

Solución de la ecuación de onda como un problema de valores iniciales usando diferencias finitas

F.S. Guzmán

Instituto de Física y Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo, Edificio C-3, Cd. Universitaria, 58040 Morelia, Michoacán, México.

Recibido el 25 de junio de 2009; aceptado el 13 de abril de 2010

Se presenta la solución de la ecuación de onda como ejemplo paradigmático de la solución de problemas de valores iniciales con condiciones de frontera usando la aproximación de diferencias finitas. Primero se desarrolla una solución elemental y una discretización directa a manera de introducción. Posteriormente se resuelve la ecuación de onda como un sistema de primer orden, se estudia la hiperbolicidad del sistema de ecuaciones resultante, se calculan los modos y velocidades características del sistema y se imponen condiciones de frontera en términos de las variables características. Se adopta el método de líneas como esquema de evolución. Además se hace especial énfasis en que los resultados numéricos necesitan un criterio de validez. En el caso de la aproximación con diferencias finitas de una ecuación diferencial parcial se presenta la convergencia a una solución correcta en el límite continuo. Finalmente, se espera que este trabajo sirva de guía para la correcta solución de problemas de valores iniciales con condiciones de frontera en general.

Descriptor: Métodos de diferencias finitas; técnicas computacionales; ecuación de onda.

The solution of the wave equation is presented as the paradigm of the solution of initial value problems with boundary conditions using the finite differences approximation. First, it is developed an elementary solution and a direct discretization in order to introduce the method. Second, the wave equation is solved as a system of first order, the hyperbolicity properties of the resulting system of equations is studied, the characteristic variables and characteristic speeds of the system are calculated and boundary conditions are imposed in terms of the characteristic variables. In this case the method of lines is used as the evolution scheme. Special attention is devoted to the fact that numerical calculations require a criterion to be valid. In the case of the approximation using finite differences of a partial differential equation, the convergence to a correct solution in the continuum limit is presented as such criterion. Finally, it is expected that this manuscript serves as a guide to solve correctly other initial value problems with boundary conditions.

Keywords: Finite differences method; computing techniques; wave equation.

PACS: 02.60.Bf; 02.70.-c

En este trabajo se presenta la solución de la ecuación de onda como la solución de un problema de valores iniciales con condiciones de frontera usando una aproximación en diferencias finitas. Se considera dicho caso como el ejemplo emblemático de la solución de problemas de evolución de sistemas hiperbólicos de ecuaciones. Adicionalmente se provee un código computacional simple que funciona, que es capaz de reproducir los resultados en este artículo y se espera que el lector pueda aplicar las ideas aprendidas en problemas más complicados relacionados con la solución de ecuaciones diferenciales parciales (EDPs) asociadas con problemas de valores iniciales de su interés.

La aproximación en diferencias finitas está basada en la discretización del dominio en el que se define el sistema de EDPs. Además de presentar dicha herramienta se muestra el tipo de pruebas que debe cumplir la solución numérica de EDP. El caso de la ecuación de onda se estudia de dos maneras:

- i) la primera utilizando una discretización simple que usa tres niveles temporales, con la cual se ilustra el funcionamiento de la aproximación en diferencias finitas, y
- ii) la segunda, en la que la ecuación de onda se descompone en un sistema de ecuaciones de primer orden en el espacio y en el tiempo, cuya solución se construye

utilizando el método de líneas con dos niveles temporales.

El segundo caso además se resuelve para sistemas de coordenadas espacio-temporales generales, se estudia la hiperbolicidad del sistema de ecuaciones de primer orden, se construyen las variables características y las velocidades características del sistema y se imponen condiciones de frontera usando las variables características. Se presenta la solución de la ecuación de onda en coordenadas muy generales, con la intención de mostrar cómo los fenómenos que son invariantes bajo cambios de coordenadas (por ejemplo los relacionados con los sistemas relativistas) pueden controlarse numéricamente si se eligen las coordenadas de manera apropiada. Para mejor entendimiento y para ejercitar los conceptos presentados aquí, los códigos utilizados se encuentran disponibles públicamente [1]. De hecho aquí se ilustran las estrategias que se usan para resolver problemas como el de la evolución de agujeros negros, en los que la elección de coordenadas se hace incluso de manera dinámica [2], pero también permite estudiar problemas de tipo onda en espacio-tiempos generales, como es el caso de la propagación de la radiación gravitacional, que obedece ecuaciones tipo onda no homogéneas.

Las herramientas que se presentan se aplican a gran variedad de problemas de la física teórica actual, como es el caso de la solución de ecuaciones diferenciales parciales de

la relatividad numérica [2], la ecuación de Schrödinger dependiente del tiempo asociada a condensados de Bose en la aproximación de campo medio o gases atómicos [3], los problemas asociados con la evolución de fluidos que obedecen las ecuaciones de Euler [4], entre otros también de gran importancia.

La solución de la ecuación de onda homogénea planteado en variables de primer orden es de la mayor importancia y relevancia por una sencilla razón: el análisis de hiperbolicidad que decide si un problema de valores iniciales es bien planteado o no se reduce al estudio de la parte principal del sistema de ecuaciones de primer orden (como se ilustra en este trabajo). Entonces, las ecuaciones de tipo onda que contienen el operador de onda, y alguna función de la función de onda, e incluso algunos términos de fuentes, tienen la estructura característica de la ecuación de onda homogénea. Por tanto, las herramientas presentadas aquí sirven para resolver una gran cantidad de problemas semejantes.

Por otra parte, existe en la actualidad la preocupación de que en la mayoría de los casos, cuando se presentan resultados numéricos, no se muestran pruebas de convergencia. Es por ello que se presenta de manera detallada el método para verificar la convergencia de los resultados calculados usando diferencias finitas.

El trabajo está organizado de la manera siguiente. En la Sec. 1 se describen los métodos numéricos relacionados con la aplicación de las diferencias finitas a la solución de EDPs. En la Sec. 2 se ejemplifica el funcionamiento de la aproximación en diferencias finitas con la ecuación de onda y una discretización simple. En la Sec. 3 se presenta la formulación de primer orden de la ecuación de onda en el espacio-tiempo de 1+1 dimensiones y la hiperbolicidad del sistema. En la Sec. 4 se presenta la solución de la ecuación de onda no homogénea correspondiente a la ecuación de Klein-Gordon con soluciones solitónicas, conocida como ecuación de Sine-Gordon. Finalmente, en la Sec. 5 se mencionan algunos comentarios finales.

1. Introducción a las diferencias finitas

La solución de alguna EDP usando la aproximación con diferencias finitas consiste en definir una versión discreta de la EDP, y además en calcular una solución a dicha aproximación sobre un dominio discreto contenido en el dominio de la EDP en el continuo. En esta sección se presentan los elementos necesarios para formular la versión discreta de una EDP, aunque nos restringimos al caso de problemas de valores iniciales con sistemas hiperbólicos de ecuaciones y se deja de lado la solución de ecuaciones elípticas.

Debido a que se resuelve la versión discreta de la EDP original es necesario indicar los criterios de validez de dichas soluciones, es decir, la relación que deben guardar tales soluciones numéricas de una ecuación que es solamente una aproximación de la EDP original con la solución de la ecuación original. Por ello se indica con claridad el procedimiento

para llevar a cabo pruebas de convergencia cuando se resuelve una EDP con el método de aproximación con diferencias finitas.

1.1. Ingredientes de la aproximación discreta de las EDPs en el continuo

Primer ingrediente

La aproximación en diferencias finitas de una ecuación diferencial parcial (EDP) consiste en definir las funciones y variables involucradas en un conjunto discreto de puntos del dominio donde se busca una solución de la ecuación.

Con el fin de ilustrar este concepto, supóngase que una EDP que involucra la función f está definida en un dominio que tiene dos coordenadas t y x (por ejemplo, el tiempo y una coordenada espacial). Una manera de discretizar la ED consiste en considerar que las variables y funciones de la ED están definidas en un conjunto de puntos tanto en el espacio $x_j = j\Delta x$ como en el tiempo $t^n = n\Delta t$, donde j y n son enteros que etiquetan los puntos donde estará definida la ecuación. Así, la función f queda definida solamente en los puntos del dominio (t^n, x_j) y denotamos dichos valores de la función por f_j^n . Se considera además que $j = 0, 1, \dots, N$ y $n = 0, 1, \dots, N_t$. Limitar los valores de las etiquetas j y n ilustra el hecho de que se ha elegido un sominio finito donde se calculará la solución. La razón principal es que si a cada elemento de la malla se asigna un valor de cada variable independiente y de cada función involucrada en la EDP, es necesario elegir N de manera que la memoria de la computadora sea suficiente para asignar valores a cada variable y función en cada x_j ; los algoritmos descritos a continuación evitan la necesidad de asignar también memoria a todas las variables y funciones involucradas para los N_t niveles de tiempo que cubren el dominio en que se desea resolver la ecuación y bastará con asignar memoria a variables y funciones para solamente dos o tres niveles de tiempo a la vez. Los valores x_0 y x_N corresponden a los valores que delimitan el dominio de x y de manera análoga los valores t^0 y t^{N_t} delimitan el dominio de t .

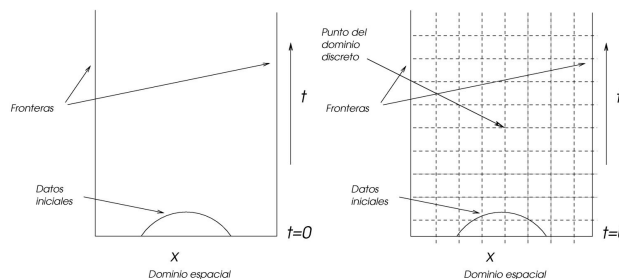


FIGURA 1. En la primera figura se muestra el dominio de un problema de valores iniciales con condiciones de frontera y en la segunda la versión discreta del dominio que define una malla sobre la que quedan definidas las funciones involucradas en la EDP.

Este procedimiento de discretización define una malla en el dominio, en cuyos nodos quedan definidas las funciones involucradas en la EDP. La malla más elemental es aquella cuyos nodos se encuentran igualmente espaciados, cuya gráfica se muestra en la Fig. 1. En este trabajo se considera solamente el caso en que la discretización es de este tipo.

El criterio para determinar una discretización queda al arbitrio de quien resuelve un problema, o de las propiedades de la ecuación diferencial que se desea resolver. Para la discretización uniformemente espaciada en nuestro caso, se definen las resoluciones espacial $\Delta x = x_{j+1} - x_j$ y temporal $\Delta t = t^{n+1} - t^n$ y se dice que la resolución espacial y temporal son mayores cuando Δx y Δt son menores.

Segundo ingrediente

Una vez definido el dominio discreto de la EDP, es necesario aproximar la ecuación diferencial, es decir, los operadores diferenciales que aparecen en dicha ecuación usando solamente los valores de las funciones involucradas que están definidos en la malla. Una condición que se demanda es que dichas funciones sean analíticas, lo cual implica que las funciones involucradas tienen definida una expansión en serie de Taylor, lo cual a su vez permite construir aproximaciones de los operadores diferenciales.

Derivada de primer orden. Dada f_j^n , una función como la definida antes con n fijo, es posible calcular valores aproximados de la función en los puntos adyacentes a partir de expansiones en serie de Taylor truncadas de f en torno a f_j^n . El número de puntos cercanos a x_j en los que se expanda la serie de Taylor y el orden del error de truncado de ésta, determinarán el orden de precisión de la aproximación. Para ilustrar la construcción de la aproximación de segundo orden se usarán las siguientes tres expansiones con error de truncado $O(\Delta x^3)$:

$$\begin{aligned} f(x_{j-1}) &= f(x_j) - \Delta x f'(x_j) + \frac{\Delta x^2}{2} f''(x_j) + O(\Delta x^3), \\ f(x_j) &= f(x_j), \\ f(x_{j+1}) &= f(x_j) + \Delta x f'(x_j) + \frac{\Delta x^2}{2} f''(x_j) + O(\Delta x^3), \end{aligned} \tag{1}$$

donde la prima denota derivada con respecto a x . A partir de estas aproximaciones es posible construir diferentes operadores diferenciales para las derivadas de f_j^n . Por ejemplo, sumando la primera y la tercera expresiones, y dividiendo por $2\Delta x$, se obtiene la expresión para la primera derivada en el punto x_j con un error de segundo orden:

$$f'(x_j) = \frac{f(x_{j+1}) - f(x_{j-1})}{2\Delta x} + O(\Delta x^2). \tag{2}$$

Conviene observar que para calcular esta aproximación es necesario conocer los valores f_{i+1}^n y f_{i-1}^n , por lo que esta aproximación se llama centrada. Un comentario adicional es conveniente aquí: la expresión (2) es la definición de derivada en cálculo de una variable en el límite cuando $\Delta x \rightarrow 0$.

Es posible calcular la derivada de primer orden usando aproximaciones no centradas, desbalanceadas, usualmente llamadas *upwind*. Para tal efecto basta considerar las expansiones siguientes:

$$\begin{aligned} f(x_j) &= f(x_j), \\ f(x_{j+1}) &= f(x_j) + \Delta x f'(x_j) + \frac{\Delta x^2}{2} f''(x_j) + O(\Delta x^3), \\ f(x_{j+2}) &= f(x_j) + 2\Delta x f'(x_j) + \frac{4\Delta x^2}{2} f''(x_j) + O(\Delta x^3), \end{aligned} \tag{3}$$

donde ahora se ha hecho la expansión hasta dos puntos a la derecha de x_j . Una combinación que cancela los términos que contienen la segunda derivada y el término de orden cero es $f(x_{j+2}) - 4f(x_{j+1}) + 3f(x_j)$, que al ser dividida por $2\Delta x$ resulta en una expresión para la derivada de f en x_j usando solamente valores de f en los puntos x_{j+1} y x_{j+2} :

$$f'(x_j) = \frac{-f(x_{j+2}) + 4f(x_{j+1}) - 3f(x_j)}{2\Delta x} + O(\Delta x^2). \tag{4}$$

De manera análoga, la construcción de la aproximación de la derivada que usa solamente puntos a la izquierda de x_j es la siguiente:

$$f'(x_j) = \frac{f(x_{j-2}) - 4f(x_{j-1}) + 3f(x_j)}{2\Delta x} + O(\Delta x^2). \tag{5}$$

Para calcular operadores con un orden de aproximación mayor basta con generalizar el procedimiento ilustrado anteriormente. Supóngase que se desea calcular la derivada de primer orden con un error de cuarto orden, entonces basta con considerar el siguiente conjunto de series truncadas:

$$\begin{aligned} f(x_{j-2}) &= f(x_j) - 2\Delta x f'(x_j) + \frac{4\Delta x^2}{2} f''(x_j) \\ &\quad - \frac{8\Delta x^3}{6} f'''(x_j) + \frac{16\Delta x^4}{24} f''''(x_j) + O(\Delta x^5), \\ f(x_{j-1}) &= f(x_j) - \Delta x f'(x_j) + \frac{\Delta x^2}{2} f''(x_j) \\ &\quad - \frac{\Delta x^3}{6} f'''(x_j) + \frac{\Delta x^4}{24} f''''(x_j) + O(\Delta x^5), \\ f(x_j) &= f(x_j), \\ f(x_{j+1}) &= f(x_j) + \Delta x f'(x_j) + \frac{\Delta x^2}{2} f''(x_j) \\ &\quad + \frac{\Delta x^3}{6} f'''(x_j) + \frac{\Delta x^4}{24} f''''(x_j) + O(\Delta x^5), \\ f(x_{j+2}) &= f(x_j) + 2\Delta x f'(x_j) + \frac{4\Delta x^2}{2} f''(x_j) \\ &\quad + \frac{8\Delta x^3}{6} f'''(x_j) + \frac{16\Delta x^4}{24} f''''(x_j) + O(\Delta x^5). \end{aligned}$$

Para calcular la primera derivada de f en x_j es necesario encontrar la combinación lineal correcta del tipo $af(x_{j-2}) + bf(x_{j-1}) + cf(x_j) + df(x_{j+1}) + ef(x_{j+2})$ tal

que los coeficientes de $f(x_j)$ y de las derivadas de segundo, tercero y cuarto orden sean cero. En el caso centrado la expresión es la siguiente:

$$f'(x_j) = \frac{1}{12\Delta x} [f(x_{j-2}) - 8f(x_{j-1}) + 8f(x_{j+1}) - f(x_{j+2})] + O(\Delta x^4), \quad (6)$$

y para combinaciones desbalanceadas se procede usando las expansiones truncadas para distintos puntos vecinos de x_j ; los resultados de varias elecciones de puntos vecinos se resumen en la Tabla I.

Derivada de segundo orden. Para obtener una aproximación de la segunda derivada de una función respecto a x en x_j , se procede nuevamente mediante la búsqueda de una combinación de las expansiones en serie de Taylor en distintos puntos vecinos de x_j , pero en esta ocasión imponiendo la condición de que los coeficientes de las derivadas de orden cero, orden uno, tercero y superiores sean cero. Procediendo de dicha manera, la derivada de segundo orden con una aproximación de segundo orden puede estimarse a partir de las expresiones (1) desarrolladas hasta orden $O(\Delta x^4)$, con la combinación $f(x_{j+1}) - 2f(x_j) + f(x_{j-1})$ dividida por Δx^2 , o sea,

$$f''(x_j) = \frac{f(x_{j+1}) - 2f(x_j) + f(x_{j-1}))}{\Delta x^2} + O(\Delta x^2), \quad (7)$$

que resulta ser una aproximación centrada. Así pues, el algoritmo para construir aproximaciones en diferencias finitas de operadores diferenciales es el siguiente:

- i) Hacer la expansión en serie de Taylor de f en puntos cercanos a aquel donde se quiere evaluar la derivada de f .
- ii) Encontrar una combinación lineal de tales expansiones que cumpla con la condición de que los coeficientes de las derivadas de orden superior e inferior a la derivada del orden deseado sean cero.

Una vez conocido el procedimiento para calcular aproximaciones de derivadas de funciones, basta con mencionar los coeficientes de las combinaciones lineales necesarias para calcular los operadores diferenciales de primero y segundo orden, centrados y desbalanceados con distintos órdenes de precisión. En en la Tabla I se muestran los coeficientes de las funciones evaluadas en los distintos puntos vecinos de x_j en las expresiones para el cálculo de las derivadas de primer orden con precisión de segundo orden y los coeficientes cuando la precisión es de cuarto orden. En la Tabla II se muestran aquellos coeficientes para las derivadas de segundo orden.

TABLA I. Coeficientes de la aproximación de la derivada de primer orden usando distintas combinaciones de puntos vecinos. Las aproximaciones con segundo orden de precisión significan $f' = 1/(2\Delta x)[...] + O(\Delta x^2)$ y $f' = 1/(12\Delta x)[...] + O(\Delta x^4)$ para las aproximaciones con cuarto orden de precisión. Las expresiones (2), (4), (5) y (6) ilustran el uso de esta tabla.

	x_{j-4}	x_{j-3}	x_{j-2}	x_{j-1}	x_j	x_{j+1}	x_{j+2}	x_{j+3}	x_{j+4}
Segundo orden				1	0	-1			
			1	-4	3				
					-3	4	-1		
Cuarto orden	3	-16	36	-48	25				
		-1	6	-18	10	3			
			1	-8	0	8	-1		
				-3	-10	18	-6	1	
					-25	48	-36	16	-3

TABLA II. Coeficientes de la aproximación de la derivada de segundo orden usando distintas combinaciones de puntos vecinos. Las aproximaciones con segundo orden de precisión son del tipo $f'' = 1/(\Delta x^2)[...] + O(\Delta x^2)$, mientras que las aproximaciones de cuarto orden de precisión van como $f'' = 1/(12\Delta x^2)[...] + O(\Delta x^4)$. La expresión (7) ejemplifica el uso de esta tabla.

	x_{j-5}	x_{j-4}	x_{j-3}	x_{j-2}	x_{j-1}	x_j	x_{j+1}	x_{j+2}	x_{j+3}	x_{j+4}	x_{j+5}
Segundo orden					1	-2	1				
			-1	4	-5	2					
						2	-5	4	-1		
Cuarto orden	-10	61	-156	214	-154	45					
		1	-6	14	-4	-15	10				
				-1	16	-30	16	-1			
					10	-15	-4	14	-6	1	
						45	-154	214	-156	61	-10

Los desarrollos hasta el momento ilustran las aproximaciones en diferencias finitas de los operadores diferenciales de una EDP. El hecho de haber elegido un dominio espacio-temporal, permite considerar directamente la discretización de la ecuación de onda:

$$\partial_{tt}\phi - \partial_{xx}\phi = 0, \tag{8}$$

donde se ha elegido la notación con la que un subíndice indica una derivada con respecto a la variable representada con el índice. Originalmente puede partirse de la ecuación $\partial_{tt}\phi - v^2\partial_{xx}\phi = 0$ donde v es la velocidad de propagación de la onda, pero se usará $v = 1$, lo que equivale a reescalar la coordenada temporal $t \rightarrow v\bar{t}$ para que la ecuación resultante sea (8); dado el caso en que $v \neq 1$, el parámetro t de la solución calculada de la Ec. (8) $\phi = \phi(x, t)$ deberá reescribirse como $\phi = \phi(x, v\bar{t})$ en términos del parámetro temporal de interés \bar{t} .

Para construir la aproximación en diferencias finitas (DF) de (8) basta con escribir la aproximación de derivadas de segundo orden y construir la aproximación de la derivada temporal, lo cual se consigue intercambiando x_j por t^n en las expresiones y cuadros anteriores. El resultado es el siguiente cuando se usan operadores con aproximación de segundo orden, que centrados en el punto (t^n, x_j) dan como resultado:

$$\frac{\phi_j^{n+1} - 2\phi_j^n + \phi_j^{n-1}}{(\Delta t)^2} - \frac{\phi_{j+1}^n - 2\phi_j^n + \phi_{j-1}^n}{(\Delta x)^2} = O(\Delta x^2, \Delta t^2), \tag{9}$$

para todo $j = 1, \dots, N - 1$ y $n = 1, \dots, N_t - 1$, los casos en que $j = 0, N$ y $n = 0, N_t$ se discuten enseguida al momento de imponer condiciones de frontera y condiciones iniciales. En este punto es preciso notar que la aproximación con diferencias finitas ha introducido un error, es decir, el lado derecho de la ecuación no es necesariamente cero, sino que es del orden $O(\Delta x^2, \Delta t^2)$. La certeza de que tal ocurre determina el tercer ingrediente.

Tercer ingrediente

Al usar este método numérico es necesario tener en mente que la aproximación en DF provee de una versión aproximada de la ecuación en el continuo y que la ecuación que se resolverá en adelante es solamente una aproximación de la original, es decir, la que es definida en el dominio continuo. Es preciso entonces establecer criterios que indiquen que cuando se resuelve la versión discreta de una EDP, al menos en el límite continuo se está calculando la solución correcta.

1.2. Convergencia

Hemos mostrado que las aproximaciones de los operadores diferenciales son solamente una aproximación con un error asociado de cierto orden (segundo y cuarto orden en los ejemplos mostrados en las Tablas I y II). Cuanto mayor sea la resolución con que se construya la malla (valores más pequeños

de Δx y Δt), cuanto menor será el error con que se están aproximando los operadores diferenciales a sus contrapartes en el continuo, y por tanto la aproximación en diferencias finitas de una ecuación diferencial es más precisa. Siendo que se resuelve la versión aproximada de una EDP, es necesario verificar si dicha solución converge a la solución de la ecuación en el continuo.

Para mostrar un criterio de convergencia consideremos como ejemplo la Ec. (9). Como veremos enseguida, es útil definir la razón $\Delta t = C\Delta x$, donde C es constante, o sea, que los errores en ambas variables son proporcionales, entonces si llamamos a la parte derecha de (9)

$$Error_1 = O(\Delta x^2, \Delta t^2) = E\Delta x^2$$

con E un coeficiente de error en la aproximación de las diferencias finitas con respecto a la ecuación continua. Si ahora se define una nueva resolución $\Delta x_\kappa = \Delta x/\kappa$ con $\kappa > 1$ se tiene que $Error_\kappa = E\frac{\Delta x^2}{\kappa^2}$. De tales expresiones se tiene que

$$\frac{Error_1}{Error_\kappa} = \kappa^2, \tag{10}$$

es decir, si la resolución aumenta por un factor de κ , la aproximación de la versión discreta se aproxima a la versión continua de manera cuadrática con el aumento de la resolución. El hecho de que sea cuadrática es consecuencia de que la aproximación que se construyó involucra errores de segundo orden. Si la aproximación se hiciera con errores de cuarto orden, la versión discreta se aproximaría a la continua de manera cuártica y así para aproximaciones de mayor precisión.

Para ilustrar el concepto de convergencia en diferencias finitas de manera más general, considérese una función f_t que es solución numérica de una EDP a un tiempo dado, y que ha sido construida bajo la discretización de dicha ecuación con una aproximación de segundo orden. Suponiendo además que se conoce la solución exacta $f_0(x)$, el resultado numérico puede escribirse en la forma

$$f(x) = f_0(x) + E(x)\Delta x^2 + O(\Delta x^3),$$

donde E denota un coeficiente del error. Dado que se conoce la solución exacta, es posible conocer el error con que se calcula la solución numérica usando distintos valores de Δx . Sean f_1 y f_2 dos soluciones numéricas calculadas usando las resoluciones Δx y $\Delta x/2$ respectivamente ($\kappa = 2$ en (10)). La razón entre los errores es la siguiente:

$$\frac{f_1 - f_0}{f_2 - f_0} = \frac{\Delta x^2 + O(\Delta x^3)}{\frac{1}{4}\Delta x^2 + O(\Delta x^3)} = 4 + O(\Delta x). \tag{11}$$

En este caso se ha supuesto que la función f depende solamente de la variable x . El número cuatro en (11) se llama factor de convergencia y debe ser evaluado en cada punto de la malla donde se ha calculado la función f . Cuando en un cálculo numérico que se ha llevado a cabo a partir de una aproximación de segundo orden el factor de convergencia es $4 = 2^2$ en (10), se dice que la solución converge con segundo

orden, es decir, cumple con lo predicho por la teoría en (11). De manera análoga, es posible mostrar que si la aproximación es de cuarto orden, el factor será $16 = 2^4$.

Por ejemplo, para la ecuación de onda (8) la solución exacta es conocida $\phi(x, t) = f(x+t) + g(x-t)$ con f y g dadas por las condiciones iniciales del problema, y bastará calcular la solución numérica con dos resoluciones distintas para determinar si los algoritmos utilizados producen una solución que converge correctamente a la exacta. Podría parecer que lo mencionado en estos párrafos es elemental, pero el siguiente ejemplo ilustra la importancia del factor de convergencia: supóngase que se discretiza el dominio $x \in [0, 1]$ con la opción (i) $\Delta x = 0.1$ y con $\Delta x = 0.05$ para estudiar el factor de convergencia en el caso de una ED elemental, es muy probable que la solución numérica no tenga factor de convergencia cuatro, pero es probable que lo tendrá cuando se usan resoluciones (ii) $\Delta x = 0.01$ y $\Delta x = 0.005$; si éste es el caso, se dice que los cálculos en la opción (ii) han sido ejecutados en el régimen de convergencia, mientras la opción (i) no aporta condiciones para que se obtenga la convergencia deseada. Un caso en que es posible visualizar esta situación es el que corresponde a un perfil inicial gaussiano de la función de onda de ancho 0.1 centrada en $x = 0.5$, entonces, en el caso (i) se pretende resolver dicho perfil con una malla de solamente diez y veinte puntos y el ancho de la campana con solamente tres y seis puntos, lo que hace imposible incluso visualizar que el perfil es gaussiano. Usando las resoluciones del caso (ii) es posible obtener cálculos suficientemente precisos. Es posible hacer estos experimentos usando el código [1].

Regularmente los cálculos que no convergen con el orden deseado se desechan (o debieran desecharse), mientras que aquellos realizados en el régimen de convergencia son aceptables. *Este ejemplo ilustra claramente el hecho de que no toda solución numérica calculada con la aproximación de DF es aceptable, y que ningún cálculo que no converja debe gozar de credibilidad.*

Para el caso en que se desconoce la solución exacta es posible hacer un estudio de convergencia usando los resultados numéricos calculados con tres distintas resoluciones (no dos como en el caso anterior) llamado estudio de autoconvergencia. Si además de f_1 y f_2 definidas antes se calcula la solución f_3 de la EDP con resolución $\Delta x/4$ se puede calcular la razón siguiente:

$$\begin{aligned} \frac{f_1 - f_2}{f_2 - f_3} &= \frac{\Delta x^2 - \frac{1}{4}\Delta x^2 + O(\Delta x^3)}{\frac{1}{4}\Delta x^2 - \frac{1}{16}\Delta x^2 + O(\Delta x^3)} \\ &= \frac{\Delta x^2 + O(\Delta x^3)}{\frac{1}{4}\Delta x^2 + O(\Delta x^3)} = 4 + O(\Delta x), \end{aligned} \quad (12)$$

donde una vez más el resultado se llama factor de convergencia, y una vez más resulta ser cuatro cuando la aproximación de la EDP es de segundo orden. De igual modo, si la aproximación fuera de cuarto orden el factor sería 16 en lugar de 4.

Finalmente, es útil saber que cuando el factor de convergencia es menor que cuatro, es necesario aceptar varias posibilidades:

- (a) hay un error en la implementación del programa,
- (b) los algoritmos no permiten la convergencia en el rango Δx elegido, es decir, no se han hecho los cálculos en el régimen de convergencia y es necesario ensayar con otros valores de Δx .

2. La ecuación de onda con una discretización simple

2.1. El problema de valores iniciales y su versión discreta

El problema de valores iniciales a resolver es entonces el siguiente:

$$\begin{aligned} \partial_{tt} - \partial_{xx} &= 0, \\ \phi(x, 0) &= \phi_0(x), \quad \partial_t \phi(x, 0) = \dot{\phi}_0(x), \\ \phi(-1, t) &= \phi_L(t), \quad \phi(1, t) = \phi_R(t), \\ x &\in [-1, 1], \quad t > 0. \end{aligned} \quad (13)$$

Para construir la solución de una EDP que admite por solución global la evolución de datos iniciales es necesario proveer dichos datos iniciales. En el caso de la ecuación de onda, la solución numérica más simple consiste en evolucionar los datos iniciales usando la Ec. (9). De hecho, dados los valores de la función de onda en toda la malla al tiempo n y $n-1$, es posible construir los valores de ϕ al tiempo $n+1$. Para conseguirlo basta con resolver ϕ_j^{n+1} en (9), lo cual implica

$$\begin{aligned} \phi_j^{n+1} &= \left(\frac{\Delta t}{\Delta x}\right)^2 [\phi_{j+1}^n - 2\phi_j^n + \phi_{j-1}^n] \\ &+ 2\phi_j^n - \phi_j^{n-1} + O(\Delta x^2, \Delta t^2), \end{aligned} \quad (14)$$

que permite conocer los valores de la función de onda para todos los puntos de la malla al tiempo $n+1$, excepto aquellos que están en los bordes $x_0 = -1$ y $x_N = +1$, debido a que este algoritmo requiere los valores de un punto a la derecha y otro a la izquierda para todo x_j en la malla. Para ilustrarlo, en la Fig. 2 se muestra el elemento contenido en la expresión (14), llamado molécula del algoritmo de evolución. En tal diagrama aparece como un círculo negro la posición del dominio discreto (t^{n+1}, x_j) donde se puede calcular ϕ en términos de los valores de ϕ en las posiciones indicadas con círculos blancos. Además, esta figura ilustra la necesidad de conocer el valor de la función de onda en todos los puntos del espacio para t^n y t^{n-1} , y en especial los valores de la función en las posiciones (t^n, x_{-1}) y (t^n, x_{N+1}) que no están definidos en la malla y corresponden al caso en que el círculo negro se localiza en (t^{n+1}, x_0) y (t^{n+1}, x_N) . Afortunadamente los valores ϕ_0^{n+1} y ϕ_N^{n+1} pueden ser calculados imponiendo una condición de frontera.

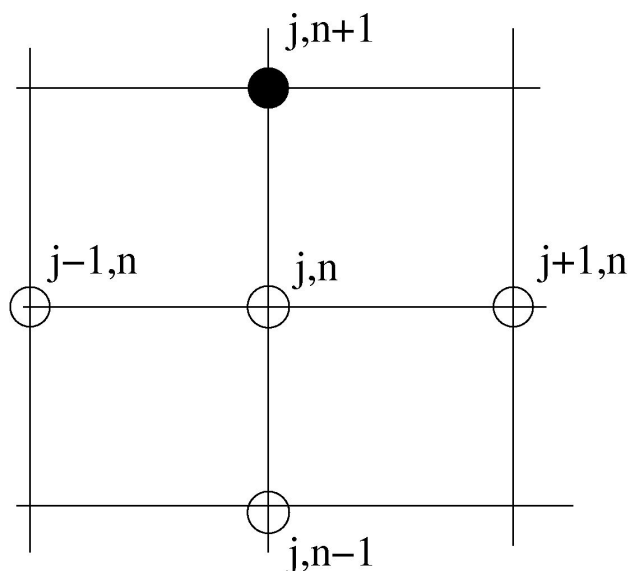


FIGURA 2. Molécula correspondiente al algoritmo de evolución (14) para la construcción de la solución al tiempo t^{n+1} a partir de datos en los tiempos t^n y t^{n-1} según la Ec. (14).

Finalmente, la expresión (14) permite la evolución de valores de la función de onda en tiempos previos a t^{n+1} , y además indica que para iniciar la evolución es necesario conocer los valores de la función en dos tiempos previos. En este punto es cuando es necesario llenar el dominio espacial con datos iniciales en dos niveles de tiempo.

Estabilidad del algoritmo.

Existe una restricción del algoritmo descrito por (14) respecto a la estabilidad de la solución que se está calculando. De dicha expresión se ve que un valor del factor $\Delta t/\Delta x$ mayor que uno implicaría que la amplitud de la función de onda en cada punto crecería al tiempo t^{n+1} con respecto al valor que tenía en el tiempo t^n y sucesivamente para valores posteriores del tiempo. Al factor $\Delta t/\Delta x$ se le conoce como factor de Courant Friedrichs Lewy (CFL) debido a que dichos autores lo definieron para la ecuación de advección [5,6]. Para que la discretización (14) sea estable es necesario que el factor CFL cumpla $\Delta t/\Delta x < 1$ (ver el apéndice para una demostración de dicha condición).

Datos iniciales.

Cuando se trata de resolver una ecuación de segundo orden es necesario proveer el valor inicial de la función y de su primera derivada temporal. En el caso de la ecuación de onda, cuya evolución está dictada por (14), dado que son necesarios los datos en dos niveles de tiempo, antes de proceder a llevar a cabo la evolución, diremos que los datos iniciales estarán definidos en los niveles de tiempo t^0 y un hipotético t^{-1} , con el fin de calcular en el primer paso de tiempo el valor de ϕ en el tiempo t^1 .

Específicamente se elige aquí un perfil gaussiano en $t = t^0 = 0$, y dado que se conoce la solución exacta de la

ecuación de onda, se tiene que

$$\begin{aligned} \phi_j^0 &= Ae^{-x_j^2/\sigma^2}, \\ \phi_j^{-1} &= \frac{1}{2}Ae^{-(x_j+\Delta t)^2/\sigma^2} + \frac{1}{2}Ae^{-(x_j-\Delta t)^2/\sigma^2}, \end{aligned} \quad (15)$$

para $j = 0, 1, \dots, N$. En el caso general, en el que se conoce ϕ_j^0 y su derivada temporal $\partial_t(\phi_j^0)$ [ver (13)] no es directo conocer el valor de la función de onda ϕ_j^{-1} . Una posibilidad es definir los datos iniciales para la función de onda como ϕ_j^0 para cada j , y para calcular ϕ_j^{-1} se hace una expansión en serie de Taylor hacia atrás en el tiempo de modo que

$$\phi_j^{-1} = \phi_j^0 - \Delta t \partial_t(\phi_j^0) + \frac{\Delta t^2}{2} \partial_{tt}(\phi_j^0) + O(\Delta t^3),$$

donde si se desea obtener precisión de tercer orden se puede reemplazar $\partial_{tt}(\phi_j^0)$ por $\partial_{xx}(\phi_j^0)$ gracias a la ecuación de onda. Para el caso de datos iniciales simétricos en el tiempo ($\partial_t \phi_j^0 = 0$) se tiene $\phi_j^{-1} = \phi_j^0 + O(\Delta t^2)$, con error de segundo orden. Estos experimentos se pueden llevar a cabo modificando el programa [1].

Técnicamente, en el programa lo que se hace en lugar de asignar memoria a las variables en todos los niveles de tiempo $n = 1 \dots N_t$ es lo siguiente: se recicla la memoria, es decir, una vez que se ha construido la solución ϕ_j^{n+1} para todo j , se reasignan los valores

$$\begin{aligned} \phi_j^n &\rightarrow \phi_j^{n-1} \\ \phi_j^{n+1} &\rightarrow \phi_j^n, \end{aligned}$$

lo que permite aplicar el algoritmo de evolución al siguiente valor de $n + 1$. Entonces solamente se asigna memoria a las variables durante tres niveles de tiempo simultáneamente.

2.2. Condiciones de frontera y solución numérica

Condiciones de frontera periódicas.

En este caso, el dominio es tal que el extremo derecho del dominio espacial se identifica con el extremo izquierdo, o bien, que la topología del dominio espacial pasa de ser la recta real a ser la circunferencia S^1 , de perímetro $x_N - x_0$. En la Fig. 3 se ilustra ésta transformación cuando se ha considerado el dominio discreto. Este tipo de condiciones de frontera son útiles para hacer pruebas de códigos numéricos, usualmente con la finalidad de estudiar la estabilidad de los algoritmos y la propagación de los errores durante la evolución, esto es, alguna señal -puede ser una función de onda como la estudiada aquí- se pone a dar vueltas en un dominio cíclico, y se estudia si el algoritmo de evolución es inestable (que la señal explote), si es disipativo (que la amplitud decrezca), que sufra dispersión, etc, todos estos posibles efectos sin involucrar los efectos de fronteras en el dominio numérico que puedan entorpecer la interpretación del análisis que se hace de un algoritmo. También existen casos en los que las condiciones periódicas son apropiadas para aproximar una situación física, como el caso

en que se estudia la propagación de una onda en un medio periódico como la propagación de fonones en una red cristalina, o la propagación de una onda en un dominio periódico, por ejemplo sobre una superficie esférica.

Condición de onda saliente.

Se trata de una condición que se impone para modelar fronteras abiertas que permiten el flujo hacia fuera de las soluciones pero no hacia dentro. Son condiciones de frontera apropiadas para representar los fenómenos radiativos de un sistema aislado, que genera señales que se propagan hacia fuera; cuando se trata de un problema de valores iniciales como los descritos aquí, en los que se imponen fronteras artificiales a una distancia finita de los datos iniciales o de las fuentes, y no se desea que la radiación regrese a la fuente o rebote hacia el dominio numérico, es muy importante imponer condiciones de frontera que impidan el rebote de señales hacia el dominio numérico, tales condiciones son las de onda saliente.

En el caso de la ecuación de onda, se sabe que la solución general es del tipo $\phi(t, x) = f(x + t) + g(x - t)$ para f y g arbitrarias, cada una de las cuales corresponde al desplazamiento hacia la izquierda y a la derecha respectivamente de los pulsos iniciales, a lo largo de las líneas características $t/x = ct$. Tales funciones f, g son soluciones de las ecuaciones

$$(\partial_t - \partial_x)f = 0, \tag{16}$$

$$(\partial_t + \partial_x)g = 0, \tag{17}$$

respectivamente. Así pues, la condición de onda saliente en el caso de la ecuación de onda se reduce a imponer en el borde izquierdo $x_0 = -1$ la condición (16) y en el borde derecho $x_N = +1$ la condición (17). Es importante mencionar que para imponer la condición del borde izquierdo solamente se cuenta con valores de la función de onda en puntos vecinos hacia la derecha, y de manera análoga, para imponer la condición en el borde derecho solo se cuenta con valores de la función de onda en puntos vecinos hacia la izquierda. Por tanto, esta condición de frontera ilustra la utilidad de las aproximaciones en diferencias finitas de derivadas que consideran puntos de un solo lado. Por ejemplo, la aproximación en DF de (16) y (17), utilizando las expresiones (4) y (5) se escriben respectivamente:

$$\frac{\phi_0^{n-1} - 4\phi_0^n + 3\phi_0^{n+1}}{2\Delta t} - \frac{-\phi_2^{n+1} + 4\phi_1^{n+1} - 3\phi_0^{n+1}}{2\Delta x} = O(\Delta x^2, \Delta t^2),$$

$$\frac{\phi_N^{n-1} - 4\phi_N^n + 3\phi_N^{n+1}}{2\Delta t} + \frac{\phi_{N-2}^{n+1} - 4\phi_{N-1}^{n+1} + 3\phi_N^{n+1}}{2\Delta x} = O(\Delta x^2, \Delta t^2),$$

donde además se ha usado la discretización para la derivada temporal en x_0 y x_N al tiempo t^{n+1} . Las moléculas involucradas en estas aproximaciones aparecen en la Fig. 4. La

información conocida en estas expresiones es la siguiente: todos los valores de ϕ en los tiempos t^n y t^{n-1} , y los puntos internos de la malla al tiempo t^{n+1} que ya han sido calculados a través de (14), de modo que los valores incógnita son aquellos valores de ϕ en los puntos negros en la Fig. 4. Afortunadamente es posible resolver las expresiones anteriores para estos valores desconocidos de la función de onda, lo que da como resultado:

$$\phi_0^{n+1} = \frac{\frac{\Delta t}{\Delta x}(-\phi_2^{n+1} + 4\phi_1^{n+1}) + 4\phi_0^n - \phi_0^{n-1}}{3(1 + \frac{\Delta t}{\Delta x})}, \tag{18}$$

$$\phi_N^{n+1} = \frac{\frac{\Delta t}{\Delta x}(-\phi_{N-2}^{n+1} + 4\phi_{N-1}^{n+1}) + 4\phi_N^n - \phi_N^{n-1}}{3(1 + \frac{\Delta t}{\Delta x})}. \tag{19}$$

El uso de estas expresiones para los valores del campo, en los bordes del dominio implica el resultado correcto. Finalmente, en la Fig. 5 se muestra la solución calculada con los dos tipos de condiciones de frontera discutidos aquí.

2.2.1. Error y convergencia.

El error de la solución numérica con respecto a la solución exacta queda definido en cada punto del dominio espacial como

$$e_j = \phi_j^{num} - \phi_j^{ex}. \tag{20}$$

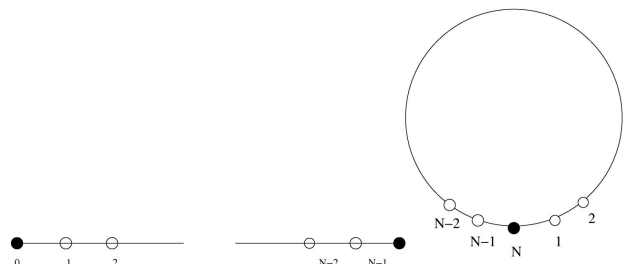


FIGURA 3. Identificación del dominio espacial que es parte de la recta real con una circunferencia. El mecanismo consiste en igualar los valores de ϕ en ambos extremos del dominio $\phi_0^n = \phi_N^n$ para todo n .

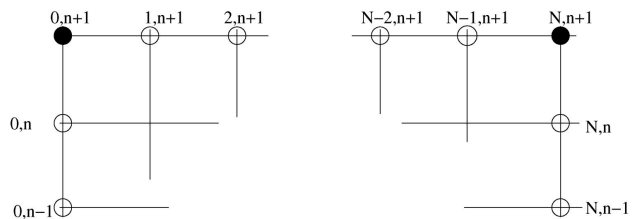


FIGURA 4. Moléculas de la versión discreta adecuadas para aplicar las condiciones de frontera de onda saliente (16,17). Al igual que antes, se desea calcular la función de onda en los puntos negros y los puntos blancos son las posiciones donde debe conocerse previamente los valores de ϕ .

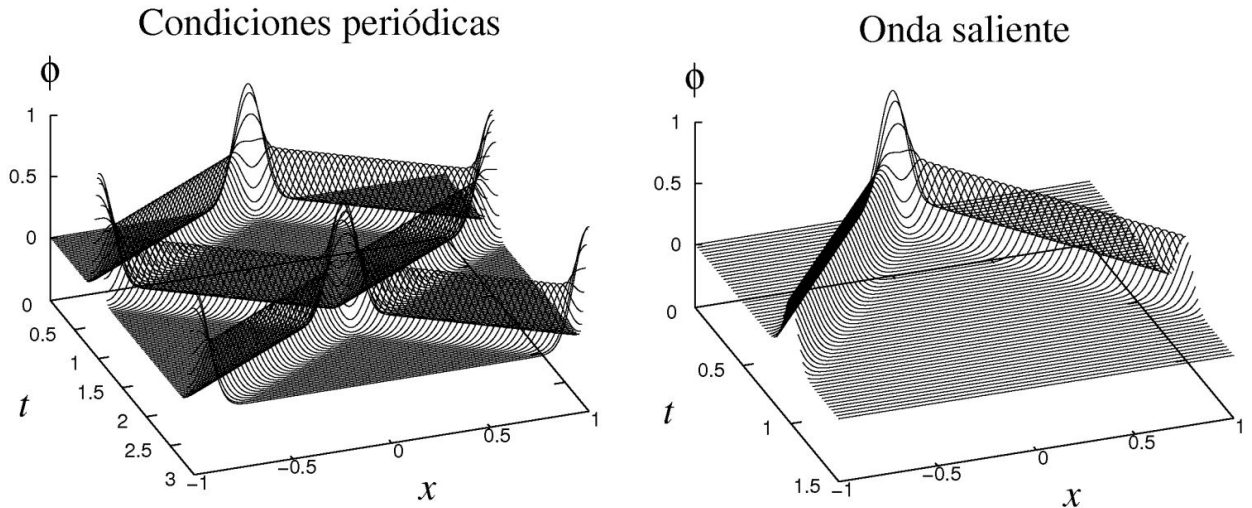


FIGURA 5. Solución numérica de la ecuación de onda que usa la discretización (14) con $A = 1$ y $\sigma = 0.1$. En el lado izquierdo se muestra el caso que usa condiciones periódicas, y la señal que sale de un lado del dominio entra por el otro. En el lado derecho se muestra el resultado de aplicar las condiciones de frontera de onda saliente (16,17), y la función de onda sale por las fronteras al tiempo $t \sim 1$, que corresponde al caso de que la onda viaja a velocidad 1.

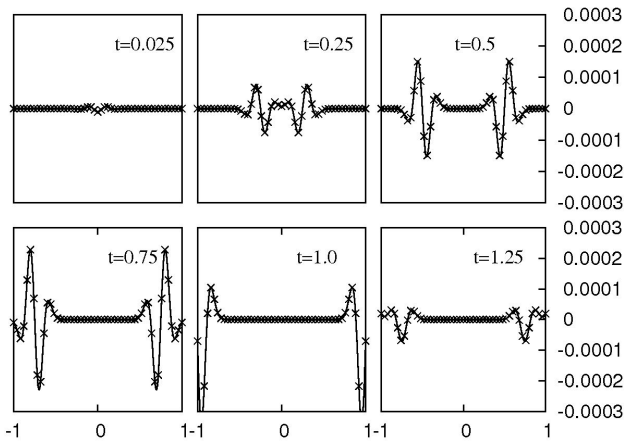


FIGURA 6. Error para distintos valores del tiempo de la solución recién mostrada para datos iniciales de perfil gaussiano simétricos en el tiempo con condiciones de frontera de onda saliente y dos resoluciones distintas $\Delta x = 0.002$ (líneas continuas) y $\Delta x = 0.001$ (cruces). Sabemos que en el límite continuo el error es cero, por lo que basta el uso de dos resoluciones para verificar la convergencia del error a cero. El error de la solución con mayor resolución ha sido multiplicado por 4 para verificar la expresión (11), o sea, que efectivamente el error es cuatro veces menor cuando se usa la resolución mayor que cuando se usa la resolución menor, por eso aparecen ambas curvas superpuestas. Esto muestra la convergencia del error a cero con segundo orden.

En la Fig. 6 se muestra el error de la solución numérica calculada usando (14,18,19) comparada con la solución exacta para el caso de condiciones de frontera de onda saliente. Para saber si los cálculos convergen a la solución exacta, basta con aprovechar el hecho de que en el límite continuo (o de resolución infinita) el error es cero. Entonces se usa la expresión (11) y se ejecuta el código con dos resoluciones distintas para mostrar la convergencia de segundo orden del error a cero.

El estudio del error de un cálculo numérico para todos los valores de tiempo es poco práctico. Es por tal razón que se estudia alguna función escalar del error y se estima la convergencia de ésta. Ejemplos de tales funciones son:

- i) El error en algún punto particular del dominio espacial, por ejemplo en las fronteras o en el máximo de la distribución gaussiana.
- ii) La norma $L_1(e)$ del error, definida como

$$L_1(e) = \sum_{j=0}^N |e_j|.$$

- iii) La norma $L_2(e)$ del error, definida como

$$L_2(e) = \sqrt{\sum_{j=0}^N e_j^2}.$$

Entonces, dichas funciones escalares del error se pueden monitorear como funciones del tiempo. Como ejemplo, en la Fig. 7 se muestran las normas $L_1(e)$ y $L_2(e)$ para las simulaciones mencionadas. Queda manifiesto que esta función escalar tiene convergencia de segundo orden a cero.

El monitoreo de la precisión a través de funciones escalares del error es un procedimiento muy eficiente, pues permite determinar si las ejecuciones y la implementación de los programas es correcta, sin necesidad de revisar directamente las soluciones numéricas en cada paso de tiempo. Desafortunadamente no siempre que se presentan simulaciones numéricas se presentan resultados de convergencia de las simulaciones. Existen teoremas que establecen condiciones que garantizan que dada la convergencia en una norma se tienen la convergencia de la solución numérica [6].

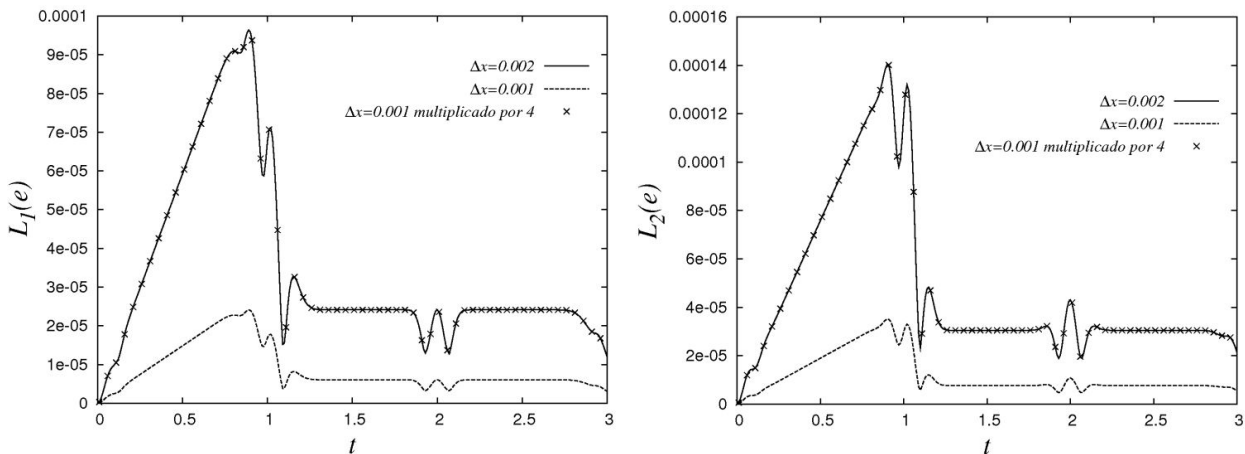


FIGURA 7. Normas L_1 y L_2 del error como función del tiempo para las simulaciones mostradas en la figura anterior. La línea continua indica la norma del error cuando se usa resolución $\Delta x = 0.002$, la discontinua corresponde a la resolución $\Delta x = 0.001$ y las cruces son una muestra de la norma calculados con resolución $\Delta x = 0.001$ multiplicada por 4. Una vez más, se verifica que las curvas continua y con cruces se superponen. De ahí que la convergencia de estas normas del error a cero es de segundo orden.

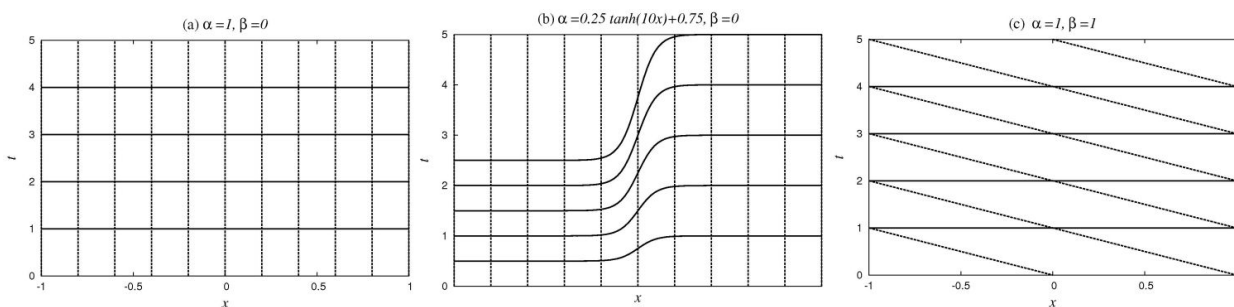


FIGURA 8. Se muestran los efectos de la elección de coordenadas en la descripción del espacio-tiempo. La líneas continuas describen curvas de coordenada temporal constante y las líneas intermitentes indican curvas de coordenada espacial constante. En el primer panel se indica el caso de coordenadas usuales, según las cuales los elementos del espacio-tiempo se etiquetan con curvas paralelas a los ejes coordenados. En el segundo panel se muestra una descripción del mismo espacio-tiempo pero con las curvas de tiempo constante separadas de una manera no uniforme que se usa como ejemplo posteriormente. Finalmente se muestra el caso en que se describe el espacio-tiempo con coordenadas tales que las curvas de posición constante no son verticales, sino que se desplazan con respecto a las coordenadas de un observador normal que viaja a lo largo de líneas verticales.

3. La ecuación de onda como sistema de primer orden

Hasta ahora, se ha mencionado que el dominio de la ecuación de onda es el espacio-tiempo, y en adelante se usará el formalismo de la física moderna que identifica a dicho dominio con el llamado espacio-tiempo de Minkowski. En esta sección se resuelve la ecuación de onda para ilustrar la solución de problemas en el espacio-tiempo plano en coordenadas que presentan distintos efectos sobre la función solución.

Una complicación adicional en esta sección consiste en que la ecuación de onda se escribirá no como una ecuación con derivadas parciales de segundo orden, sino que se transformará en un sistema de ecuaciones con derivadas de primer orden en el espacio y en el tiempo, porque se trata de un esquema que se aplica comunmente en una gran variedad de problemas de la física, incluidas las ecuaciones de Euler para la evolución de fluidos, las ecuaciones de Einstein para la evolución de la geometría del espacio-tiempo, las ecuaciones de Maxwell, etcétera.

Adicionalmente, la solución de la ecuación de onda muestra el proceso para construir la solución de una EDP hiperbólica de manera completa y la serie de ingredientes necesarios para la solución de un problema de valores iniciales con condiciones de frontera.

Así pues, se parte de que la ecuación de onda está definida en un subconjunto del espacio de Minkowski en 1+1 dimensiones, etiquetado por las coordenadas del espacio-tiempo (\tilde{x}, \tilde{t}) . El elemento de línea de dicho espacio-tiempo está dado por $ds^2 = -d\tilde{t}^2 + d\tilde{x}^2$ en tales coordenadas, o bien $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ donde $g_{\mu\nu}$ son las componentes del tensor métrico [7]. Dicho elemento de línea indica que el espacio-tiempo se etiqueta con rectas de \tilde{t} constante y \tilde{x} constante como se muestra en el primer panel de la Fig. 8. Una manera más general de etiquetar el espacio-tiempo (foliar el espacio-tiempo) implica la necesidad de etiquetar con curvas de coordenada temporal y espacial constantes mucho más generales que rectas. Esto se consigue mediante una transformación de coordenadas, por ejemplo, si se desea que en lugar de etiquetar el espacio-tiempo con rectas de tiempo

constante, se puede elegir una coordenada temporal tal que $dt = \alpha(x)d\tilde{t}$. Como ejemplo se muestra en la Fig. 8b, el caso en que el espacio-tiempo se etiqueta con curvas de tiempo constante que no son rectas horizontales en un diagrama de espacio-tiempo. De manera más general, se puede definir una transformación de coordenadas del tipo $dt = \alpha d\tilde{t}$ y $dx = d\tilde{x} - \beta d\tilde{t}$, donde α juega el papel de etiqueta entre los valores del tiempo considerados y β es una velocidad a la que se desplaza la coordenada x , es decir, ahora las curvas de coordenada espacial constante ya no son necesariamente líneas rectas verticales como en el caso de la Fig. 8a, y se ilustra el caso en que $\alpha = 1$ y $\beta = 1$ en el tercer panel. Al usar dichas coordenadas generales es posible por ejemplo, controlar el ritmo de la evolución de la función de onda en una parte determinada del dominio. De la Fig. 8b por ejemplo, se infiere que la evolución de cualquier fenómeno en la parte izquierda del dominio será más lenta en ese sistema coordinado que en la parte derecha, y en la Fig. 8c se muestra un caso en que las coordenadas espaciales se desplazan a lo largo del dominio. Entonces, la elección de las funciones α y β está en función del problema que se desea resolver.

En dichas coordenada generales, el tensor métrico resultante y su inversa en forma matricial son los siguientes:

$$g_{\mu\nu} = \begin{pmatrix} (-\alpha^2 + \beta^2) & \beta \\ \beta & 1 \end{pmatrix}, \tag{21}$$

$$g^{\mu\nu} = \begin{pmatrix} -1/\alpha^2 & \beta/\alpha^2 \\ \beta/\alpha^2 & (1 - \beta^2/\alpha^2) \end{pmatrix}, \tag{22}$$

donde $\alpha > 0$ es la función lapso y β es la única componente del llamado vector de corrimiento o *shift*, además $\mu, \nu = t, x$. En general el operador de onda o bien el operador d'Alambertiano para un espacio-tiempo dado está definido por

$$\square\phi = \frac{1}{\sqrt{-g}}\partial_\mu[\sqrt{-g}g^{\mu\nu}\partial_\nu\phi]$$

donde $g = \det(g_{\mu\nu})$. Para el caso general (21) se tiene que $\sqrt{-g} = \alpha$ y la ecuación de onda en su forma general se escribe

$$\begin{aligned} 0 &= \square\phi \\ &= \frac{1}{\sqrt{-g}}\partial_\mu[\sqrt{-g}g^{\mu\nu}\partial_\nu\phi] \\ &= \frac{1}{\alpha}\partial_t[\alpha g^{t\nu}\partial_\nu\phi] + \frac{1}{\alpha}\partial_x[\alpha g^{x\nu}\partial_\nu\phi] \\ &= \frac{1}{\alpha}\partial_t[\alpha g^{tt}\partial_t\phi + \alpha g^{tx}\partial_x\phi] + \frac{1}{\alpha}\partial_x[\alpha g^{xt}\partial_t\phi + \alpha g^{xx}\partial_x\phi] \\ &= \frac{1}{\alpha}\partial_t\left[-\frac{1}{\alpha}\partial_t\phi + \frac{\beta}{\alpha}\partial_x\phi\right] \\ &+ \frac{1}{\alpha}\partial_x\left[\frac{\beta}{\alpha}\partial_t\phi + \alpha\left(1 - \frac{\beta^2}{\alpha^2}\right)\partial_x\phi\right]. \end{aligned} \tag{23}$$

Obsérvese que cuando $\alpha = 1$ y $\beta = 0$ se recupera el operador de onda sencillo (8). Es deseable escribir un sistema de

EDPs que contenga derivadas solamente de primer orden. La razón es que se simplifica el estudio de las propiedades del sistema de ecuaciones con la idea de construir una solución global a partir de la evolución de datos iniciales. La expresión anterior (23) sugiere la definición de dos nuevas variables $\psi := \partial_x\phi$ y $\pi := (\partial_t\phi - \beta\partial_x\phi)/\alpha$. Obsérvese que π es el argumento de la derivada temporal de primer orden en (23).

Con la definición de estas dos nuevas variables es posible escribir la ecuación de onda como un sistema de dos ecuaciones de primer orden para estas dos nuevas variables. La primera de tales ecuaciones se obtiene directamente de (23):

$$\partial_t\pi = \partial_x(\alpha\psi + \beta\pi). \tag{24}$$

Si ahora se supone que ϕ es al menos de clase C^2 , la ecuación para ψ es $\partial_t\psi = \partial_t(\partial_x\phi) = \partial_x(\partial_t\phi)$, lo que implica

$$\partial_t\psi = \partial_x(\alpha\pi + \beta\psi). \tag{25}$$

Las Ecs. (24-25) constituyen la versión de primer orden de la ecuación de onda con la constricción $\psi = \partial_x\phi$, la cual debe satisfacerse con el grado de precisión de la aproximación en diferencias finitas. Esto recuerda que la función incógnita original ϕ no aparece en el sistema de ecuaciones, pero puede reconstruirse a partir de la definición de π una vez calculados π y ψ , o sea $\partial_t\phi = \alpha\pi + \beta\psi$.

Finalmente, el problema de valores iniciales con condiciones de frontera por resolver es el siguiente:

$$\begin{aligned} \partial_t\pi &= \partial_x(\alpha\psi + \beta\pi), \quad \partial_t\psi = \partial_x(\alpha\pi + \beta\psi) \\ \pi(x, 0) &= \pi_0(x), \quad \psi(x, 0) = \psi_0(x), \\ \pi(-1, t) &= \pi_L(t), \quad \pi(1, t) = \pi_R(t), \\ \psi(-1, t) &= \psi_L(t), \quad \psi(1, t) = \psi_R(t). \\ x &\in [-1, 1], \quad t > 0, \end{aligned} \tag{26}$$

con la constricción $\partial_x\phi = \psi$, misma que debe satisfacerse en el límite continuo.

3.1. Análisis característico

Al análisis característico de un sistema de ecuaciones de primer orden es muy útil y en algunos casos indispensable. Entre las aplicaciones que tiene dicho análisis se encuentran las siguientes:

- En algunos casos (como el presente) permite escribir el sistema de ecuaciones como un sistema de ecuaciones de advección, es decir, del tipo $\partial_t u = \partial_x u$, y de hecho aquí se hará con el sistema de ecuaciones de evolución en (26).
- Las variables que obedecen dichas ecuaciones de advección son las llamadas variables características, que se construyen aquí para el caso de la ecuación de onda.

- Dichas variables características permiten imponer condiciones de frontera de manera clara. En el caso de la ecuación de onda, tales variables características corresponden a los modos de la ecuación de onda que se propagan “a la derecha” y “a la izquierda”, y en el presente trabajo se hace para coordenadas incluso con $\alpha = \alpha(x)$ y $\beta = \beta(x)$.
- En el caso de problemas de valores iniciales con ecuaciones de evolución menos conocidas, permite determinar si el problema se encuentra bien planteado, o bien formulado. Esto es, si dados unos datos iniciales para el problema se obtiene una solución a un tiempo dado t_1 , y se resuelve nuevamente la evolución de datos iniciales como los originales pero con alguna ligera perturbación, la solución obtenida al tiempo t_1 será una solución como la obtenida originalmente pero con una ligera diferencia. Que un problema sea bien planteado estará relacionado con las propiedades de hiperbolicidad del sistema de ecuaciones [8], pero no discutiremos aquí su relación con la estabilidad de las soluciones.

La importancia de este tipo de análisis aplicado a la solución numérica de problemas de valores iniciales puede quedar ilustrada con el histórico caso de la solución del problema del choque de agujeros negros, el cual se había formulado como un problema de valores iniciales en una descomposición 3+1 del espacio-tiempo que obedecía un sistema de ecuaciones que no es fuertemente hiperbólico (ver abajo) y por tanto no permitía un problema de valores iniciales bien planteado; a la postre, después de varios años, mediante la manipulación de las ecuaciones y la introducción de nuevas variables de evolución se construyó otro sistema de ecuaciones de evolución fuertemente hiperbólico, lo que corresponde a un problema bien planteado y por tanto permitió la evolución estable de sistemas de agujeros negros en una computadora [2].

Para lograr el análisis del caso -que estamos tratando de las ecuaciones (26), se define el vector de estado $\mathbf{u} = (\pi, \psi)^T$, y es posible escribir las ecuaciones de evolución como

$$\partial_t \mathbf{u} + \mathbf{A} \partial_x \mathbf{u} = -\partial_x (\mathbf{A} \mathbf{u}), \quad (27)$$

donde

$$\mathbf{A} = - \begin{pmatrix} \beta & \alpha \\ \alpha & \beta \end{pmatrix}. \quad (28)$$

Las direcciones características, o sea, las direcciones locales de propagación de las señales en el plano xt se obtienen de los valores propios de la matriz \mathbf{A} , para lo cual se resuelve la ecuación $\det(\mathbf{A} - I_2 \lambda) = 0$ con $\lambda = dx/dt$, donde I_2 es la matriz identidad 2×2 . Los resultados son los siguientes

$$\lambda_{\pm} = -\beta \pm \alpha, \quad (29)$$

es decir, las velocidades de propagación dependen de las coordenadas que se han elegido para etiquetar el espacio-tiempo. Dado que los dos valores propios son distintos y reales, y como veremos enseguida los vectores propios forman un sistema completo de vectores, el sistema de ecuaciones es fuertemente hiperbólico [8]. Esta propiedad garantiza que se tiene un problema de valores iniciales bien planteado [8]. En el caso $\beta = 0$ y $\alpha = 1$ (la ecuación de onda usual) se obtienen las velocidades características $\lambda = \pm 1$, lo que implica que las trayectorias de los datos iniciales son rectas con pendientes ± 1 que definen el cono $x = x_0 \pm t$, cuyas rectas corresponden a las características de la solución, es decir, aquellas curvas (rectas en este caso) a lo largo de las cuales el valor de la solución es el mismo que el valor al tiempo inicial [9].

Los vectores propios correspondientes a λ_{\pm} son $u_1 = (1, -1)^T$ y $u_2 = (1, 1)^T$. La matriz que diagonaliza \mathbf{A} es por tanto

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \mathbf{P}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \quad (30)$$

De aquí, \mathbf{A} puede escribirse como $\mathbf{A} = \mathbf{P} \mathbf{\Lambda} \mathbf{P}^{-1}$ con $\mathbf{\Lambda} = \text{diag}(\lambda_+, \lambda_-)$. Multiplicando la Ec. (27) por \mathbf{P}^{-1} se tiene

$$\begin{aligned} \mathbf{P}^{-1} \partial_t \mathbf{u} + \mathbf{P}^{-1} \mathbf{A} \partial_x \mathbf{u} &= -\partial_x (\mathbf{P}^{-1} \mathbf{A} \mathbf{u}) \\ \partial_t \mathbf{w} + \mathbf{\Lambda} \partial_x \mathbf{w} &= -\partial_x (\mathbf{\Lambda} \mathbf{w}), \end{aligned} \quad (31)$$

donde

$$\mathbf{w} = \mathbf{P}^{-1} \mathbf{u} = \frac{1}{2} (\pi - \psi, \pi + \psi)^T = (R, L)^T \quad (32)$$

es el vector cuyas componentes son las *variables características*. De este modo, las dos Ecs. (24) y (25) se desacoplan, y la dinámica del campo escalar ha quedado descompuesta en un modo que se mueve a la derecha ($R = 1/2(\pi - \psi)$) y otro que se mueve a la izquierda ($L = 1/2(\pi + \psi)$).

3.2. Datos iniciales

En la versión de primer orden de la ecuación de onda es necesario conocer al tiempo inicial los valores de $\pi(0, x)$ y $\psi(0, x)$. Esto es equivalente a demandar valores para ϕ y su derivada temporal al tiempo inicial nuevamente. Con el fin de enfocarnos en la evolución de datos iniciales es posible elegir simplemente datos simétricos en el tiempo para un perfil gaussiano:

$$\begin{aligned} \phi(0, x) &= A e^{-(x-x_0)^2/\sigma^2}, \\ \psi(0, x) &= -2 \frac{(x-x_0)}{\sigma^2} \phi(0, x), \\ \pi(0, x) &= 0. \end{aligned} \quad (33)$$

Se sabe que la solución de la ecuación de onda es la superposición de un pulso que se mueve a la derecha y otro que se mueve a la izquierda ($\phi(t, x) = f(x+t) + g(x-t)$). Entonces la evolución de los datos iniciales (33) debe mostrar

la descomposición de los datos iniciales en dos gaussianas. De hecho para el caso usual ($\alpha = 1, \beta = 0$) y los datos iniciales (33) se tendría

$$\begin{aligned} \phi(0, x) &= \frac{1}{2} A e^{-([x+t]-x_0)^2/\sigma^2} \Big|_{t=0} \\ &+ \frac{1}{2} A e^{-([x-t]-x_0)^2/\sigma^2} \Big|_{t=0}, \end{aligned} \quad (34)$$

que involucra dos pulsos gaussianos superpuestos que al tiempo $t = 0$ aparecen como un solo pulso.

3.3. Evolucionando los datos iniciales

La evolución de datos consiste en el cálculo de una función f_j^{n+1} a partir de datos en los niveles de tiempo previos como se describió en la sección anterior. Para ilustrarlo considérese la Ec. (24) escrita en la forma siguiente para $j = 1, \dots, N - 1$:

$$(\partial_t \pi)_j = \frac{(\alpha \psi + \beta \pi)_{j+1} - (\alpha \psi + \beta \pi)_{j-1}}{2\Delta x}. \quad (35)$$

Esta versión de la ecuación se conoce como versión semidiscreta, es decir, que solamente la parte derecha aparece en forma discreta, denotada usualmente por $rhs(\pi)$. Esto es, (35) es una ecuación ordinaria en t de π para cada valor de j , o sea se tienen $N - 1$ ecuaciones diferenciales ordinarias en t de π . Entonces basta con un integrador de ecuaciones diferenciales ordinarias para evolucionar datos de un tiempo al siguiente en cada valor de j . Este método de evolución es conocido como “método de líneas” (MoL) y se trata de una herramienta muy utilizada para la evolución en problemas de valores iniciales. La molécula usada para construir la solución al tiempo $n + 1$ se muestra en la Fig. 9.

El integrador de ecuaciones ordinarias se elige en términos de la precisión deseada, la disipación que introduce en los cálculos y sus propiedades de estabilidad, lo que implica restricciones en el valor del factor CFL $\Delta t/\Delta x$ (para aprender más acerca de las propiedades y estabilidad de los algoritmos de evolución conviene revisar las Refs. 6 y 8). Para obtener los resultados de este trabajo se ha usado un algoritmo Runge-Kutta de tercer orden, según el cual, si la función f satisface una ecuación del tipo $\partial_t f = rhs(f) = S$, donde S es la *parte derecha* de la ecuación de evolución para f y que puede incluir funciones y derivadas de funciones conocidas en su versión discreta, el algoritmo para calcular f^{n+1} en términos de valores de funciones en el tiempo previo se resume así:

$$\begin{aligned} f^* &= f^n + \Delta t S^n, \\ f^{**} &= \frac{3}{4} f^n + \frac{1}{4} f^* + \frac{\Delta t}{4} S^*, \\ f^{n+1} &= \frac{1}{3} f^n + \frac{2}{3} f^{**} + \frac{2}{3} \Delta t S^{**}. \end{aligned}$$

Tal algoritmo es usado comunmente porque requiere de tres iteraciones solamente, es preciso y estable para valores suficientemente pequeños del factor CFL.

3.4. Condiciones de frontera

Como se puede ver de la Ec. (35) y de la Fig. 9, el valor de la variable que se desea actualizar puede calcularse solamente en los puntos interiores del dominio, y no en aquellos localizados en las fronteras x_0, x_N ; la razón es que en la *parte derecha* de π aparece una derivada espacial. Este no es un obstáculo sino nuevamente una oportunidad para imponer condiciones de frontera que involucran derivadas espaciales. Como ejemplos de condiciones de frontera para la ecuación de onda se considera solamente las de onda saliente. En el caso de las condiciones periódicas se procede de manera análoga a la implementada en la sección anterior, es decir se cambia la topología del dominio finito a la de S^1 . Sin embargo, en el caso de las condiciones de onda saliente, el operador de onda general usado aquí no se factoriza tan fácilmente como se hizo en (16,17), pues ahora se trata de coordenadas generales. Sin embargo, la descomposición en modos derecho e izquierdo de la solución permiten implementar de manera elegante las condiciones de frontera.

La condición de onda saliente en la frontera izquierda consiste en la eliminación del modo que viaja hacia la derecha ($R = 0$), lo que significa que no se permite la reflexión hacia dentro del dominio, y la condición en la frontera derecha consiste en eliminar el modo que viaja a la izquierda ($L = 0$). Explícitamente, en la frontera izquierda x_0 se exige

$$\begin{aligned} \frac{1}{2}(\pi_0^n + \psi_0^n) &= L_0, \\ \frac{1}{2}(\pi_0^n - \psi_0^n) &= R_0 = 0, \end{aligned} \quad (36)$$

cuya solución es $\pi_0^n = \psi_0^n = L$. De modo equivalente, la condición en la frontera derecha es $\pi_N^n = R_N$ y $\psi_N^n = -R_N$. El problema se reduce a calcular L_0, R_0, L_N y R_N . Para conocer estos valores basta con hacer una extrapolación usando los puntos internos de la malla.

3.5. Resultados

Con todos estos conceptos en mente: datos iniciales, un algoritmo de evolución para las funciones de un tiempo al siguiente, condiciones de frontera y un código que contiene todos estos ingredientes, es posible construir soluciones de la ecuación de onda usando distintas coordenadas. Aquí, se presentan algunos ejemplos ilustrativos. Nuevamente, se elige el dominio $[-1, 1] \times [0, t)$ y datos iniciales con $A = 1$ y $\sigma = 0.1$.

En la Fig. 10 se presentan dos casos con $\alpha = 1$. En el primer caso $\beta = 0$, lo que corresponde al caso usual de la ecuación de onda. De hecho, puede verse que la gaussiana inicial se divide en dos gaussianas pequeñas que llegan a las fronteras al mismo tiempo (alrededor $t \sim 1$) al igual que en la Sec. 2. Sin embargo, en el segundo caso $\beta = 1$, esto es, las coordenadas se mueven hacia $x > 0$ a la velocidad de propagación de la onda; entonces el sistema de coordenadas sigue

a uno de los pulsos, lo que en la figura se traduce en un pulso centrado permanentemente en $x = 0$; el otro pulso alcanza la frontera en la mitad del tiempo que toma en el caso anterior.

Un ejemplo que ilustra el papel del lapso α es el mostrado en la Fig. 11. En tal caso el lapso tiene el perfil de una función escalón suavizada tal que el lapso va de 0.5 a 1.0. El efecto que produce esta elección es que -siendo α^2 el coeficiente de dt^2 - determina qué tan separados se encuentran los niveles de tiempo (ver la Fig. 8b). Por tanto, la evolución en la región donde $\alpha = 0.5$ ($x < 0$) es más lenta que en la región donde $\alpha = 1$ ($x > 0$). De hecho, en el primer caso al pulso le toma un tiempo ~ 2 para llegar a la frontera, mientras que en el segundo caso el pulso llega a la frontera en un tiempo ~ 1 . Este tipo de comportamiento es muy útil en escenarios de relatividad general. Por ejemplo, cuando se forma un hoyo negro debido al colapso de algún tipo de materia, lo que se encuentra es que los invariantes geométricos comienzan a diverger en la región de la singularidad, y por tanto una posibilidad para suavizar los efectos de gradientes grandes de funciones consiste en elegir unas coordenadas (en términos de α y β) que comprime la separación temporal entre un tiempo dado y el siguiente en la región cercana al origen demandando que $\alpha \rightarrow 0$ en dicha región, de modo que la evolución tiende a congelarse ahí.

Como ejemplo final, en la Fig. 12 se presenta la solución para el caso $\alpha = 1$ y $\beta = x$. En este caso las coordenadas viajan a la velocidad de la onda en las fronteras $x = \pm 1$, porque en esos dos puntos $\beta = \pm 1$. Esto implica que las señales nunca llegarán a las fronteras numéricas. Otro efecto es que los pulsos en estas coordenadas se comprimen conforme se acercan a las fronteras. Por una parte ésta es una ventaja, porque

no es necesario imponer condiciones de frontera (las señales nunca llegarán a las fronteras), y por otra parte, los pulsos están siendo resueltos con cada vez menos puntos del dominio espacial, lo que afecta la precisión de los cálculos. En cualquier caso, este ejemplo ilustra las posibilidades que se tienen cuando se puede elegir un lapso y un *shift* no triviales.

En la Fig. 13 se presenta la autoconvergencia de la solución para el último de los casos mostrado, y se indica que a partir de cierto momento, la solución deja de converger y no se debe confiar en el resultado numérico a partir de entonces. Si se desea que la solución converja durante más tiempo, basta -para este caso- con aumentar la resolución, ya que el hecho de que no converja la solución es debido a que los pulsos se comprimen en las coordenadas escogidas conforme se aproximan a las fronteras, y el número de puntos con los que se desea resolver cada pulso disminuye.

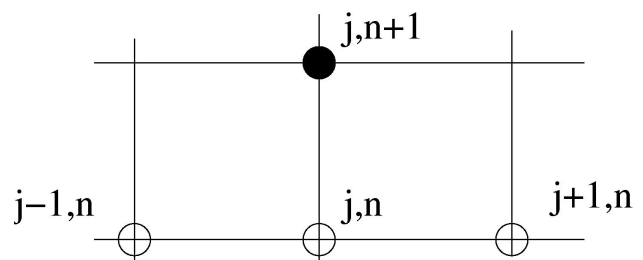


FIGURA 9. Ilustración de la molécula usada en la construcción de la solución al tiempo $n + 1$ con dos niveles de tiempo. El círculo negro indica el punto de la malla donde se desea calcular el valor de una función, mientras que los círculos vacíos indican la posición de los puntos de la malla donde es necesario conocer los valores de las funciones involucradas.

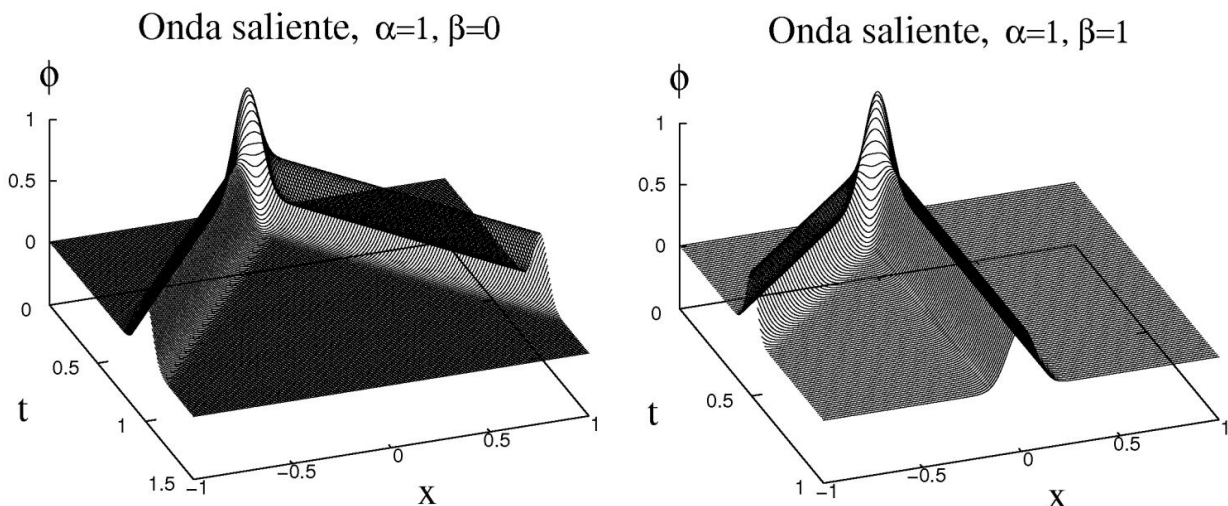


FIGURA 10. Solución de la ecuación de onda para dos sistemas de coordenadas. (Izquierda) la ecuación de onda usual $\alpha = 1, \beta = 0$; la gaussiana inicial se parte en dos gaussianas de amplitud igual a la mitad de la original y mismo ancho, cada una de las cuales viaja hacia las fronteras. (Derecha) la ecuación de onda con $\alpha = 1$ y $\beta = 1$, lo que implica que las coordenadas viajan a la velocidad de propagación de uno de los pulsos, por eso uno de los pulsos aparece siempre en la misma posición.

4. Solitones, un caso de ecuación de onda no homogénea

Ecuaciones de tipo onda no homogéneas son objeto de estudio en distintas ramas de la física, siendo la de Klein-Gordon (KG) una de las mejor estudiadas. La ecuación de KG se escribe en general como

$$\square\phi + \frac{dV(\phi)}{d\phi} = 0,$$

donde ϕ se llama campo escalar y $V(\phi)$ es el potencial al que se somete el campo escalar ϕ , y determina sus propiedades físicas. Las aplicaciones de dicha ecuación van desde la propagación de fonones en los materiales y ecuaciones de perturbación de sistemas astrofísicos (como los agujeros negros [12]), hasta la evolución del universo temprano durante el período llamado inflación y algunos modelos que explican la energía oscura en el universo [11].

Como ejemplo ilustrativo se considera un tipo de ecuación de KG que admite soluciones solitónicas, es decir, soluciones cuya densidad de energía -y en algunos casos el perfil del campo escalar- es independiente del tiempo (condición de onda solitaria) y que al interactuar con otras soluciones del mismo tipo en el mismo espacio-tiempo el perfil de densidad antes y después de la interacción se mantiene sin cambio; en la Ref. 13 se expone con detalle la definición matemática de los solitones. En particular aquí se muestra el funcionamiento

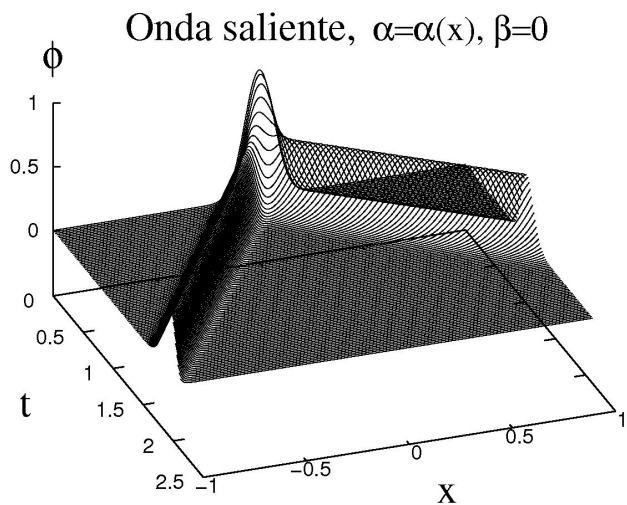


FIGURA 11. Solución de la ecuación de onda para el caso $\beta = 0$ y $\alpha = 0.25 \tanh(10x) + 0.75$, que es una versión suave de una función escalón que salta de 0.5 a 1.0. Puede observarse que en la región donde $\alpha = 0.5$ ($x < 0$) la onda se propaga con baja velocidad en estas coordenadas (el pulso tarda más en llegar que el pulso en la otra mitad del dominio). Esto es así porque en dicha región se ha usado unas coordenadas tales que el intervalo de tiempo es $0.5dt$, de modo que los intervalos de tiempo en esa región se encuentran más cercanos uno al otro, mientras que en la región con $x > 0$, para el cual el intervalo de tiempo es como en los casos anteriores $1.0dt$, ver la Fig. 8b.

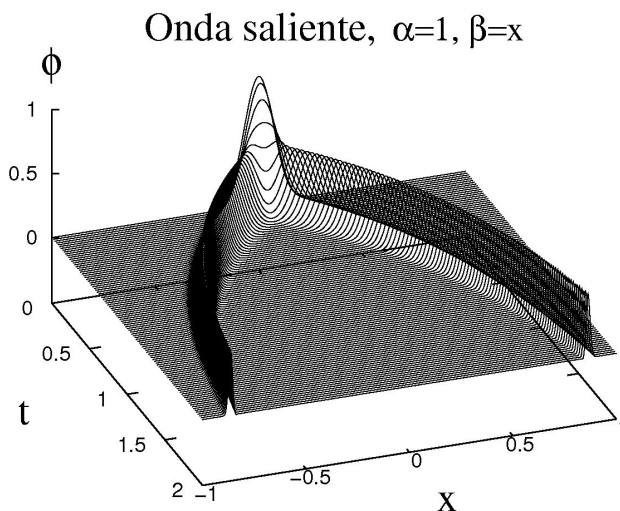


FIGURA 12. Solución para el caso $\beta = x$ y $\alpha = 1$. Debido a que el valor del *shift* iguala al valor de la velocidad de propagación de la onda en las fronteras, las señales nunca llegarán a las fronteras. Las gaussianas se comprimen conforme se aproximan a las fronteras.

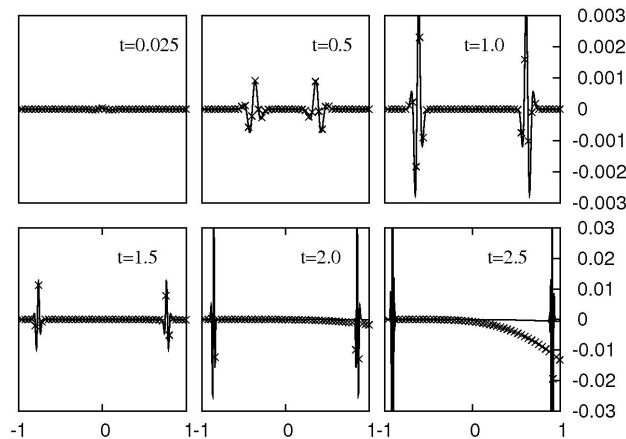


FIGURA 13. Se presenta la autoconvergencia de la solución para el caso en que $\beta = x$ y $\alpha = 1$. Para tal efecto se usa el resultado (12). Para construir esta gráfica se usan las resoluciones $\Delta x_1 = 0.002$, $\Delta x_2 = 0.001$ y $\Delta x_3 = 0.0005$. La línea continua corresponde a la resta de la solución calculada con Δx_1 y Δx_2 , mientras que la muestra de puntos \times corresponde a la resta de las soluciones calculadas con Δx_2 y Δx_3 multiplicada por 4. De los resultados se tiene que los resultados convergen hasta $t \sim 2$ pues se superponen ambas curvas, y a partir de entonces las dos curvas no se superponen, por lo que no se puede confiar en los resultados a partir de entonces pues no convergen con el orden esperado.

de los métodos numéricos descritos en el caso de la ecuación de Sine-Gordon, es decir, la ecuación de KG con potencial $V(\phi) = 1 - \cos \phi$, es decir, la ecuación $\square\phi + \sin \phi = 0$.

Se conocen soluciones exactas de dicha ecuación, mismas que usaremos como datos iniciales para resolver el problema de evolución. Para lograrlo se adopta la versión de variables de primer orden descrito en la sección anterior. Así pues, como datos iniciales se superponen dos soluciones solitónicas

en el mismo dominio al tiempo inicial:

$$\begin{aligned} \phi(x, 0) = & 4 \tan^{-1} \left[\exp \left(\frac{x - x_0 - vt}{\sqrt{1 - v^2}} \right) \right]_{t=0} \\ & - 4 \tan^{-1} \left[\exp \left(\frac{x + x_0 + vt}{\sqrt{1 - v^2}} \right) \right]_{t=0} \end{aligned} \quad (37)$$

donde v es la velocidad de propagación del campo. Estos datos iniciales corresponden a dos solitones centrados en $\pm x_0$ y con velocidad v tal que se mueven uno en dirección del otro. Dado que se desea ver el fenómeno desde un sistema de coordenadas fijo y centrado en el origen se eligen $\alpha = 1, \beta = 0$, y los valores iniciales para las dos variables de primer orden se reducen a

$$\begin{aligned} \psi = \partial_x \phi = & 2 \operatorname{sech} \left[\frac{-x + x_0 + vt}{\sqrt{1 - v^2}} \right] \\ & - 2 \operatorname{sech} \left[\frac{-x - x_0 - vt}{\sqrt{1 - v^2}} \right], \\ \pi = \partial_t \phi = & -2v \operatorname{sech} \left[\frac{-x + x_0 + vt}{\sqrt{1 - v^2}} \right] \\ & + 2v \operatorname{sech} \left[\frac{-x - x_0 - vt}{\sqrt{1 - v^2}} \right]. \end{aligned}$$

Las ecuaciones de evolución son el conjunto de Ecs. (26) que aquí se reescriben con la corrección en la ecuación para π que contiene la derivada del potencial del campo escalar:

$$\begin{aligned} \partial_t \psi &= \partial_x (\alpha \pi + \beta \psi), \\ \partial_t \pi &= \partial_x (\alpha \psi + \beta \pi) - \alpha \sin \phi. \end{aligned}$$

La versión discreta de $rhs(\pi)$ es

$$[(\alpha \psi + \beta \pi)_{j+1} - (\alpha \psi + \beta \pi)_{j-1}] / 2\Delta x - \alpha_j \sin \phi_j.$$

En la Fig. 14 se presenta la evolución de la función de onda calculada como un problema de valores iniciales usando $x_0 = 10$ y $v = 0.5$. También se presenta la densidad de energía asociada a la función de onda, escrita en general como [13]

$$\rho = \frac{1}{2} \left(\frac{\partial \phi}{\partial t} \right)^2 + \frac{1}{2} \left(\frac{\partial \phi}{\partial x} \right)^2 + V. \quad (38)$$

También se muestra la gráfica de convergencia del error de la solución numérica comparada con la solución exacta (37). La propiedad de que el perfil de densidad de energía

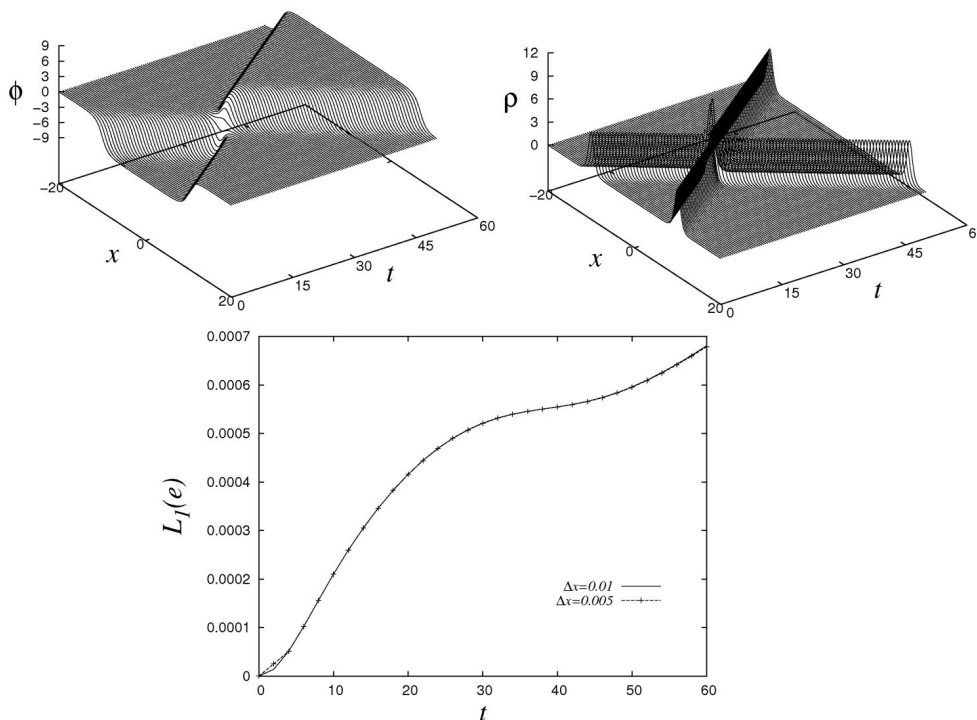


FIGURA 14. Se muestra la evolución de dos solitones superpuestos que obedecen la ecuación de Sine-Gordon. En el primer panel se muestra el perfil del campo escalar y cómo en el momento de la colisión cambia de signo aproximadamente a $t \sim 15$. Se muestra además el perfil de densidad del campo escalar según la Ec. (38), el cual se conserva para ambos solitones después de la colisión. Un trabajo detallado relacionado con el choque de solitones debería mostrar que los perfiles se preservan y aquí solamente se muestra en una sola gráfica el comportamiento de la densidad de energía. También se ha comparado la solución numérica con la exacta (37) y se muestra la norma L_1 del error usando distintas resoluciones, donde se ha multiplicado por cuatro el error calculado con la resolución alta para verificar la convergencia de segundo orden a cero del error. El dominio es $x \in [-20, 20], t \in [0, 60]$.

se preserva antes y después de la colisión es el signo más importante de que se trata de dos soluciones solitónicas [13].

De igual modo se pueden explorar las propiedades solitónicas de otros tipos de ecuación de onda cuya solución exacta sea desconocida [3]. Por ejemplo, en la Ref. 14 se descubrió que en modelos que consideran que la materia oscura es un condensado de Bose, las estructuras (como galaxias o cúmulos de galaxias) se comportan como solitones y se trasladan cuando son sometidos a condiciones de colisión.

Una propiedad importante de este tipo de sistemas con operadores de onda en ecuaciones no homogéneas es que la estructura característica es la misma que en el caso homogéneo, pues los términos del sistema de ecuaciones que se estudian en el análisis característico son los que corresponden a la parte principal del sistema de ecuaciones, lo que garantiza que las velocidades de propagación y los modos característicos de este sistema se definen del mismo modo que en el caso de la ecuación de onda homogénea.

5. Comentarios finales

En este trabajo se ha presentado tanto la manera más simple de resolver la ecuación de onda usando diferencias finitas, como la más completa que involucra el análisis característico.

Un punto importante relacionado con los algoritmos es que en los ejemplos desarrollados en este manuscrito, la resolución en todos los casos es uniforme. Esta condición es suficiente para resolver los problemas de valores iniciales con una dimensión espacial como los mostrados aquí, pero no necesariamente lo es en casos más generales. La razón es que cuando las ecuaciones involucran tres dimensiones espaciales la memoria de la computadora se convierte en una limitante. Es entonces necesario optimizarla y hacer uso de algoritmos de refinamiento de mallas que permiten asignar mayor resolución solamente en aquellas regiones donde los gradientes de las funciones involucradas en las EDPs son grandes, por ejemplo en nuestro caso puede ser en las regiones donde los pulsos de ϕ tienen amplitudes mayores. Un algoritmo comúnmente usado para este propósito es el conocido como el de "caja en caja" de Berger y Oliger [15], que consiste en que a partir de un dominio espacial de un problema de valores iniciales con resolución Δx , se eligen subdominios en los cuales la resolución es mayor, por ejemplo $\Delta x/2$, y dentro de estos subdominios se definen otros con resolución aún mayor, digamos $\Delta x/4$ y así sucesivamente. La particularidad de dicho algoritmo es que las fronteras de cada subdominio descansan sobre puntos de la malla de la siguiente resolución menor.

El elemento más importante en los cálculos numéricos basados en la aproximación en diferencias finitas es la convergencia de la solución numérica. No sobra insistir en que si se resuelve un problema con la aproximación de las diferencias finitas es necesario indicar ya sea la convergencia o la autoconvergencia de los resultados numéricos.

Respecto a la solución de un sistema de primer orden, es importante tener en cuenta la noción de velocidades características y modos característicos de propagación. Estos

permiten entender si un problema de valores iniciales es bien planteado o no, además de la utilidad que tienen dichos elementos para imponer condiciones de frontera correctamente.

Agradecimientos

El autor agradece a Alejandro Cruz Osorio, Fabio D. Lora Clavijo y a Jesús M Rueda Becerril por haber hecho observaciones importantes que mejoraron el texto. Este trabajo recibe apoyo parcial de los proyectos CIC-UMSNH-4.9 y CO-NACyT 106466.

Apéndice A

Aquí se presenta una prueba de estabilidad para el esquema de discretización (14) y las restricciones del factor CFL.

En esta prueba se estudia el crecimiento del error con el tiempo. Así a $t = 0$ en $x \in [0, N\Delta x]$ se considera que el error en cada punto del dominio espacial es $e_j = e(j\Delta x)$. Debido a que la aproximación en diferencias finitas (14) es lineal bastará con calcular la propagación del error para un solo j . Se propone la expresión para el error con propagación exponencial del error al tiempo inicial como

$$e_j^n = e^{i\mu j \Delta x} e^{\nu k n} = e^{i\mu j \Delta x} \xi^n \quad (\text{A.1})$$

donde para $t = 0$ se tiene $e_j = e^{i\mu j \Delta x}$. Se propone $\xi = e^{\nu k}$ y $\nu \in \mathbb{C}$, lo que permite que si ν es imaginario el error solamente oscilará en el tiempo, pero no crecerá, y por el contrario, si ν es real positivo el error crecerá exponencialmente. El criterio de estabilidad se reduce a pedir que $|\xi| \leq 1$ para un esquema estable. De (14) se tiene que el error satisface la misma ecuación que ϕ , por lo que es útil introducir la expresión (39) en (14):

$$\begin{aligned} e_j^{n+1} - 2e_j^n + e_j^{n-1} &= \rho^2 (e_{j+1}^n - 2e_j^n + e_{j-1}^n) \Rightarrow \\ e^{i\mu j \Delta x} \xi^n \left(\xi - 2 + \frac{1}{\xi} \right) &= \rho^2 (e^{i\mu \Delta x} - 2 + e^{-i\mu \Delta x}) \Rightarrow \\ \xi^2 - 2\xi A + 1 &= 0 \end{aligned} \quad (\text{A.2})$$

donde $\rho = \Delta t / \Delta x$ y $A = 1 - 2\rho^2 \sin^2(\nu \Delta x / 2)$. Las raíces de dicha ecuación son: $\xi_+ = A + \sqrt{A^2 - 1}$ y $\xi_- = A - \sqrt{A^2 - 1}$. Ahora, $A \leq 1$ porque para todo valor de ν se cumple $1 - 2\rho^2 \sin^2(\nu \Delta x / 2) \leq 1$. Se tienen pues dos casos i) $|A| \leq 1$, lo que implica que $|\xi_+| = |\xi_-| = 1$, y basta para indicar que el esquema es estable, y ii) $A < -1$, lo que implica que $|\xi_-| > 1$, lo cual basta para que el esquema sea inestable.

Entonces hay solamente una condición que se debe cumplir para que el esquema sea estable, $|A| \leq 1$, o sea $-1 \leq 1 - 2\rho^2 \sin^2(\nu \Delta x / 2) \leq 1$, siendo la primera desigualdad la única que implica una restricción sobre el valor en ρ , a saber, para que el esquema sea estable es necesario que $\rho^2 \leq 1$.

Finalmente, retomando la notación de la discretización se tiene que el factor CFL está sometido a la restricción $\Delta t/\Delta x \leq 1$. El lector puede usar el código para resolver

la ecuación de onda con distintos valores del factor CFL y verificar en la práctica ésta condición de estabilidad.

-
1. <http://www.ifm.umich.mx/guzman/Grupo/grupo.html>
 2. M. Alcubierre, *Introduction to 3+1 Numerical Relativity*, (Oxford University Press, 2008).
 3. R. Becerril, F. S. Guzmán, A. Rendón-Romero, and S. Valdez-Alvarado, *Rev. Mex. Fis.* **E 54** (2008) 120.
 4. E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics*, (Springer 1999).
 5. R. Courant, K. Friedrichs and H. Lewy, “On the partial difference equations of mathematical physics”, (IBM Journal, March 1967). pp. 215.
 6. J.W. Thomas, *Numerical Partial Differential Equations*, (Texts in Applied Mathematics. Springer, 1995).
 7. G.B. Arfken and H.J. Weber, *Mathematical Methods for Physicists*. (Academic Press, 2005).
 8. B. Gustafsson, H-O. Kreiss, and J. Olinger, *Time Dependent Problems and Difference Methods*. (Wiley-Interscience, 1996).
 9. R.J. LeVeque, in *Numerical methods for conservation laws*. (Birkhauser, Basel 1992).
 10. W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in Fortran*. (Cambridge University Press, 1992).
 11. A.R. Liddle and D.H. Lyth *Cosmological Inflation and Large-Scale Structure*. (Cambridge University Press, 2000).
 12. S. Chandrasekhar, *The Mathematical Theory of Black Holes*, (Oxford University Press, 1998).
 13. R. Rajaraman, *Solitons and Instantons*, (North-Holland Personal Library, Elsevier 1984).
 14. A. Bernal and F.S. Guzmán, *Phys. Rev.* **D 74** (2006) 103002.
 15. M.J. Berger and J. Olinger, *J.Comp.Phys.* **53** (1984) 484.