beta(\pi_{i+1}^{n-1} - \pi_{i-1}^{n-1})}{2\Delta x} \quad (4.10)$$

y

$$\partial_t \psi = \frac{\pi_{i+1}^{n-1} - \pi_{i-1}^{n-1} + \beta(\psi_{i+1}^{n-1} - \psi_{i-1}^{n-1})}{2\Delta x} \quad (4.11)$$

Para encontrar los valores de  $\pi$  y  $\psi$  resolveremos las ecuaciones (4.10) y (4.11) utilizando el método de Runge-Kutta de cuarto orden.

Los resultados numéricos se presentan en el siguiente capítulo.

### 4.1.2. Ecuación de Poisson

En esta sección expondremos la solución de la ecuación de Poisson con Diferencias Finitas usando los métodos Jacobi, Gauss-Seidel y Sobre Relajación Sucesiva “SOR”.

La ecuación de Poisson en dos dimensiones  $(x, y)$  se expresa como sigue:

$$\nabla^2 F(x, y) = f(x, y) \quad (4.12)$$

donde

$$\nabla^2 F(x, y) = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2}$$

Aplicando la ecuación 2.9 al lado izquierdo de la ecuación 4.12, obtenemos como resultado la siguiente expresión

$$\frac{F_{i+1,j} + F_{i-1,j} + F_{i,j+1} + F_{i,j-1} - 4F_{i,j}}{h^2} = f(x_i, y_j). \quad (4.13)$$

donde  $h$  es tamaño de las divisiones del eje espacial. Utilizaremos los tres métodos arriba mencionados para dar solución a la ecuación 4.13.

Jacobi, Gauss-Seidel y SOR son métodos sistemáticos iterativos útiles para resolver numerosos sistemas de ecuaciones lineales algebraicas. Estos métodos consisten en utilizar tantas iteraciones como sean necesarias para obtener la solución mejor aproximada. Se propone una solución de la ecuación (solución de prueba) la cual es utilizada para calcular la solución en la primer iteración, entonces, el valor que se obtuvo en la primer iteración es utilizado para obtener el valor de la solución en la segunda iteración y este a su vez es utilizado para obtener el valor el la tercera iteración, y así sucesivamente hasta llegar a la última iteración la cual nos proporcionará la solución más aproximada a la solución exacta acotada por alguna definición de error. El procedimiento iterativo se dice ser convergente cuando las diferencias entre la solución exacta y las aproximaciones sucesivas tienden a cero cuando el número de iteraciones aumenta.

De expresiones como 4.13 podemos obtener un sistema de ecuaciones de la siguiente forma ( como ejemplo, un sistema de 4 ecuaciones),

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 &= b_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 &= b_4 \end{aligned} \quad (4.14)$$

donde  $a_{ii} \neq 0$  para  $i = 1, \dots, 4$ , podemos escribir las ecuaciones anteriores de la siguiente manera,

$$\begin{aligned}
x_1 &= \frac{1}{a_{11}}(b_1 - a_{12}x_2 - a_{13}x_3 - a_{14}x_4) \\
x_2 &= \frac{1}{a_{22}}(b_2 - a_{21}x_1 - a_{23}x_3 - a_{24}x_4) \\
x_3 &= \frac{1}{a_{33}}(b_3 - a_{31}x_1 - a_{32}x_2 - a_{34}x_4) \\
x_4 &= \frac{1}{a_{44}}(b_4 - a_{41}x_1 - a_{42}x_2 - a_{43}x_3)
\end{aligned}
\tag{4.15}$$

### 4.1.3. Método de Jacobi

Indicaremos las iteraciones de  $x_i$  de la forma  $x_i^{(n)}$  con  $n$  el número de iteraciones. Asumiendo que la  $n$ -ésima iteración es conocida para  $i = 1, \dots, 4$ , entonces el Método de Jacobi expresa la iteración  $n + 1$  en términos de la iteración  $n$ -ésima. Las ecuaciones correspondientes a este método son:

$$\begin{aligned}
x_1^{n+1} &= \frac{1}{a_{11}}(b_1 - a_{12}x_2^n - a_{13}x_3^n - a_{14}x_4^n) \\
x_2^{n+1} &= \frac{1}{a_{22}}(b_2 - a_{21}x_1^n - a_{23}x_3^n - a_{24}x_4^n) \\
x_3^{n+1} &= \frac{1}{a_{33}}(b_3 - a_{31}x_1^n - a_{32}x_2^n - a_{34}x_4^n) \\
x_4^{n+1} &= \frac{1}{a_{44}}(b_4 - a_{41}x_1^n - a_{42}x_2^n - a_{43}x_3^n)
\end{aligned}$$

En general para el caso de  $m$  ecuaciones,

$$x_i^{n+1} = \frac{1}{a_{ii}} \left\{ b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(n)} - \sum_{j=i+1}^m a_{ij}x_j^{(n)} \right\} \quad i = 1, \dots, m \tag{4.16}$$

Aplicando la ecuación (4.16) a (4.13)

$$F_{i,j}^{n+1} = \frac{1}{4}(F_{i+1,j}^n + F_{i-1,j}^n + F_{i,j+1}^n + F_{i,j-1}^n - h^2 f_{i,j}) \quad (4.17)$$

#### 4.1.4. Método de Gauss-Seidel

En este método el valor de la iteración  $(n+1)$  será utilizado tan pronto como esté disponible. Las ecuaciones equivalentes a las ecs. (4.15) para este método son

$$\begin{aligned} x_1^{n+1} &= \frac{1}{a_{11}}(b_1 - a_{12}x_2^n - a_{13}x_3^n - a_{14}x_4^n) \\ x_2^{n+1} &= \frac{1}{a_{22}}(b_2 - a_{21}x_1^{n+1} - a_{23}x_3^n - a_{24}x_4^n) \\ x_3^{n+1} &= \frac{1}{a_{33}}(b_3 - a_{31}x_1^{n+1} - a_{32}x_2^{n+1} - a_{34}x_4^n) \\ x_4^{n+1} &= \frac{1}{a_{44}}(b_4 - a_{41}x_1^{n+1} - a_{42}x_2^{n+1} - a_{43}x_3^{n+1}) \end{aligned}$$

En el caso de  $m$  ecuaciones

$$x_i^{n+1} = \frac{1}{a_{ii}} \left\{ b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{n+1} - \sum_{j=i+1}^m a_{ij}x_j^{(n)} \right\} \quad i = 1, \dots, m \quad (4.18)$$

Aplicando la ecuación (4.18) a la ecuación (4.13) obtenemos

$$F_{i,j}^{n+1} = \frac{1}{4}(F_{i+1,j}^n + F_{i-1,j}^n + F_{i,j+1}^n + F_{i,j-1}^{n+1} - h^2 f_{i,j}) \quad (4.19)$$

#### 4.1.5. Método de Sobre Relajación Sucesiva (SOR)

Si sumamos y restamos el término  $x_i^{(n)}$  al lado derecho de la ecuación (4.18) nos queda la siguiente expresión

$$x_i^{n+1} = x_i^{(n)} + \left[ \frac{1}{a_{ii}} \left\{ b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{n+1} - \sum_{j=i}^m a_{ij} x_j^{(n)} \right\} \right] \quad (4.20)$$

La expresión dentro de los corchetes es la corrección o cambio hecho a  $x_i^{(n)}$  para  $i = 1, \dots, m$  para una iteración de Gauss-Seidel. Si las correcciones sucesivas son todas de un solo signo, sería razonable esperar que la convergencia se acelere si a cada una de las ecuaciones 4.20 se les da un término de corrección más grande que el definido en dicha ecuación.

Así es como surge la iteración sobre relajada sucesiva o SOR la cual se define por las ecuaciones (en el caso de un sistema de 4 ecuaciones)

$$\begin{aligned} x_1^{n+1} &= x_1^{(n)} + \frac{\omega}{a_{11}} \{ b_1 - a_{11}x_1^{(n)} - a_{12}x_2^{(n)} - a_{13}x_3^{(n)} - a_{14}x_4^{(n)} \} \\ x_2^{n+1} &= x_2^{(n)} + \frac{\omega}{a_{22}} \{ b_2 - a_{21}x_1^{(n)} - a_{22}x_2^{(n)} - a_{23}x_3^{(n)} - a_{24}x_4^{(n)} \} \\ x_3^{n+1} &= x_3^{(n)} + \frac{\omega}{a_{33}} \{ b_3 - a_{31}x_1^{(n)} - a_{32}x_2^{(n)} - a_{33}x_3^{(n)} - a_{34}x_4^{(n)} \} \\ x_4^{n+1} &= x_4^{(n)} + \frac{\omega}{a_{44}} \{ b_4 - a_{41}x_1^{(n)} - a_{42}x_2^{(n)} - a_{43}x_3^{(n)} - a_{44}x_4^{(n)} \}. \end{aligned}$$

Donde  $\omega$  es el parámetro de aceleración o factor de relajación, el cual yace dentro del rango  $1 < \omega < 2$ . Para el caso de un sistema de  $m$  ecuaciones la ecuación que describe al método de SOR se tiene:

$$x_i^{(n+1)} = x_i^{(n)} + \frac{\omega}{a_{ii}} \left\{ b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(n+1)} - \sum_{j=i}^m a_{ij} x_j^{(n)} \right\} \quad i = 1, \dots, m \quad (4.21)$$

Aplicando la ecuación (4.21) a la ecuación (4.13) obtenemos

$$F_{i,j}^{n+1} = (1 - \omega) F_{n,i,j} + \frac{\omega}{4} (F_{i+1,j}^n + F_{i-1,j}^n + F_{i,j+1}^n + F_{i,j-1}^{n+1} - h^2 f_{i,j}) \quad (4.22)$$



Entonces para la Ecuación de Poisson, las ecuaciones a resolver son: (4.17), (4.19) y (4.22). Los resultados numéricos de dichas ecuaciones se expondrán en el capítulo siguiente.

## 4.2. Método Espectral de Colocación

### 4.2.1. Ecuación de Poisson

Ahora expondremos la manera de dar solución a la ecuación de Poisson con el Método de Colocación Espectral.

Como ya se mencionó anteriormente, queremos encontrar una solución de la forma

$$u(x, y) = \sum_{n=0}^N \sum_{m=0}^N a_{nm} T_n(x) T_m(y) \quad (4.23)$$

Como los términos  $T_n(x)$  y  $T_m(y)$  (polinomios de Chebyshev) ya son conocidos, lo que nos resta por hacer, es encontrar los coeficientes  $a_{nm}$ , y para ello sustituiremos la expansión de la ecuación de Poisson,

$$\nabla^2 F(x, y) = f(x, y) \quad (4.24)$$

asumiendo que la ecuación se cumple para los puntos de colocación, entonces sustituimos la expansión y la evaluamos en dichos puntos obteniendo

$$\sum_{n=0}^N \sum_{m=0}^N a_{nm} \left[ \frac{d^2 T_n(x_i)}{dx^2} T_m(y_j) + T_n(x_i) \frac{d^2 T_m(y_j)}{dy^2} \right] = f(x_i, y_j) \quad (4.25)$$

Junto con las condiciones de frontera (proporcionadas para cada ejemplo en particular) se tendrán  $(N+1)^2$  ecuaciones lineales para las  $(N+1)^2$  incógnitas

$a_{nm}$ . De esta manera tendremos que resolver una ecuación matricial de la forma

$$\mathbf{A}\vec{\mathbf{X}} = \vec{\mathbf{b}} \quad (4.26)$$

con el vector columna  $\vec{\mathbf{X}} = (a_{00}, a_{01}, \dots, a_{NN})$ .

Haciendo referencia a la ecuación 4.25 podemos ver que los términos de la matriz  $\mathbf{A}$  están formados por  $\left[ \frac{d^2 T_n(x_i)}{dx^2} T_m(y_j) + T_n(x_i) \frac{d^2 T_m(y_j)}{dy^2} \right]$ , donde la primera fila se obtiene cuando  $i = 0, j = 0$ , la segunda cuando  $i = 0, j = 1$  y así sucesivamente.

De la ecuación 4.26 vemos que solo es necesario invertir la matriz  $\mathbf{A}$  y podremos encontrar cada uno de los coeficientes  $a_{nm}$  y así encontrar la solución  $u(x, y)$  ec. 4.23.

Para invertir la matriz  $\mathbf{A}$  se utilizó el método de descomposición  $LU$ . En este método se escribe  $\mathbf{A}$  como el producto de dos matrices

$$\mathbf{A} = \mathbf{L}\mathbf{U} \quad (4.27)$$

donde  $\mathbf{L}$  es una matriz triangular inferior (tiene elementos diferentes de cero solo en la diagonal y abajo de la diagonal) y  $\mathbf{U}$  es una matriz triangular superior (tiene elementos diferentes de cero solo en la diagonal y arriba de la diagonal). Usamos la descomposición (4.27) para resolver el sistema  $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$

$$\mathbf{A}\vec{\mathbf{x}} = (\mathbf{L}\mathbf{U}) \cdot \vec{\mathbf{x}} = \mathbf{L}(\mathbf{U}\vec{\mathbf{x}}) = \vec{\mathbf{b}}$$

en dos pasos. Primero se resuelve

$$\mathbf{L}\vec{\mathbf{y}} = \vec{\mathbf{b}}$$

y después se resuelve

$$\mathbf{U}\vec{\mathbf{x}} = \vec{\mathbf{y}}$$

para encontrar  $\vec{\mathbf{x}}$ . La ventaja de esta descomposición es que resolver un sistema triangular de ecuaciones, es muy simple. Los detalles de cómo se realiza esta descomposición se encuentran en el capítulo 2 del libro “Numerical Recipes in C” [10]. Los resultados numéricos se presentan en el siguiente capítulo.

### 4.3. Método Espectral Tau

#### 4.3.1. Ecuación de Burgers

Vamos a resolver la ecuación diferencial de Burgers usando el método de Tau. Esta ecuación es

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2}, \quad |x| \leq 1 \quad (4.28)$$

con las condiciones de frontera

$$u(\pm 1, t) = 0 \quad (4.29)$$

y condición inicial

$$u(x, 0) = -\sin \pi x \quad (4.30)$$

donde  $\nu$  es la viscosidad. Para viscosidades pequeñas, la solución se desarrolla como una onda de “diente de sierra” en el origen. La solución teórica de este problema es conocida, la obtuvo J. D. Cole [8], y fue compilada por Benton y Platzmann [9]:

$$u(x, t) = 4\pi\nu \left[ \frac{\sum_{n=1}^{\infty} na_n e^{-n^2\pi^2 t\nu} \sin n\pi x}{a_0 + 2 \sum_{n=1}^{\infty} a_n e^{-n^2\pi^2 t\nu} \cos n\pi x} \right], \quad (4.31)$$

donde  $a_n = (-1)^n I_n(1/2\pi\nu)$  e  $I_n(\xi)$  denota las funciones de Bessel modificadas del primer tipo. A veces se piensa que una solución analítica es siempre mucho mejor que una numérica, pero esto no es siempre así, para graficarla con una computadora, la solución (4.31) es intratable para valores pequeños de  $t$  y  $\nu$ , donde  $I_n(\xi)$ , cuando  $\xi \rightarrow \infty$ , se comporta asintóticamente como  $e^{\xi}(2\pi\xi)^{-1/2}$  independiente de  $n$ .

Una representación manejable de la solución es

$$u(x, t) = -\frac{I(x, t)}{J(x, t)} \quad (4.32)$$

donde  $I(x, t)$  y  $J(x, t)$  están definidas por (ver [11])

$$I(x, t) = \int_{-\infty}^{\infty} \sin[\pi(x - \eta)] f(x - \eta) \exp(-\eta^2/4\nu t) d\eta \quad (4.33)$$

$$J(x, t) = \int_{-\infty}^{\infty} f(x - \eta) \exp(-\eta^2/4\nu t) d\eta \quad (4.34)$$

donde  $f(\xi) = \exp(-\cos(\pi\xi/2\pi\nu))$ . Esta solución exacta nos permitirá comparar nuestra solución numérica. Usaremos el método trapezoidal con límites finitos. Debido al hecho de que los integrandos decaen muy rápidamente cuando  $|\eta|$  aumenta, la integración se hará en un intervalo entre  $[-n, n]$  y luego en  $[-2n, 2n]$  y si no cambia el resultado dentro de una tolerancia previamente establecida (por ejemplo  $10^{-10}$ ), entonces ahí nos detenemos.

### 4.3.2. Implementación del método Tau

Como se ha dicho, empezamos desarrollando  $u(x, t)$  en términos de los primeros  $N + 1$  polinomios de un conjunto ortogonal de funciones. En este caso usamos los polinomios de Chebyshev

$$u_N(x, t) = \sum_{n=0}^N a_n(t) T_n(x) \quad (4.35)$$

A fin de minimizar el residuo

$$R(x, t) = \frac{\partial u_N}{\partial t} + u_N \frac{\partial u_N}{\partial x} - \nu \frac{\partial^2 u_N}{\partial x^2} \quad (4.36)$$

uno demanda que  $R(x, t)$  sea ortogonal al espacio expandido por  $\{T_k(x)\}_{k=0}^{N-2}$ , es decir

$$\int_{-1}^1 R(x, t) T_k(x) \frac{dx}{\sqrt{1-x^2}} = 0 \quad 0 \leq k \leq N-2 \quad (4.37)$$

esta integral, debido a la ortogonalidad de los polinomios  $T_k(x)$ , nos conduce a

$$\dot{a}_k = \hat{N}_k + \nu a_k^{(2)} \quad 0 \leq k \leq N-2 \quad (4.38)$$

donde  $\hat{N} = -(U_N \partial_x u_N)_k$ , y  $a_k^{(2)}$  son los coeficientes de la expansión de la segunda derivada de  $u(x, t)$  con respecto a la variable espacial  $x$ . La manera más eficiente de calcularlos es usando la relación de recurrencia

$$c_k a_k^{(2)} = a_{k+2}^{(2)} + 2(k+1) a_{k+1}^{(1)}. \quad (4.39)$$

El sistema de ecuaciones 4.38 junto con las condiciones de frontera  $u(x = -1, t) = u(x = 1, t) = 0$ , constituyen un conjunto completo para calcular los coeficientes de la expansión a todo tiempo con la condición inicial dada. Las condiciones de frontera en terminos de los coeficientes son

$$\sum_{n=0}^N a_n = \sum_{n=0}^N (-1)^n a_n = 0 \quad (4.40)$$

que son equivalentes al par de ecuaciones

$$\sum_{\substack{n=0 \\ n \text{ par}}} a_n = 0, \quad \sum_{\substack{n=1 \\ n \text{ impar}}} a_n = 0. \quad (4.41)$$

El sistema 4.38 evoluciona a los coeficientes  $a_n$  para  $0 \leq N - 2$ , y con las condiciones de frontera 4.41, se calculan  $a_{N-1}$  y  $a_N$  a cada paso de tiempo.

El termino no lineal de 4.38 requiere de trato especial. Tanto  $u(x, t)$  como  $\partial_x u(x, t)$  tienen su expansión en terminos de los polinomios  $T_n$  y tienen sus coeficientes respectivos  $a_n$  y  $a_n^{(1)}$ . La pregunta es, cuáles son los coeficientes  $\langle u \partial_x u \rangle_k = b_k$  de la expansión del producto  $u \partial_x u$

$$u_N \frac{\partial u_N}{\partial x} = \sum_{k=0}^N b_k T_k(x), \quad (4.42)$$

en términos de  $a_k$  y  $a_k^{(1)}$ . En principio podría usarse la relación

$$b_k = \langle u \partial_x u \rangle_k = \frac{1}{2} \left[ \sum_{p+q=k} a_p a_q^{(1)} + \sum_{|p-q|=k} a_p a_q^{(1)} \right] \quad (4.43)$$

que puede demostrarse facilmente con ayuda de las propiedades de los polinomios de Chebyshev. Sin embargo, es mucho más eficiente utilizar un cálculo pseudo-espectral con la ayuda de las Transformadas de Fourier Rápidas (TFR). El proceso del cálculo se sistematiza a continuación

Cálculo de  $\langle u \partial_x u \rangle_k = b_k$  dados los coeficientes  $a_k$

1. Use la relación de recurrencia

$$c_k a_k^{(1)} = a_{k+2}^{(1)} + 2(k+1)a_{k+1} \quad (4.44)$$

para calcular los coeficientes de la expansión de la primera derivada con respecto a  $x$  de  $u(x, t)$

2. Las TFR necesitan de “entrada” los coeficientes de la expansión de una función  $u(x, t)$ , y dan de salida la función valuada en varios puntos (de “colocación”)  $u_N(x_i)$ . Así que, teniendo  $a_k$  y  $a_k^{(1)}$ , utilice las TFR y obtenga  $u_N(x_i)$  y  $(\partial u_N / \partial x)(x_i)$  (esto es,  $u$  y su derivada con respecto a  $x$  en el espacio físico).

$$a_k \longrightarrow u_N(x_i) \quad a_k^{(1)} \longrightarrow \frac{\partial u_N(x_i)}{\partial x} \quad (4.45)$$

3. Realice los productos en el espacio físico:

$$u_N(x_i) \frac{\partial u_N}{\partial x}(x_i) \quad (4.46)$$

4. A través de la TFR inversas, regrese al espacio espectral. La TFR inversa tienen por entrada los valores dados por (4.46), y por salida los coeficientes  $\langle u \partial_x u \rangle_k = b_k$

$$\frac{\partial u_N(x_i)}{\partial x} \longrightarrow \langle u \partial_x u \rangle_k = b_k \quad (4.47)$$

Esta “danza” entre el espacio espectral y el físico, se realiza a cada paso de tiempo. La evolución temporal se realizará usando el esquema Adams-Bashforth para el término no lineal, y el esquema Crank-Nicolson para el término lineal. De este modo se tiene el esquema ABCN

$$\begin{aligned} a_n(t + \Delta t) &= a_n(t) + \frac{\Delta t}{2} [b_n(t - \Delta t) - 3b_n(t)] \\ &+ \frac{\nu \Delta t}{2} [a_n^{(2)}(t) + a_n^{(2)}(t + \Delta t)] \quad 0 \leq n \leq N - 2 \end{aligned} \quad (4.48)$$

que puede re-escribirse como

$$(a_n^{(2)} - \lambda a_n)(t + \Delta t) = f_n(a_n, b_n, t, t - \Delta t) \quad 0 \leq n \leq N - 2 \quad (4.49)$$

donde  $\lambda = 2/\nu\Delta t$  y la función  $f_n$  es explícitamente

$$f_n = -\lambda a_n(t) - \frac{1}{\nu} [b_n(t - \Delta t) - 3b_n(t)] - a_n^{(2)}(t) \quad (4.50)$$

Si usamos la ecuación (??) en (4.49), ésta última se escribiría como

$$\frac{1}{c_k} \sum_{\substack{p=k+2 \\ (p+k) \text{ par}}}^N p(p^2 - k^2)a_p - \lambda a_k = f_k \quad k = 0, \dots, N-2 \quad (4.51)$$

Usando (4.41) y (4.51) llegamos a un sistema de ecuaciones lineales de la forma

$$\mathcal{L}\mathbf{X} = \mathbf{b} \quad (4.52)$$

donde  $\mathcal{L}$  es una matriz triangular superior, cuya solución requiere  $N^2$  operaciones. Un procedimiento de solución más eficiente se consigue reorganizando las ecuaciones (4.49) con ayuda de la relación de recurrencia

$$2ka_k^{(1)} = c_{k-1}a_{k-1}^{(2)} - a_{k+1}^{(2)} \quad (4.53)$$

la ecuación (4.49) queda

$$2ka_k^{(1)} = c_{k-1} [-f_{k-1} + \lambda a_{k-1}] - [-f_{k+1} + \lambda a_{k+1}] \quad k = 1, \dots, N-3. \quad (4.54)$$

Después, usando la relación de recurrencia (3.25) para la primera derivada, en combinación con la (4.54), resulta una ecuación con solamente los coeficientes  $a_k$ , a saber

$$\begin{aligned} & \frac{c_{n-2}\lambda}{4n(n-1)}a_{n-2} - \left[1 + \frac{\lambda\beta_n}{2(n^2-1)}\right] a_n + \frac{\lambda\beta_{n+2}}{4n(n+1)}a_{n+2} \\ = & -\frac{c_{n-2}}{4n(n-1)}f_{n-2} + \frac{\beta_n}{2(n^2-1)}f_n - \frac{\beta_{n+2}}{4n(n+1)}f_{n+2}, \quad k = 2, \dots, \end{aligned} \quad (4.55)$$



donde

$$\beta_n = \begin{cases} 1 & 0 \leq n \leq N - 2 \\ 0 & n > N - 2. \end{cases} \quad (4.56)$$

Note que los coeficientes pares e impares están desacoplados en (4.55) y en (4.41). El sistema de ecuaciones que surge para los coeficientes pares e impares resulta ser uno con una matriz tridiagonal con una “trenza” (en la primera fila). Un procedimiento muy eficiente (mucho más que el de Gauss) se expone en el apéndice A [6], y es el que se usó para integrar en el tiempo la ecuación de Burgers. Para manejar el error de “aliasing” explicado en el apéndice A, se usó la regla de  $2N/3$ . Es decir, en la expansión de  $u_N(x, t)$  se tienen 32 términos, realmente se usarán 21; si se tienen 64, realmente se usarán 42, etc.

# Capítulo 5

## Resultados Numéricos

En este capítulo se dan a conocer los resultados numéricos que se obtuvieron al resolver las ecuaciones de Onda, de Poisson y Burgers, con los Métodos de Diferencias Finitas y Espectrales.

### 5.1. Ecuación de Onda

#### 5.1.1. Resultados Numéricos aplicando Diferencias Finitas

Para encontrar el valor de  $\phi$  por medio de las ecuaciones descritas en el capítulo anterior, se realizó un programa en lenguaje C. En éste programa también se calcula el error (Norma 1) que hay entre la solución numérica y la solución exacta. Para realizar los cálculos se utilizó el valor de  $c^2 = 1$ .

$$e = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (\phi_{exac}(x_i, t_j) - \phi_{num}(x_i, t_j))^2} \quad (5.1)$$

Se utilizaron varios tipos de condiciones de frontera, los resultados obtenidos

con cada una de éstas se dan a conocer a continuación.

En las condiciones de frontera periódicas para encontrar el valor de  $\psi$  y  $\pi$  en las fronteras del dominio  $[-1,1]$  en un tiempo  $t_j$  según las ecuaciones (4.10) y (4.11), necesitamos conocer los valores de estas variables en las posiciones  $(i + 1)$  e  $(i - 1)$  en un tiempo  $t_{j-1}$ . Entonces por ejemplo, para encontrar los valores de la  $\psi$  en el tiempo  $t_1$  en la frontera  $x = -1$ , se deben conocer  $\psi(x_1)$  y  $\psi(x_{-1})$ , para  $x = 1$  debemos saber los valores de  $\psi(x_{N+1})$  y  $\psi(x_{N-1})$  en el tiempo  $t = 0$ . De las condiciones iniciales podemos obtener los valores de  $\psi(x_1)$  y  $\psi(x_{N-1})$ , pero los otros dos valores no los conocemos. Lo mismo sucede cuando queremos conocer los valores de las fronteras para  $\pi$ .

Debido a lo anterior, se hizo la siguiente asignación a las posiciones que no se conocían,

$$\begin{aligned} \psi(x_{-1}) &= \psi(x_{N-1}) & \pi(x_{-1}) &= \pi(x_{N-1}) \\ \psi(x_{N+1}) &= \psi(x_1) & \pi(x_{N+1}) &= \pi(x_1) \end{aligned} \quad (5.2)$$

para todo tiempo  $t > 0$ .

Para  $\beta$  usamos los valores de 0, 1 y -1. Obteniendo así los siguientes resultados.

Las figuras 5.1 - 5.3, muestran la solución de la ecuación de onda utilizando en método de RK4 con las condiciones periódicas y los distintos valores de  $\beta$ .

Al ir aumentando el valor de la  $N$ , cuando ejecutábamos el programa nos dimos cuenta de que el error que obteníamos iba decreciendo, como se esperaba, pues para valores de  $N$  mayores, el valor de  $k$  es menor y la aproximación es más precisa. Lo anterior se muestra en la tabla 5.1 y la representación grá-

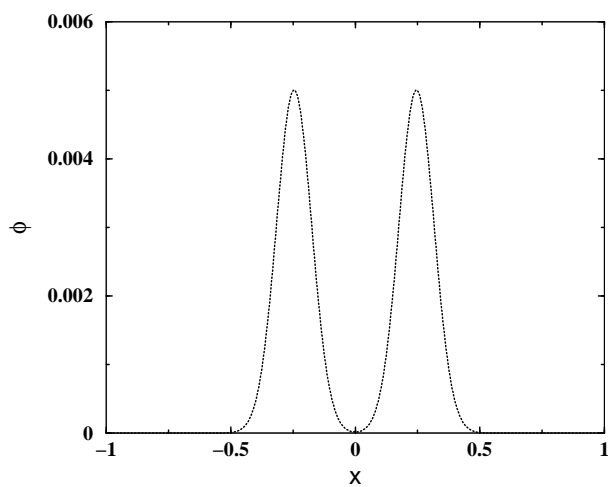


Figura 5.1: Gráfica de  $\phi(x)$  en un tiempo  $t \approx 0.18$ . Con  $\phi(x) = Ae^{x^2/\sigma^2}$  como condición inicial, usando condiciones de frontera periódicas,  $N=500$  y  $\beta=0$ .

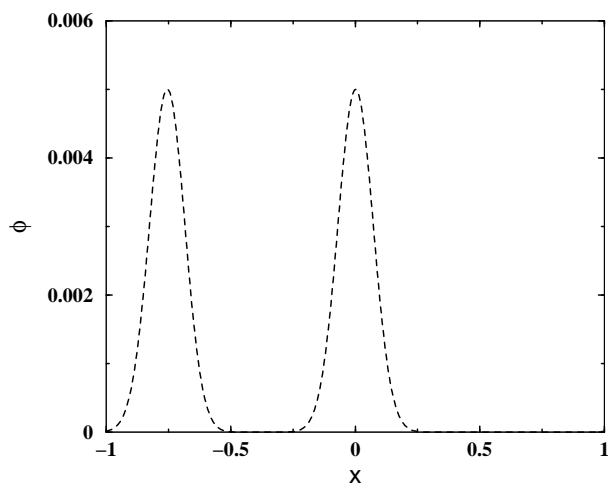


Figura 5.2: Gráfica de  $\phi(x)$  en un tiempo  $t \approx 0.26$ . Con  $\phi(x) = Ae^{x^2/\sigma^2}$  como condición inicial y usando condiciones de frontera periódicas,  $N=500$  y  $\beta=1$ .

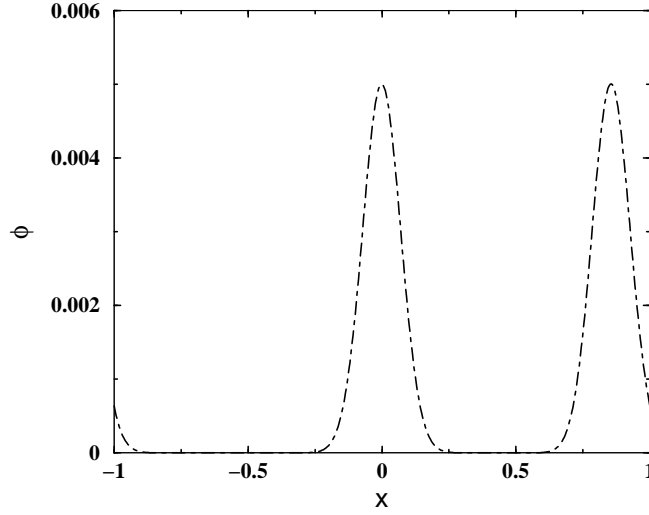


Figura 5.3: Gráfica de  $\phi(x)$  en un tiempo  $t \approx 0.30$ . Con  $\phi(x) = Ae^{x^2/\sigma^2}$  como condición inicial usando condiciones de frontera periódicas,  $N=500$  y  $\beta = -1$ .

fica en 5.4.

Para definir las condiciones de frontera reflejantes, nos interesa que  $\phi$  sea estática en las fronteras, es decir que la  $\frac{\partial \phi}{\partial t} = 0$  y de la definición de  $\pi$  con  $\beta = 0$  obtenemos que  $\pi = \frac{\partial \phi}{\partial t}$ , entonces  $\pi = 0$  en las fronteras, de esto tenemos que  $\frac{\partial \pi}{\partial t} = 0$  y como  $\frac{\partial \pi}{\partial t} = \frac{\partial \psi}{\partial x}$ , llegamos a  $\frac{\partial \psi}{\partial x} = 0$ . Es por eso que podemos definir las condiciones de frontera reflejantes de la siguiente forma:

$$\psi(x_0) = \psi(x_1) \quad \psi(x_N) = \psi(x_{N-1}) \quad \pi(x_0) = \pi(x_N) = 0 \quad (5.3)$$

Pero con estas condiciones  $\beta$  solo podía tomar el valor de 0.0, entonces se buscaron condiciones de frontera reflejantes válidas para valores de  $\beta \neq 0$  y se llegó a lo siguiente,

Cuadro 5.1: En la tabla se muestra el error que hay entre la solución exacta y la numérica obtenido para algunos valores de  $N$ , con condiciones de frontera periódicas a un tiempo  $t \approx 0.035$ .

N	$\beta=0.0$	$\beta=1, -1$
50	5.2100e-05	1.0755e-04
300	2.9292e-05	5.8509e-05
550	2.2019e-05	4.4005e-05
800	1.8367e-05	3.6716e-05
1000	1.6470e-05	3.2927e-05

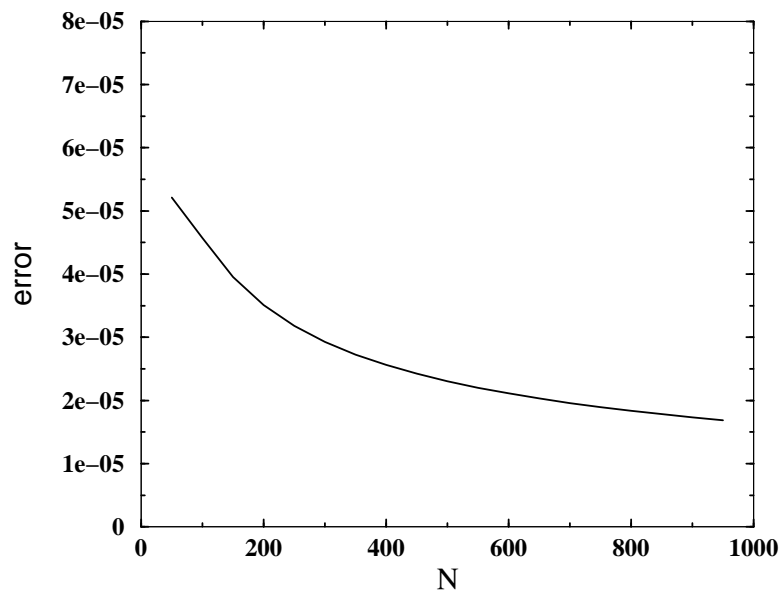


Figura 5.4: El error disminuye conforme la  $N$  aumenta. Condiciones de frontera periódicas con  $\beta=0.0$  y  $t \approx 0.035$ .

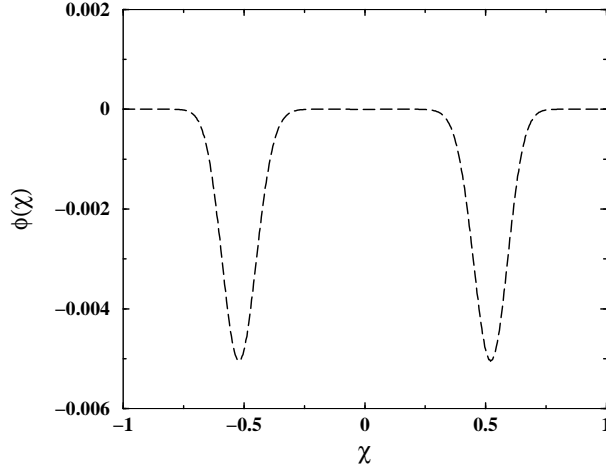


Figura 5.5: Gráfica de  $\phi(x)$  en un tiempo  $t \approx 1.7$ . Utilizando  $\phi(x) = Ae^{x^2/\sigma^2}$  como condición inicial y con condiciones de frontera reflejantes,  $N = 500$  y  $\beta=0.0$ .

$$\begin{aligned} \psi(x_0) &= \frac{\psi(x_1)(1 - \beta^2) + 2\beta\pi(x_1)}{(1 - \beta^2)} & \pi(x_0) &= -\beta\psi(x_0) \\ \psi(x_N) &= \frac{\psi(x_{N-1})(1 + \beta^2) + 2\beta\pi(x_N - 1)}{(1 - \beta^2)} & \pi(x_N) &= -\beta\psi(x_N). \end{aligned} \quad (5.4)$$

De las condiciones de frontera podemos observar que  $\beta$  no puede tomar los valores  $\pm 1$ . Los valores para  $\beta$  que se utilizaron en este trabajo son: 0.0, 0.4 y -0.4. Podemos apreciar la solución de  $\phi(x)$  con estas condiciones de frontera y los valores para  $\beta$  mencionados en las figuras 5.5-5.9.

## 5.2. Ecuación de Poisson

### 5.2.1. Resultados Numéricos aplicando Diferencias Finitas

Utilizamos  $-\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$  y  $(n^2 + m^2)e^{nx+my}$  como las funciones  $f(x, y)$  en las ecuaciones (4.17), (4.19) y (4.22).

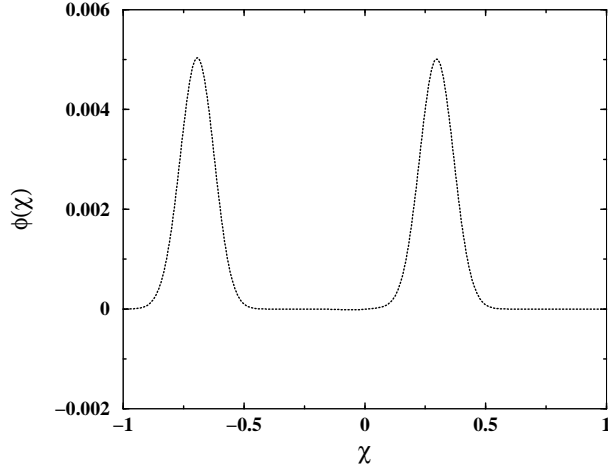


Figura 5.6: Gráfica de  $\phi(x)$  en un tiempo  $t \approx 0.13$ . Utilizando  $\phi(x) = Ae^{x^2/\sigma^2}$  como condición inicial y con condiciones de frontera reflejantes,  $N = 500$  y  $\beta=0.4$ .

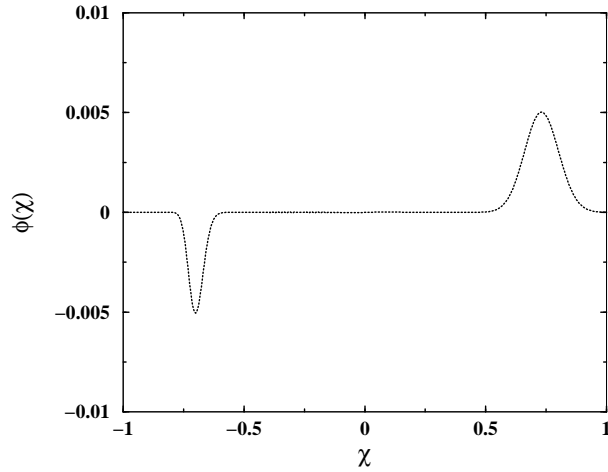


Figura 5.7: Gráfica de  $\phi(x)$  en un tiempo  $t \approx 1.42$ . Utilizando  $\phi(x) = Ae^{x^2/\sigma^2}$  como condición inicial y con condiciones de frontera reflejantes,  $N = 500$  y  $\beta=0.4$ .



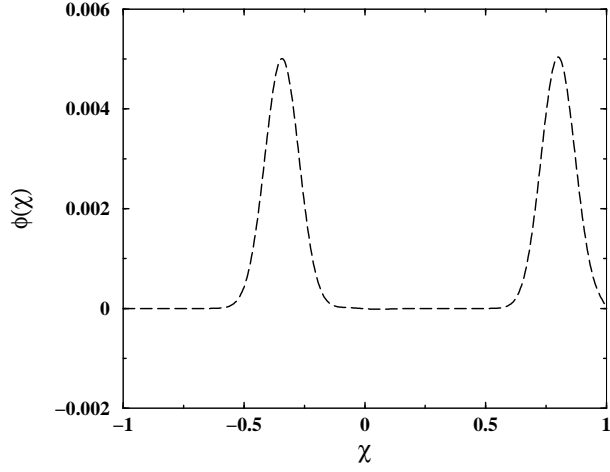


Figura 5.8: Gráfica de  $\phi(x)$  en un tiempo  $t \approx 0.12$ . Condiciones de frontera reflejantes,  $N = 500$  y  $\beta = -0.4$ .

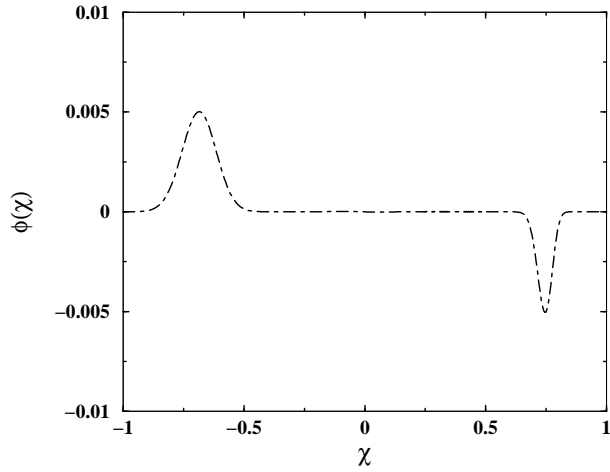


Figura 5.9: Gráfica de  $\phi(x)$  en un tiempo  $t \approx 1.31$ . Utilizando  $\phi(x) = Ae^{x^2/\sigma^2}$  como condición inicial y con condiciones de frontera reflejantes,  $N = 500$  y  $\beta = -0.4$ .

Cuadro 5.2: Error obtenido con algunos valores de  $N$ , para condiciones reflejantes con un tiempo  $t \approx 0.017$ .

N	$\beta=0.4, -0.4$
50	2.1949e-04
250	1.4738e-05
500	1.0313e-05
750	8.3824e-06
1000	7.2411e-06

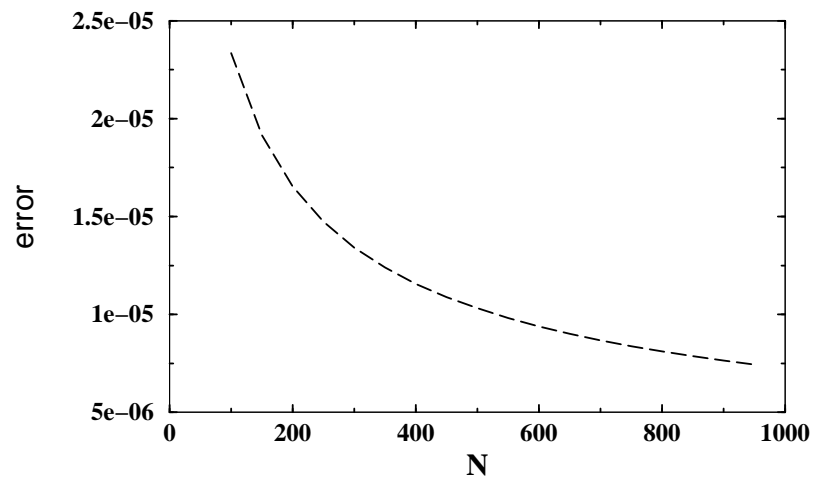


Figura 5.10: El error disminuye conforme la  $N$  aumenta. Condiciones de frontera reflejantes con  $\beta=-0.4$  y  $t \approx 0.017$ .

Se realizó un programa en lenguaje C, para obtener la solución de las ecuaciones: ec. (4.17)(Método de Jacobi), ec. (4.19) (Método de Gauss-Seidel) y ec.( 4.22) (Método de SOR). En la figura 5.11 se muestra  $\phi(x, y)$  para la solución de prueba  $-\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$  utilizando el método de SOR.

Este programa también nos proporciona el error que resulta de comparar la solución exacta y la solución numérica (Norma 2) obtenida con los tres métodos antes mencionados, el error se calcula para cada iteración,

$$e_k = \sqrt{\sum_{i,j=1}^{J-1} (\phi_{exact}(x_i, y_j) - \phi_{num}(x_i, y_j))^2} \quad (5.5)$$

Con  $f(x, y) = -\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$  utilizamos como solución de “prueba” las siguientes funciones:  $0.0$ ,  $0.5 \sin(\pi nx) \sin(\pi my)$  y  $0.9 \sin(\pi nx) \sin(\pi my)$ , y las condiciones iniciales son  $F(-1, y) = F(1, y) = F(x, -1) = F(x, 1) = 0.0$ . Al hacer funcionar el programa con cada una de las soluciones de “prueba” para cada método, nos dimos cuenta que el error era el mismo para las soluciones de prueba, con cada uno de los tres métodos, la única diferencia era el número de iteraciones en el que se obtuvo el error, es decir, que con una solución de “prueba” el número de iteración fue más pequeño que con las otras. Lo anterior se muestra en las tablas (5.3) y (5.4).

El programa se compiló con distintas  $N$ , para ver cómo era el comportamiento del error para cada una de éstas, y lo que se obtuvo fue que al ir aumentando el valor de la  $N$ , el error se hacía más pequeño, ésto se muestra en las tablas 5.3 y 5.4. Estos cambios se hicieron para las tres soluciones de prueba con cada uno de los tres métodos y el resultado fue el mismo. La grafica 5.12 nos muestra lo anterior.

Tomando como referencia las tablas 5.3 y 5.4, hacemos una comparación

'seidel.dat' ———

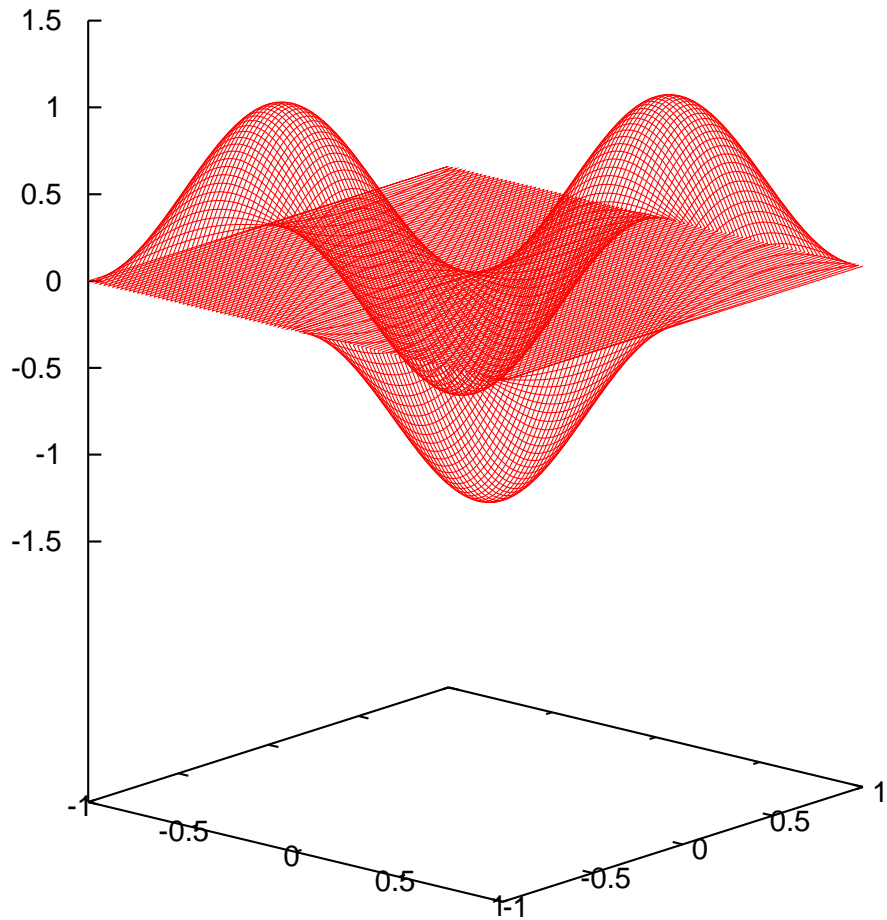


Figura 5.11: Gráfica de  $F(x,y)$  para  $x,y \in [-1,1]$ , con el método de SOR utilizando  $N = 100$ ,  $n = m = 1$  para la solución de prueba  $-\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi mx)$ .

Cuadro 5.3: En la tabla se muestra el número de iteraciones que se realizó en cada método para obtener un error  $\approx 4.64e-4$  con  $N = 60$  (número de divisiones de cada eje de la red) y con  $n = m = 1$  para las soluciones de prueba  $0$ ,  $0.5\sin(nx) \sin(my)$  y  $0.9\sin(nx) \sin(my)$ .

Método	0.0	$0.5\sin(nx) \sin(my)$	$0.9\sin(nx) \sin(my)$
Jacobi	5632	5508	5216
Gauss-Seidel	3117	2941	2727
SOR	998	970	911

Cuadro 5.4: Número de iteraciones para un error  $\approx 1.66e-4$  con  $N = 100$  y  $n = m = 1$ , para las tres soluciones de prueba.

Método	0.0	$0.5\sin(nx) \sin(my)$	$0.9\sin(nx) \sin(my)$
Jacobi	15149	14798	13985
Gauss-Seidel	9541	8989	8173
SOR	2521	2434	2261

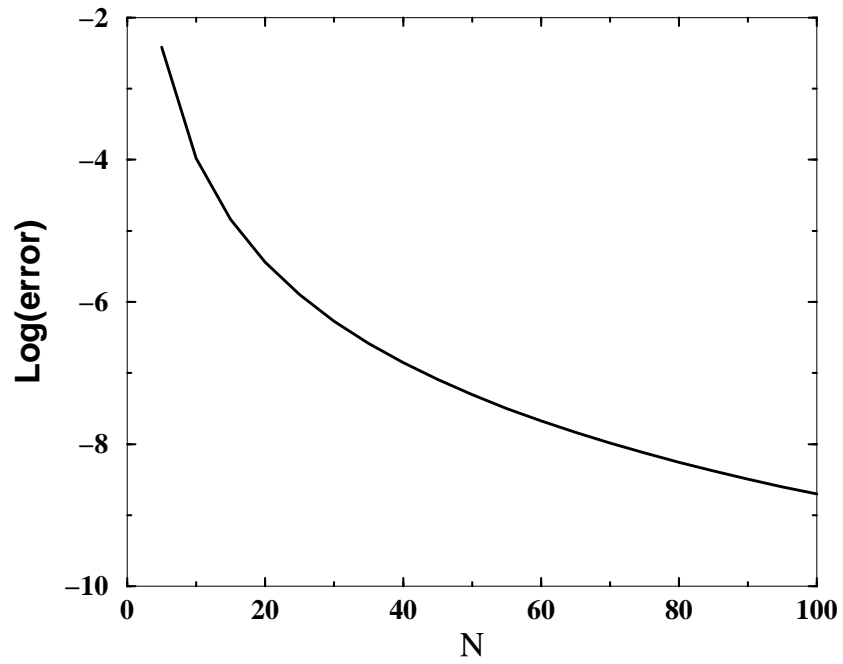


Figura 5.12: Para valores más grandes de  $N$  (no. de divisiones de cada eje de la red), el error es más pequeño, es decir, la solución numérica se acerca más a la solución exacta, con  $n = m = 1.0$ , para la solución de prueba  $-\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$

Cuadro 5.5: Comparación de los tres Métodos, con un error  $\approx 1.6618e-4$ ,  $N = 100$ ,  $m = n = 1$  y tomando como solución de prueba a  $0.9\sin(nx)\sin(my)$ .

Método	Iteración
Jacobi	13985
Gauss-Seidel	8173
SOR	2261

Cuadro 5.6: Número de iteraciones para un error  $\approx 3.63e-04$  con  $N = 20$ ,  $n = 1$  y  $m = 1$  con las funciones de prueba  $0.0$ ,  $0.5 e^{nx+my}$  y  $0.9 e^{nx+my}$ .

Método	0.0	$0.5e^{nx+my}$	$0.9e^{nx+my}$
Jacobi	2609	2553	2427
Gauss-Seidel	1135	1307	1242
SOR	162	149	148

entre los tres métodos para ver cual de éstos es más rápido para obtener la solución numérica más aproximada a la solución exacta y llegamos a la conclusión que es el Método de SOR (Tabla 5.5), por lo que es con éste con el que se hará posteriormente la comparación entre Diferencias Finitas y el Método de Colocación Espectral.

Cuando aplicamos  $f(x, y) = (n^2 + m^2)e^{nx+my}$ , utilizamos como funciones de prueba:  $0.0$ ,  $0.5 e^{nx+my}$  y  $0.9 e^{nx+my}$  y con condiciones iniciales  $F(-1, y) = (n^2 + m^2)e^{-n+my}$ ,  $F(1, y) = (n^2 + m^2)e^{n+my}$ ,  $F(x, -1) = (n^2 + m^2)e^{nx-m}$  y  $F(x, 1) = (n^2 + m^2)e^{nx+m}$ , en las tablas 5.6 y 5.7 se muestran los resultados que se obtuvieron para este ejemplo. La gráfica 5.13 muestra la solución  $\phi$  para  $0.9e^{(nx+my)}$  con el método de SOR.

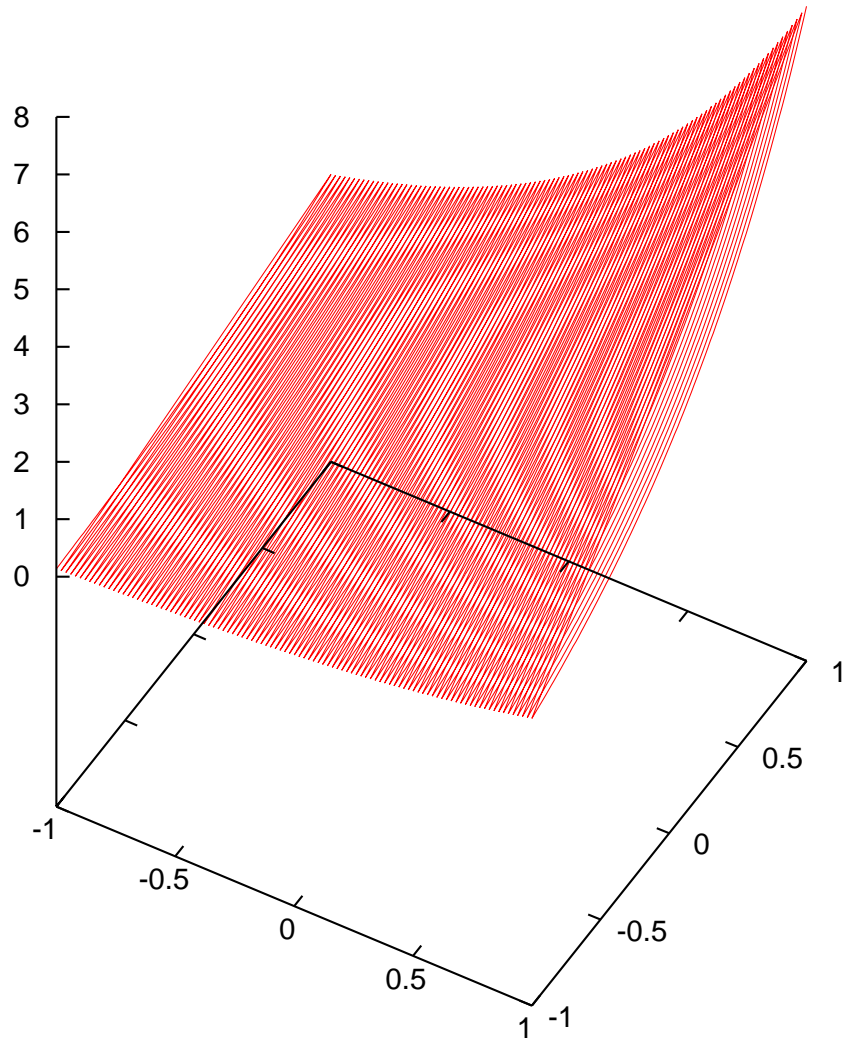


Figura 5.13: Gráfica de  $F(x, y)$ , para  $x, y \in [-1, 1]$ , utilizando el método de SOR con  $N = 100$  y  $n = m = 1,0$  y con la función de prueba  $0.9e^{(nx+my)}$ .



Cuadro 5.7: Número de iteraciones para un error  $\approx 1.39e-05$  con  $N = 100$ ,  $n = 1$  y  $m = 1$  y con las soluciones de prueba  $0.0$ ,  $0.5 e^{nx+my}$  y  $0.9 e^{nx+my}$ .

Método	0.0	$0.5e^{nx+my}$	$0.9e^{nx+my}$
Jacobi	59065	57638	54386
Gauss-Seidel	30131	29429	27798
SOR	3547	3461	3295

Cuadro 5.8: Comparación de los tres Métodos, con un error  $\approx 1.39e-05$ ,  $N = 100$ ,  $m = 1$ ,  $n = 1$  y tomando como solución de prueba a  $0.9e^{nx+my}$ .

Método	Iteración
Jacobi	54386
Gauss-Seidel	27798
SOR	3295

Las tablas (5.6) y (5.7) muestran lo ya dicho en el ejemplo anterior, al calcular el error entre la solución exacta y la numérica resulta ser el mismo para todas las soluciones de “prueba” con cada uno de los métodos, pero con diferentes iteraciones. Ésto se muestra en la figura 5.14.

Es evidente que el método que menos iteraciones realiza para llegar a la solución numérica más cercana a la solución exacta es el método de SOR (Tabla 5.8), de la misma manera que en el ejemplo anterior utilizaremos el método de SOR para hacer la comparación con el método de Colocación Espectral.

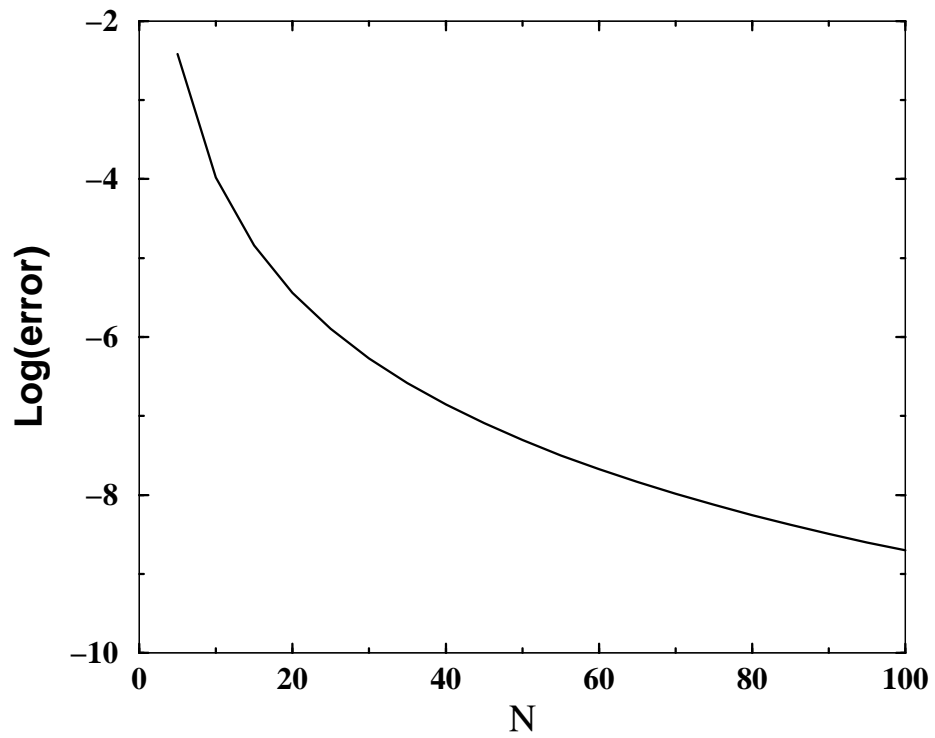


Figura 5.14: Al ir aumentando el valor de la  $N$ , el valor del error se reduce, (Método de SOR con  $n = m = 1$  y función de prueba  $0.9e^{nx+my}$ )

### 5.2.2. Resultados numéricos aplicando el Método de Colocación Espectral

Para este método también trabajaremos con las funciones  $f(x, y)$  que se presentan en la ecuación (4.12),  $-\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$  y  $(n^2 + m^2)e^{nx+my}$ , para  $x, y \in [-1, 1]$ .

Al igual que con Diferencias Finitas, también se realizó un programa en C [?], con el que obtenemos la solución numérica.

Dentro del mismo programa se calcula el error entre la solución numérica y la solución exacta.

$$e = \sqrt{\int_{-1}^1 \int_{-1}^1 (\phi_{exact}(x, y) - \phi_{num}(x, y))^2 dx dy} \quad (5.6)$$

En el caso de la función  $f(x, y) = -\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$ , las condiciones de frontera son:

$$u(-1, y) = u(1, y) = u(x, -1) = u(x, 1) = 0.0$$

En la tabla (5.9) se muestran los errores que se obtuvieron al ejecutar el programa con distintos valores para la  $N$ .

Podemos ver de la tabla (5.9) que al aumentar el valor de la  $N$  vamos a obtener un error cada vez más pequeño, entonces tendremos que la solución numérica se aproxima muy rápido a la solución exacta conforme aumenta el valor de la  $N$  (fig. 5.15).

Trabajando con la función  $(n^2 + m^2)e^{nx+my}$ , utilizamos como condiciones de frontera:  $u(-1, y) = (n^2 + m^2)e^{-n+my}$ ,  $u(1, y) = (n^2 + m^2)e^{n+my}$ ,  $u(x, -1) = (n^2 + m^2)e^{nx-m}y$ ,  $u(x, 1) = (n^2 + m^2)e^{nx+m}$

Cuadro 5.9: Error obtenido con distintos valores de  $N$  y con  $n = m = 1$ , para la función  $f(x, y) = -\pi^2(n^2 + m^2) \sin(\pi nx)$ .

N	Error
2	1.000000e+00
8	6.776337e-04
14	2.865277e-09
20	2.282127e-15
26	8.611846e-16

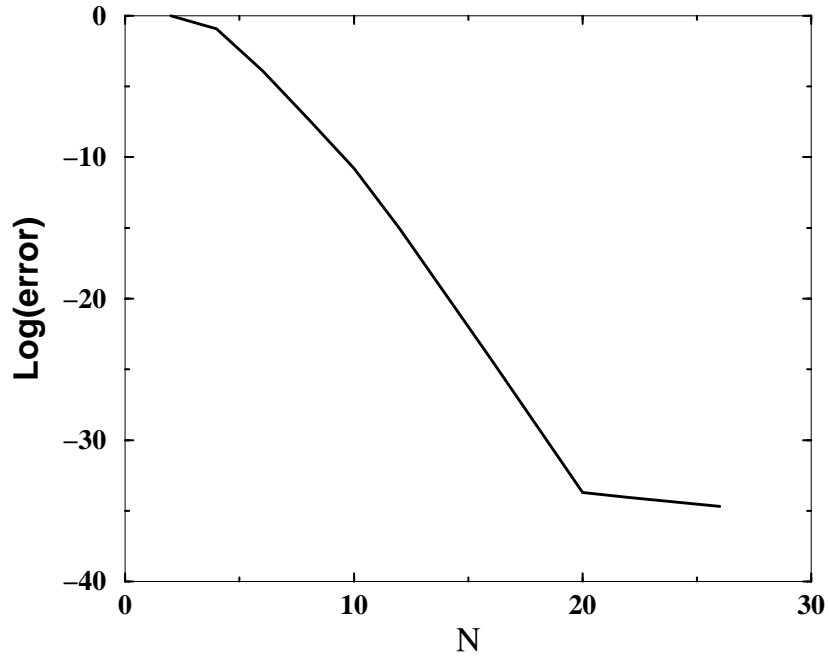


Figura 5.15: A mayor valor para  $N$ , el valor del error disminuye, con  $n = m = 1$ , para la función  $f(x, y) = -\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$ . Para  $N \approx 20$  en adelante el error que la “máquina” calcula es cero.

Cuadro 5.10: Error obtenido con distintos valores de  $N$  y con  $n = 1$  y  $m = 1$ , para la función  $(n^2 + m^2)e^{nx+my}$ .

N	Error
2	2.223901e-01
4	3.210818e-03
8	5.514271e-08
10	1.635970e-13
14	3.764055e-15

En la tabla 5.10, se puede ver como al ir aumentando el valor para la  $N$ , el error cada vez se hace más pequeño.

### 5.2.3. Comparación entre los Métodos Colocación Espectral y de Diferencias Finitas para la ecuación de Poisson

En esta sección se hará una comparación entre los errores obtenidos con cada uno de los Métodos. Primero utilizando la Norma 2 para realizar el cálculo del error en los mismos puntos para los dos Métodos. Después se utilizará la Norma 2 sólo en Diferencias Finitas y para el Método de Colocación, el error se obtendrá con la ecuación 5.6.

Las tablas 5.11 y 5.12 muestran los errores obtenidos para los dos Métodos con la Norma, también se le dieron los mismos valores a la  $N$ . Para Diferencias Finitas la  $N$  es el número de particiones de cada eje de la red y para Espectrales la  $N$  es el número de términos de la expansión.

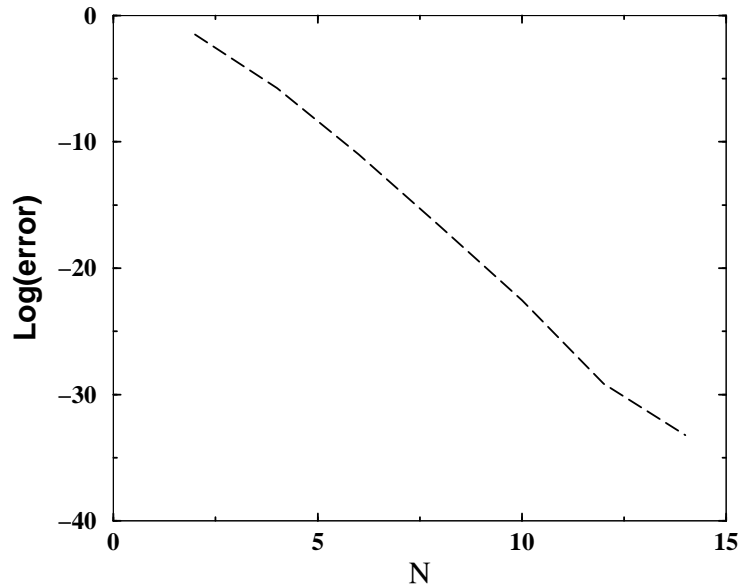


Figura 5.16: El error disminuye conforme aumenta el valor de la  $N$ , con  $n = m = 1$  ( $f(x, y) = (n^2 + m^2)e^{nx+my}$ ) (ME).

Cuadro 5.11: Errores obtenidos con los métodos Diferencias Finitas (SOR) y Espectrales para el ejemplo de  $-\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$ . Usando Norma 2.

Método	Error	N
Diferencias Finitas	2.5380e-03	26
Colocación Espectral	4.8618e-16	26

Cuadro 5.12: Errores obtenidos con los métodos de Diferencias Finitas (SOR) y Espectrales para el ejemplo de  $(n^2 + m^2)e^{(nx+my)}$ . Usando Norma 2

Método	Error	N
Diferencias Finitas	7.5596e-04	14
Colocación Espectral	1.7002e-15	14

Cuadro 5.13: Errores obtenidos con los métodos Diferencias Finitas (SOR) y Espectrales para el ejemplo de  $-\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$ .

Método	Error	N
Diferencias Finitas	1.661877e-4	100
Colocación Espectral	8.611846e-16	26

Cuadro 5.14: Errores obtenidos con los métodos de Diferencias Finitas (SOR) y Espectrales para el ejemplo de  $(n^2 + m^2)e^{(nx+my)}$ .

Método	Error	N
Diferencias Finitas	1.399846e-05	100
Colocación Espectral	3.764055e-15	14

En las tablas 5.13 y 5.14 se hace una comparación de del error obtenido con los métodos de Diferencias Finitas (método de SOR) utilizando Norma 2 y Espectrales utilizando la ecuación 5.6, para las funciones  $-\pi^2(n^2 + m^2) \sin(\pi nx) \sin(\pi my)$  y  $(n^2 + m^2)e^{(nx+my)}$ . En este caso la  $N$  toma el valor con el que se obtiene el error más pequeño para cada método.

### 5.3. Ecuación de Burgers

La ecuación de Burgers con las condiciones de frontera que usamos, tiene una solución que parece un diente de sierra. Así que en el centro (en  $x = 0$ ) hay una pendiente que crece con el tiempo (fig. 5.17). Los métodos espectrales funcionan mejor cuando se tienen funciones bien comportadas. No obstante, aun cuando se presenta este gradiente creciente alrededor del centro, el método Tau funciona bien agregando un mayor número de terminos en la expansión de  $u(x, t)$ . La figura (5.18) muestra la solución numérica de la

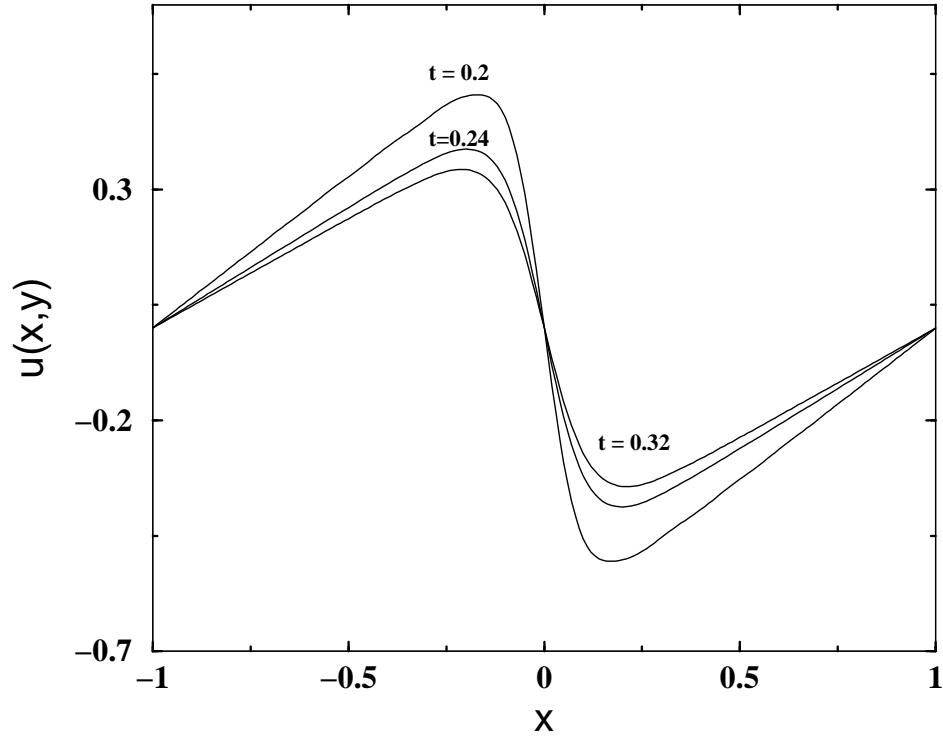


Figura 5.17: Al aumentar el tiempo la pendiente crece. Solución numérica en distintos tiempos con  $\nu = 0.1$ ,  $N=40$ .

ecuación de Burgers con  $N = 32$ , para un  $t = 0,624$ . Lo que las simulaciones numéricas mostraron es que, a medida que aumentamos el valor de  $N$ , la solución numérica se aproxima a la solución exacta. Es posible disminuir el problema con el gradiente en el origen, integrando la ecuación en  $[0, 1]$ , esto es posible por la simetría del problema, pero esto ya no lo hicimos en esta tesis.

Lo que también constatamos con las simulaciones, es que el gradiente se hace menos pronunciado si la viscosidad aumenta.



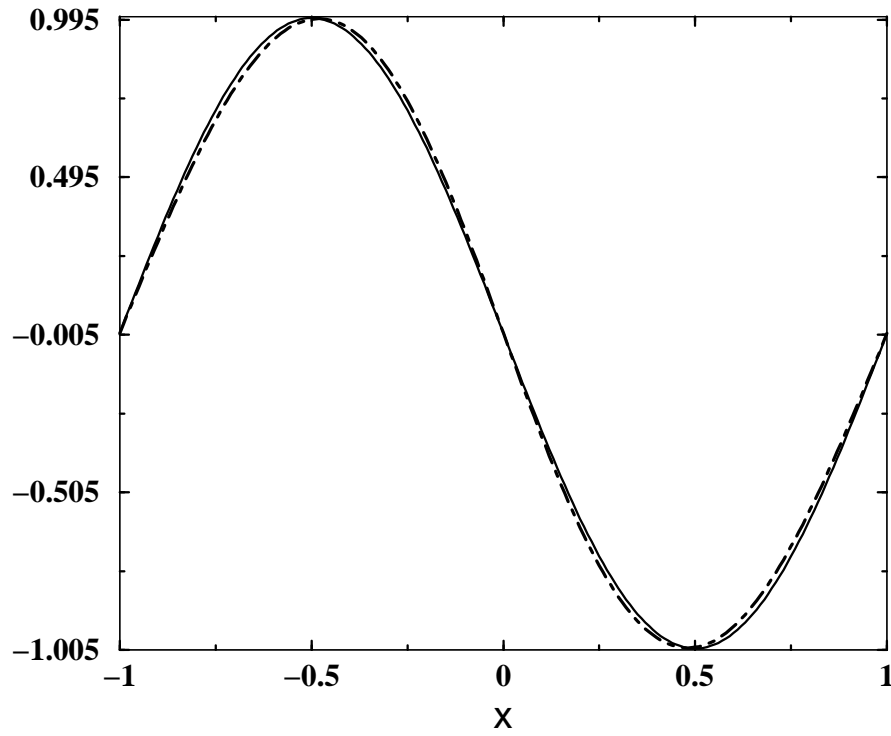


Figura 5.18: Se muestran la solución exacta (línea sólida) y la solución numérica (línea punteada) con  $\nu=0.03/\pi$ . Al ir aumentando el valor de  $N$  la solución numérica se aproxima a la solución exacta. Solución numérica  $u(x, t)$  con  $N = 40$ .

## *Conclusiones:*

En este trabajo se expusieron las generalidades de los Métodos de Diferencias Finitas (MDF) y los Espectrales (ME), y se ilustró su uso en varios ejemplos: la ecuación de calor, la ecuación de onda, la de Burgers, y la de Poisson, todas éstas de amplio uso en Física. Los MDF son los más conocidos y utilizados desde hace muchos años. Los ME empezaron a popularizarse desde hace 30 años y han empezado a usarse en más ramas de la Física más allá de su uso inicial en la Dinámica de los Fluidos. En esta revisión, constatamos que los ME dan resultados más precisos cuando se usa un número  $N$  de términos en la expansión de la solución aproximada  $u_N(x, t)$  comparada con el mismo número de particiones del dominio espacial usado en los MDF. Aunque los ME si son mas laboriosos de implementar, el trabajo extra se compensa pues la convergencia que presentan es exponencial.

Este trabajo sirvió como un entrenamiento para el estudio de las soluciones numéricas de ecuaciones diferenciales que aparecen en Física, y deja espacio para estudios más detallados que revisen aspectos de la influencia de los errores de “aliasing”, estabilidad y convergencia, que en este trabajo, por la extensión del mismo, ya no se realizó con más detenimiento.

El estudio de estos métodos provee una herramienta que seguramente estaremos usando en los próximos dos años en problemas de astrofísica y relatividad numérica, que es una de las áreas de reciente aplicación de los ME.

# Apéndice A

## Normas de Vectores y Matrices

### A.1. Normas de Vectores

La norma de un vector  $\mathbf{x}$  es un número real positivo, el cual proporciona una muestra del "tamaño" del vector y se denota por  $\|\mathbf{x}\|$ . La norma de un vector satisface los siguientes axiomas

- (i)  $\|\mathbf{x}\| > 0$  si  $\mathbf{x} \neq \mathbf{0}$  y  $\|\mathbf{x}\| = 0$  si  $\mathbf{x} = \mathbf{0}$ .
- (ii)  $\|c\mathbf{x}\| = |c|\|\mathbf{x}\|$  para un número escalar real o complejo  $c$ .
- (iii)  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

.

Si el vector  $\mathbf{x}$  tiene componentes  $\{x_1, x_2, \dots, x_n\}$ , entonces las tres normas más usadas comunmente se definen de la siguiente forma:

La 1-norma de  $\mathbf{x}$ , es la suma de los módulos de las componentes del vector, es decir,

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| + \dots + |x_n| = \sum_{i=1}^n |x_i| \quad (\text{A.1})$$

La norma infinita de  $\mathbf{x}$ , es el módulo máximo de las componentes de  $\mathbf{x}$

$$\|\mathbf{x}\|_\infty = \max_i |x_i| \quad (\text{A.2})$$

La 2-norma de  $\mathbf{x}$ , es la raíz cuadrada de la suma de los cuadrados de los módulos de las componentes del vector.

$$\|\mathbf{x}\|_2 = (|x_1|^2 + |x_2|^2 + \dots + |x_n|^2)^{1/2} = \left[ \sum_{i=1}^n |x_i|^2 \right]^{1/2} \quad (\text{A.3})$$

## A.2. Normas de una Matriz

La norma de una matriz  $\mathbf{A}$  es un número real positivo, el cual nos da una muestra del "tamaño" de la matriz y satisfacen los siguientes axiomas

- (i)  $\|\mathbf{A}\| > 0$  si  $\mathbf{A} \neq \mathbf{0}$  y  $\|\mathbf{A}\| = 0$  si  $\mathbf{A} = \mathbf{0}$ .
- (ii)  $\|c\mathbf{A}\| = |c|\|\mathbf{A}\|$  para algún número real o complejo  $c$ .
- (iii)  $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ .
- (iv)  $\|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$ .

## A.3. Normas Compatibles

Vectores y matrices, juntos satisfacen una condición equivalente a la condición (iv) de la norma de matrices. Como consecuencia, la norma de una matriz y la de un vector se dicen ser compatibles si

$$\|\mathbf{Ax}\| \leq \|\mathbf{A}\|\|\mathbf{x}\|, \quad \mathbf{x} \neq \mathbf{0}. \quad (\text{A.4})$$

## A.4. Normas de una Matriz Subordinada

Sea  $\mathbf{A}$  una matriz de  $(n \times n)$  y  $\mathbf{x}$  un miembro del conjunto  $S$  de los vectores de norma 1, i. e.,  $\mathbf{x} \in S$  si  $\|\mathbf{x}\| = 1$ . En general, la norma del vector  $\|\mathbf{Ax}\|$  varía así como  $\mathbf{x}$  varía,  $\mathbf{x} \in S$ . Ahora supongamos que  $\mathbf{x}_0$  es miembro del conjunto  $S$ , el cual hace que  $\|\mathbf{Ax}_0\|$  alcance su máximo valor. Entonces la norma de la matriz se define por

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|. \quad (\text{A.5})$$

Esta norma de la matriz se dice ser subordinada de la norma del vector y automáticamente satisface la condición de compatibilidad, pues si  $\mathbf{x}_1$  es cualquier otro miembro de  $S$ ,  $\|\mathbf{Ax}_1\| \leq \|\mathbf{Ax}_0\| = \|\mathbf{A}\| = \|\mathbf{A}\|\|\mathbf{x}_1\|$  ya que  $\|\mathbf{x}_1\| = 1$ . De lo anterior se sigue que, para toda norma de matriz subordinada,

$$\|\mathbf{I}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ix}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{x}\| = 1 \quad (\text{A.6})$$

donde  $\mathbf{I}$  es la matriz unitaria. Las normas 1, 2 e  $\infty$  con  $\|\mathbf{x}\| = 1$ , se definen a continuación;

La 1-norma de la matriz  $\mathbf{A}$  es la suma máxima de los módulos de los elementos de columna de la matriz  $\mathbf{A}$

La norma infinita de  $\mathbf{A}$  es la suma máxima de los módulos de los elementos de renglón de la matriz.

La 2-norma de la matriz  $\mathbf{A}$  es la raíz cuadrada del radio espectral de  $\mathbf{A}^H \mathbf{A}$ , donde el radio espectral  $\rho(\mathbf{A}^H \mathbf{A})$  es el módulo mayor de los valores propios de la matriz  $\mathbf{A}^H \mathbf{A}$ .

Cuando  $\mathbf{A}$  es real y simétrica se cumple que  $\mathbf{A}^T = \mathbf{A}$ , entonces

$$\|\mathbf{A}\|_2 = [\rho(\mathbf{A}^2)]^{1/2} = [\rho^2(\mathbf{A})]^{1/2} = \rho(\mathbf{A}) = \max_i |\lambda_i|. \quad (\text{A.7})$$

## A.5. Método de Runge-Kutta

En esta sección se describe el método de Runge-Kutta de orden 4 (RK4). Sea un problema de valor inicial de la siguiente forma

$$y' = f(t, y), y(t_0) = y_0 \quad (\text{A.8})$$

Entonces el método de Runge-Kutta para este problema es dado por la siguiente ecuación:

$$y_{n+1} = y_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) \quad (\text{A.9})$$

donde

$$\begin{aligned} k_1 &= f(t_n, y_n) \\ k_2 &= f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_1\right) \\ k_3 &= f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_2\right) \\ k_4 &= f(t_n + h, y_n + hk_3) \end{aligned}$$

Así, el valor  $y_{n+1}$  es determinado por el valor  $y_n$  más el producto  $h$  (*tamaño del intervalo*) por una pendiente estimada. Dicha pendiente es el promedio de las siguientes pendientes:

$k_1$ , es la pendiente evaluada en el punto inicial del intervalo.

$k_2$ , es la pendiente en el punto medio del intervalo usando el valor de  $k_1$  que determina el valor de  $y$  en el punto  $t_n + h/2$  usando el método de Euler.

$k_3$ , es de nuevo la pendiente en el punto medio del intervalo, pero ahora usando el valor de  $k_2$  el cual determina el valor de  $y$ .

$k_4$ , es la pendiente en el valor final del intervalo y usamos el valor de  $k_3$  para determinar el valor de  $y$ .

Cuando las cuatro pendientes son promediadas, se les da más peso a las pendientes del punto medio

$$p = \frac{k_1 + 2k_2 + 2k_3 + k_4}{6}. \quad (\text{A.10})$$

Que el método RK4 sea de orden cuatro, significa que el error por paso está en el orden de  $h^5$ , mientras que el error total acumulado es de orden  $h^4$ .

## A.6. La Expansión Discreta de Fourier

En muchas aplicaciones prácticas, no podemos implementar métodos numéricos basados sobre series de Fourier. Algunas de las dificultades son: los coeficientes de Fourier para una función arbitraria no se conocen en forma cerrada y deben por lo tanto ser aproximados de alguna manera, que sea eficiente para recuperar en el espacio físico la información que es calculada en el espacio transformado; pero las simples no linealidades son las que conducen a las complicaciones extremas.

La forma de superar estas dificultades es el uso de la transformada discreta de Fourier y las relacionadas series discretas de Fourier.

Para cualquier entero  $N > 0$ , al conjunto de puntos

$$x_j = \frac{2\pi j}{N} \quad j = 0, \dots, N-1, \quad (\text{A.11})$$

se les llama nodos o puntos de la red. Los coeficientes discretos de Fourier de una función valuada en los complejos  $u$  dentro de  $[0, 2\pi]$  con respecto a estos puntos son

$$\tilde{u}_k = \frac{1}{N} \sum_{j=0}^{N-1} u(x_j) e^{-ikx_j} \quad -N/2 \leq k \leq N/2 - 1. \quad (\text{A.12})$$

Debido a la relación de ortogonalidad

$$\frac{1}{N} \sum_{j=0}^{N-1} e^{ipx_j} = \begin{cases} 1 & \text{si } p = Nm, \quad m = 0, \pm 1, \pm 2, \dots \\ 0 & \text{otros casos} \end{cases} \quad (\text{A})$$

tenemos la fórmula de inversión

$$u(x_j) = \sum_{k=-N/2}^{N/2-1} \tilde{u}_k e^{ikx_j} \quad j = 0, \dots, N-1 \quad (\text{A.13})$$

Consecuentemente, el polinomio

$$I_N u(x) = \sum_{k=-N/2}^{N/2-1} \tilde{u}_k e^{ikx} \quad (\text{A.14})$$

es la interpolación trigonométrica de grado  $N/2$  de  $u$  en los nodos (A.12), i.e.,  $I_N u(x_j) = u(x_j)$ ,  $j = 0, \dots, N-1$ . Este polinomio es solo conocido como la serie discreta de Fourier de  $u$ .



Los valores de  $\tilde{u}'_k$ s dependen solo de los  $N$  valores de  $u$  en los nodos (A.12). La transformada discreta de Fourier (DFT) es el mapeo entre los  $N$  números complejos  $u(x_j)$ ,  $j = 0, \dots, N-1$  y los  $N$  números complejos  $\tilde{u}_k$ ,  $k = -N/2, \dots, N/2-1$ . Las dos formas convencionales para la DFT son dadas en (A.12) y (A.13), con la última ecuación a veces referida como el inverso DFT. Éstas muestran que la transformada discreta de Fourier es una transformación ortogonal en  $C^N$ .

El operador de interpolación  $I_N$  puede ser visto como una proyección sobre el espacio  $S_N$  de los polinomios trigonométricos de grado  $N/2$ , con respecto a la aproximación discreta del producto interno

$$(u, v) = \int_0^{(2\pi)} u(x)\overline{v(x)}dx \quad (\text{A.15})$$

Realmente, la forma bilineal

$$(u, v) = \frac{2\pi}{N} \sum_{j=0}^{N-1} u(x_j)\overline{v(x_j)} \quad (\text{A.16})$$

coincide con el producto interno (A.15) si  $u$  y  $v$  son polinomios de grado  $N/2$ .

$$(u, v)_N = (u, v) \quad \forall u, v \in S_N. \quad (\text{A.17})$$

Como una consecuencia, (A.16) es un producto interno sobre  $S_N$ . La interpolación  $I_N u$  de una función continua  $u$  satisface la identidad

$$(I_N u, v)_N = (u, v)_N \quad \forall v \in S_N. \quad (\text{A.18})$$

Los coeficientes discretos de Fourier pueden ser expresados también en términos de los coeficientes de Fourier exactos de  $u$ . Si  $S_u$  ( $S_u$  es la serie de Fourier de  $u$ ) converge a  $u$  en cada nodo (A.12), entonces por (A.13) conseguimos

$$\tilde{u}_k = \hat{u}_k + \sum_{\substack{m=-\infty \\ m \neq 0}}^{+\infty} \hat{u}_{k+Nm} \quad k = -N/2, \dots, N/2 - 1 \quad (\text{A.19})$$

La fórmula (A.19) muestra que el  $k$ -ésimo modo de la interpolación trigonométrica de  $u$  depende no solo del  $k$ -ésimo modo de  $u$ , sino también de todos los modos de  $u$  el cual "alias" el  $k$ -ésimo de la red discreta. La frecuencia  $(k + Nm)^n$  "alias" la frecuencia  $k^n$  sobre la red; estos son indistinguibles en cada nodo ya que  $\phi_{k+Nm}(x_j) = \phi_k(x_j)$ .

## A.7. Solución de un Sistema Tridiagonal con la primer fila distinta de cero

Sea

$$MW = F$$

donde

$$M = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 & \cdots & \cdots & c_N \\ P_1 & q_1 & r_1 & & & & \\ & P_2 & q_2 & r_2 & & & 0 \\ & & \ddots & \ddots & \ddots & & \\ & 0 & & & P_{N-1} & q_{N-1} & r_{N-1} \\ & & & & & P_N & q_N \end{bmatrix}, W = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{N-1} \\ w_N \end{bmatrix}, F = \begin{bmatrix} g \\ f_1 \\ f_2 \\ \vdots \\ f_{N-1} \\ f_N \end{bmatrix}$$

esto es,

$$c_0 w_0 + c_1 w_1 + \cdots + c_N w_N = g \quad (\text{A.20})$$

$$P_i w_{i-1} + q_i w_i + r_i w_{i+1} = f_i \quad i = 1, \dots, N-1 \quad (\text{A.21})$$

$$P_N w_{N-1} + q_N w_N = f_N \quad (\text{A.22})$$

Introduzca la relación de recurrencia [6]:

$$w_{i+1} = x_i w_i + y_i \quad i = 0, \dots, N-1 \quad (\text{A.23})$$

Obtenemos una expresión para  $x_i$  y  $y_i$ . Considere las ecuaciones (A.21) y (A.23):

$$\begin{aligned} P_i w_{i-1} + q_i w_i + r_i w_{i+1} &= f_i \\ w_{i+1} &= x_i w_i + y_i \end{aligned}$$

eliminando  $w_{i+1}$

$$P_i w_{i-1} + q_i w_i + r_i (x_i w_i + y_i) = f_i$$

que re-escribimos como

$$(q_i + r_i x_i) w_i + P_i w_{i-1} = f_i - r_i y_i \quad (\text{A.24})$$

dividiendo entre  $q_i + r_i x_i$  se tiene

$$w_i + \frac{P_i}{q_i + r_i x_i} w_{i-1} = \frac{f_i - r_i y_i}{q_i + r_i x_i} \quad (\text{A.25})$$

de la ec.(A.23), tenemos

$$w_i = x_{i-1} w_{i-1} + y_{i-1} \quad (\text{A.26})$$

Entonces, comparando (A.25) y (A.26) se tiene que:

$$x_{i-1} = \frac{-P_i}{q_i + r_i x_i}, \quad y_{i-1} = \frac{f_i - r_i y_i}{q_i + r_i x_i} \quad i = 1, \dots, N-1. \quad (\text{A.27})$$

Comenzando el cálculo de  $x_i, y_i$  a partir de  $x_{N-1}$  y  $y_{N-1}$  se usa la ecuación (A.27) para el cálculo hacia atrás, y así obtener todas las  $x_i$  e  $y_i$

Para obtener  $x_{N-1}, y_{N-1}$  usamos la ec. (A.22) y (A.23) con  $i = N-1$ :

$$\begin{cases} P_N w_{N-1} + q_N w_N = f_N \\ w_N = x_{N-1} w_{N-1} + y_{N-1} \end{cases}$$

de donde se obtiene

$$w_N = -\frac{P_N}{q_N} w_{N-1} + \frac{f_N}{q_N}. \quad (\text{A.28})$$

Por inspección de (A.28) y (A.26) observamos que

$$x_{N-1} = -\frac{P_N}{g_N} \quad y \quad y_{N-1} = \frac{f_N}{g_N} \quad (\text{A.29})$$

Así, conociendo  $x_{N-1}$  y  $y_{N-1}$ , la ecuación (A.27) nos permite calcular  $x_{N-2}, y_{N-2}, \dots, x_0, y_0$ .

Entonces, encontramos  $w_1, \dots, w_N$  con tal de que el valor de  $w_0$  se conozca. En seguida indicamos cómo realizar el cálculo de  $w_0$ . Considere la ecuación (A.20):

$$c_0 w_0 + c_1 w_1 + \dots + c_{N-1} w_{N-1} + c_N w_N = g.$$

Elimine  $w_1, \dots, w_N$  expresándolas en términos de  $w_0$ , usando la una fórmula de recurrencia de la forma:

$$w_j = \xi_j w_0 + \eta_j \quad j = 0, 1, \dots, N \quad (\text{A.30})$$

Si,  $j = 0$  tenemos  $w_0 = \xi_0 w_0 + \eta_0$ , así,

$$\xi_0 = 1, \quad \eta_0 = 0. \quad (\text{A.31})$$

Ahora, sustituyendo la ecuación (A.30) en la ecuación (A.23), se llega a:

$$w_j = x_{j-1} \xi_{j-1} w_0 + x_{j-1} \eta_{j-1} + y_{j-1}. \quad (\text{A.32})$$

Comparando (A.32) con (A.30) obtenemos

$$\begin{cases} \xi_j = x_{j-1} \xi_{j-1} \\ \eta_j = x_{j-1} \eta_{j-1} + y_{j-1} \end{cases} \quad j = 1, \dots, N$$

Ahora, sustituyendo (A.30) en (A.20)

$$c_0 w_0 + c_1 (\xi_1 w_0 + \eta_1) + \dots + c_N (\xi_N w_0 + \eta_N) = g$$

o, en forma compacta

$$\Xi w_0 + H = g \quad (\text{A.33})$$

donde

$$\Xi = \sum_{j=0}^N c_j \xi_j, \quad H = \sum_{j=1}^N c_j \eta_j \quad (\text{A.34})$$

y finalmente,

$$w_0 = \frac{g - H}{\Xi}. \quad (\text{A.35})$$

Resumiremos el proceso de solución de  $MW = F$  en el siguiente algoritmo

### A.7.1. Algoritmo:

1.  $x_i, y_i$  para  $i = 0, \dots, N - 1$  calculadas de las ecuaciones (A.27) and (A.29).
2.  $w_0$  Calculado de acuerdo a:
  - i)  $\xi_0 = 1, \eta_0 = 0$ .
  - ii)  $\xi_j, \eta_j$  para  $j = 1, \dots, N$  calculados de (A.7).
  - iii)  $\Xi$  and  $H$  calculado de from(A.34)
  - iv)  $w_0$  obtenido de (A.35).
3.  $w_j$  para  $j = 1, \dots, N$  calculado de (A.30).

## A.8. Método de Eliminación de Gauss

Este método lo vamos a explicar utilizando las ecuaciones vistas en el capítulo del Método Implícito.

Cuando en las ecuaciones de Crack-Nicolson (2.18) hay  $N - 1$  puntos internos a lo largo de cada fila, podemos escribirlas de la forma general,

$$\begin{array}{rcl}
+b_1u_1 - c_1u_2 & & = d_1, \\
-a_2u_1 + b_2u_2 - c_2u_3 & & = d_2, \\
& \vdots & \vdots \\
& -a_iu_{i-1} + b_iu_i - c_iu_{i+1} & = d_i, \\
& \dots & \dots \\
& & -a_{n-1}u_{N-2} + b_{N-1}u_{N-1} = d_{N-1}
\end{array}$$

donde las  $a$ ,  $b$ ,  $c$  y  $d$  son conocidas. De la expresión anterior, podemos utilizar la primer ecuación junto con la segunda ecuación para eliminar a  $u_1$  obteniendo una nueva ecuación y esta usarla junto con la ecuación tres para eliminar a  $u_2$  resultándonos otra ecuación la cual junto con la ecuación cuatro nos sirve para elminimar a  $u_3$  y así sucesivamente, hasta que la última ecuación que obtengamos quede solo en función de  $u_{N-1}$ .

Notemos que el coeficiente  $c$  en cada nueva ecuación es el mismo que en la ecuación anterior correspondiente. Entonces, después de que realizamos una eliminación resulta,

$$\begin{array}{rcl}
\alpha_{i_1} u_{i-1} - c_{i-1} u_i & = & S_{i-1}, \\
-a_i u_{i-1} + b_i u_i - c_i u_{i+1} & = & d_i,
\end{array} \tag{A.36}$$

donde  $\alpha_1 = b_1$ ,  $S_1 = d_1$ . Al eliminar  $u_{i-1}$  llegamos a

$$\left( b_i - \frac{a_i c_{i-1}}{\alpha_{i-1}} \right) u_i - c_i u_{i+1} = d_i + \frac{\alpha_i S_{i-1}}{\alpha_{i-1}}$$

es decir,

$$\alpha_i u_i - c_i u_{i+1} = S_i, \quad (\text{A.37})$$

donde

$$\alpha_i = b_i \frac{a_i c_{i-1}}{\alpha_{i-1}} \quad y \quad S_i = d_i + \frac{\alpha_i S_{i-1}}{\alpha_i - 1} \quad (i = 2, 3, \dots).$$

El último par de ecuaciones son:

$$\begin{aligned} \alpha_{N-2} u_{N-2} - c_{N-2} u_{N-1} &= S_{N-2} \\ -a_{N-1} u_{N-2} + b_{N-1} u_{N-1} &= d_{N-1}. \end{aligned} \quad (\text{A.38})$$

Eliminando  $u_{N-2}$  nos da,

$$\left( b_{N-1} - \frac{a_{N-1} c_{N-2}}{\alpha_{N-2}} \right) u_{N-1} = d_{N-1} + \frac{a_{N-1} S_{N-2}}{\alpha_{N-2}},$$

es decir,

$$\alpha_{N-1} u_{N-1} = S_{N-1}. \quad (\text{A.39})$$

Las ecuaciones (A.37) y (A.39) muestran que la solución puede ser calculada de

$$\begin{aligned} u_{N-1} &= \frac{S_{N-1}}{\alpha_{N-1}} \\ u_i &= \frac{1}{\alpha_i} (S_i + c_i u_{i+1}) \quad (i = N-2, N-3, \dots, 1) \end{aligned} \quad (\text{A.40})$$

donde las  $\alpha$  y  $S$  son dadas por:



$$\begin{aligned}\alpha_1 &= b_1; & \alpha_i &= b_i - \frac{a_i}{\alpha_{i-1}}, \\ S_1 &= d_1; & S_i &= d_i + \frac{a_i}{\alpha_{i-1}}S_{i-1}.\end{aligned}\tag{A.41}$$

En muchos problemas  $\alpha_i$  y  $a_i/\alpha_{i-1}$  son independientes del tiempo y solamente son calculados una vez, independientemente del paso del tiempo.

# Bibliografía

- [1] "Física Vol. 1". Robert Resnick, CECSA, Tercera Edición (1996).
- [2] "Campos Electromagnéticos". Roald Wangsness, LIMUSA (1999).
- [3] "Numerical Solution of Partial Differential Equations *Finite Difference Methods*". G. D. Smith, OXFORD, Third Edition (1999).
- [4] "The Wave Equation in  $1+1$ ". P. Diener, F. S. Guzmán, <http://numrel.aet.mpg.de/Education/Tutorials/>, (September 11, 2003).
- [5] "A Book on C". Al Kelley, Ira Pohl, Addison-Wesley, Fourth Edition (1998)
- [6] "Spectral Methods and Application to Convection". Roger Peyter, University Nice, Center for Fluid Mechanics Lecture Serie, Brown University, Providence (1988)
- [7] "Spectral Methods, Fundamentals in Single Domains" C. Canuto, M.Y. Hussaini, A. Quarteroni, T. Zang, Springer-Verlag (2006).
- [8] "On a quasi-linear parabolic equation occurring in aerodynamics" J. D. Cole. *Quart. Appl. Math.* vol. 9, pags. 225-236 (1951).
- [9] E.R. Benton and G.N. Platzmann, "A table of solutions of one-dimensional Burgers equation" *Quart. Appl. Math.* vol. 29, pags 195-212 (1972).

- [10] "Numerical Recipes in C" William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery, Cambridge University Press, Second Edition (1997).
- [11] "Spectral and Finite Difference Solution of the Burgers Equation" C.Basdevant et.al. Computers and Fluids vol 14, pags. 23-41 (1986).