

Efectos negativos de la aproximación y representación falsa de ondas sobre una malla

Yuri N. Skiba y David Parra-Guevara

Centro de Ciencias de la Atmósfera

Universidad Nacional Autónoma de México

Circuito Exterior, Ciudad Universitaria

México, D.F., 04510

MÉXICO

skiba@servidor.unam.mx

Resumen

Se considera el problema de la aproximación y representación falsa de ondas sobre una malla. Se muestran varios efectos negativos de este fenómeno en el problema de extrapolación (exactitud de un pronóstico), en el análisis de datos (precisión de la estimación del período de un campo usando mediciones discretas), y en la modelación numérica (dispersión de ondas e inestabilidad no lineal causada por la acumulación falsa de la energía en escalas pequeñas). También se discute la dispersión numérica y la conexión lógica entre la aproximación, la viscosidad numérica y la estabilidad.

1. Representación falsa de ondas sobre una malla

1.1. Un caso particular

Un dominio D , por ejemplo, un intervalo espacial o temporal $[a, b]$ consiste de un número infinito e innumerable de puntos x , además, el rango de funciones suaves definidas sobre tal dominio también está formado por un número infinito e innumerable de valores. Por otra parte,

cualquier computadora tiene memoria acotada y, por lo tanto puede operar sólo con elementos de un espacio de dimensión finita como los vectores, las matrices, etc. Por lo tanto, para resolver numéricamente (con una computadora) un problema continuo (diferencial) cuya solución es una función, es indispensable aproximar dicho problema por un problema discreto cuya solución ya es un vector de dimensión finita. En el proceso de aproximación (discretización) del problema continuo, se introduce en $[a, b]$ una malla (regular o no regular) con nodos x_i ($i = 1, 2, \dots, N$). La distancia entre dos nodos vecinos $x_i - x_{i-1}$ se llama el tamaño de malla. A continuación, denotaremos por h y τ los tamaños de mallas regulares en espacio y tiempo. Como resultado, en lugar de los valores de una función $f(x)$ en todos los puntos $a \leq x \leq b$ se consideran sólo sus valores $f(x_i)$ en N nodos x_i de la malla ($i = 1, 2, \dots, N$). Ya que el número N de nodos es acotado, la función continua $f(x)$ se cambia por una función de malla (un vector) $\{f(x_1), f(x_2), \dots, f(x_N)\}$. Si la discretización se realiza por el método de diferencias finitas, entonces todas las funciones y sus derivadas que aparecen en el modelo continuo se aproximan por las funciones de malla correspondientes. En este trabajo consideraremos varios efectos (a menudo, negativos) relacionados con el uso de las mallas.

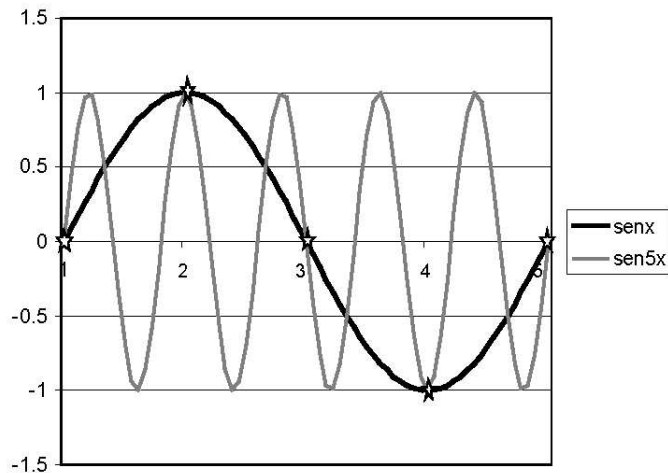


Figura 1: Los gráficos de las funciones $\text{sen } x$ y $\text{sen } 5x$ (en una malla regular).

Primero, consideramos fenómenos periódicos en espacio y tiempo. Cuando los valores $f(x_i)$ de una función $f(x)$ se conocen sólo en los nodos x_i de una malla espacial o temporal ($i = 1, 2, \dots, N$), es muy fácil

obtener una conclusión errónea sobre el comportamiento de la función. En efecto, en el plano xy hay un número infinito de diferentes curvas continuas que pasan a través de los puntos $(x_i, f(x_i))$, $i = 1, 2, \dots, N$. En particular, si la función $f(x)$ es periódica, es difícil determinar su período. Por ejemplo, en la Figura 1 se presentan las dos senoides $\text{sen}x$ y $\text{sen}5x$, con los periodos 2π y $2\pi/5$, respectivamente. Sin embargo, ambas funciones tienen los mismos valores en los nodos x_i ($i = 1, 2, \dots, 5$) de una malla regular y, por lo tanto, sus gráficos pasan a través de los mismos puntos (marcados con estrellas en Fig.1). Evidentemente, tendremos el mismo problema en el caso cuando $f(x)$ es descrita por sus valores en los nodos de una malla no regular (Fig.2).

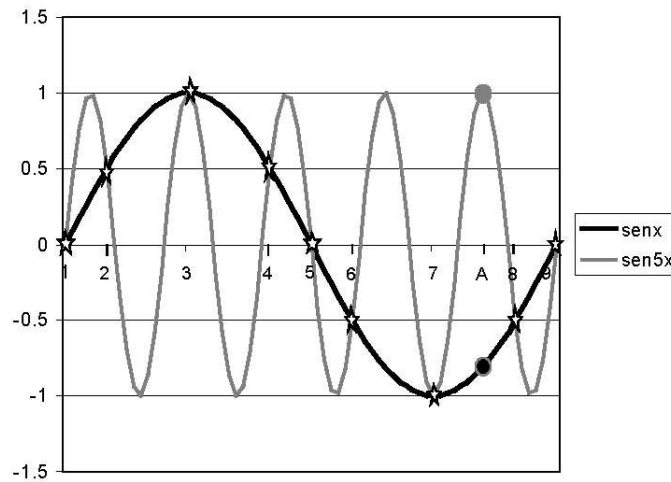


Figura 2: Los gráficos de las funciones $\text{sen}x$ y $\text{sen}5x$ (en una malla irregular).

Sea ahora $x \equiv t$ la variable temporal, y sean $f(t_i)$ los valores de un proceso no estacionario periódico $f(t)$ dado (por ejemplo, observado) en los nodos t_i ($i = 1, \dots, 7$) de una malla. Supongamos que estamos en el momento marcado por el punto t_7 y queremos pronosticar el valor $f(A)$ del proceso en un momento futuro A usando el método de extrapolación a través de una función cuyos valores coinciden con los datos observados en los puntos pasados t_1, t_2, \dots, t_7 . Ya que los valores $\text{sen}t_i$ y $\text{sen}5t_i$ de ambas senoides coinciden en los puntos mencionados, entonces podemos aproximar el proceso $f(t)$ por cualquiera de las dos senoides. Podemos, por ejemplo, usar el valor $\text{sen}A$ de la senoide $\text{sen}t$ para pronosticar el valor $f(A)$ del proceso en un momento futuro A (marcado en Fig.2 por la letra A). Sin embargo, si, en lugar de $\text{sen}t$,

el proceso auténtico se describe por la senoide $\text{sen}5t$ (cuyo grafico pasa a través de los mismos puntos históricos), entonces el valor $f(A)$ en el punto A puede ser totalmente diferente (compare los valores marcados con los discos negro y gris en Fig.2).

1.2. El caso general

Consideramos ahora dos ejemplos para demostrar que la representación falsa de ondas en una malla es algo muy típico.

Ejemplo 1. Sean $\mathbf{i} = \sqrt{-1}$, j el número del punto x_j de una malla, y h el tamaño de paso de la malla regular $h = x_j - x_{j-1}$. Aplicamos la fórmula $\exp(\mathbf{i}\sigma) = \cos \sigma + \mathbf{i}\text{sen}\sigma$ y la relación

$$\exp(\mathbf{i}kjh) = \{\exp(\mathbf{i}2m\pi)\}^j \exp(\mathbf{i}kjh) = \exp\{\mathbf{i}jh(k + 2m\pi/h)\},$$

donde k y m son dos números enteros arbitrarios. Así, es imposible distinguir las ondas $\exp(\mathbf{i}kx)$ y $\exp\{\mathbf{i}(k + 2m\pi/h)x\}$ dadas en puntos de la malla jh . Claro que sus períodos son totalmente diferentes.

Ejemplo 2. Sea n un número entero arbitrario, y sean $x_i = ih$ ($i = 0, 1, \dots, n+1$) puntos de una malla regular con tamaño h . Consideramos dos senoides

$$\text{sen}\frac{2\pi}{(n+1)h}x \quad \text{y} \quad -\text{sen}\frac{2\pi n}{(n+1)h}x$$

con los periodos $(n+1)h$ y $[(n+1)/n]h$, respectivamente. Es fácil demostrar que en cada punto x_i de la malla las senoides tienen los mismos valores:

$$f(x_i) = \text{sen}\frac{2\pi i}{(n+1)}, \quad i = 0, 1, \dots, n+1.$$

En efecto, sea $K = \pi/h$ el número de la senoide más corta que puede ser presentada en la malla (su período es igual a dos veces el tamaño de paso de la malla). Sustituyendo K en la identidad $\text{sen}kx \equiv \text{sen}[2K - (2K - k)]x$, y usando la fórmula correspondiente a $\text{sen}(\alpha - \beta)$, obtenemos

$$\text{sen}kx = \text{sen}\frac{2\pi}{h}x \cdot \cos\left(\frac{2\pi}{h} - k\right)x - \cos\frac{2\pi}{h}x \cdot \text{sen}\left(\frac{2\pi}{h} - k\right)x.$$

Ya que

$$\text{sen}\frac{2\pi}{h}ih = 0 \quad \text{y} \quad \cos\frac{2\pi}{h}ih = 1$$

en los puntos $x_i = ih$, tenemos

$$\text{sen}kih = -\text{sen}k^*ih, \quad (1)$$

donde

$$k^* = 2K - k. \quad (2)$$

Se deduce de esta forma que si se conocen sólo los valores en los puntos de la malla, entonces es imposible distinguir la onda con el número k de la onda con el número $2K - k$. Esto significa que si $k > K$, entonces la senoide con el número de onda k será presentada falsamente por la senoide con el número de onda $k^* = 2K - k$. Por ejemplo, si

$$k = \frac{2\pi}{[(n+1)h]}, \quad \text{entonces} \quad k^* = \frac{2\pi}{h} - \frac{2\pi}{(n+1)h} = 2\pi \left(\frac{n+1}{n} \right)^{-1}. \quad (3)$$

Así, la senoide $\text{sen}kih$ con el periodo de $(n+1)h$ puede ser interpretada erróneamente como la senoide $-\text{sen}k^*ih$ con el período de $[(n+1)/n]h$, y viceversa. En el caso particular, cuando $n = 3$ [1], ambas senoideas están presentadas en la Fig.3.

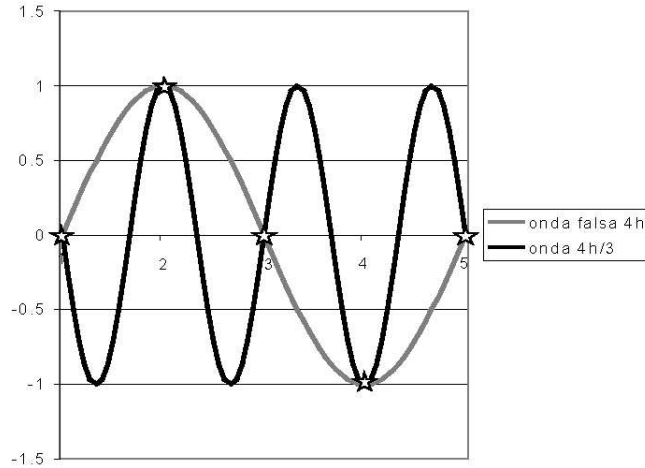


Figura 3: Representación falsa de la onda con el período $T = 4h/3$ por la onda con el período $T = 4h$ ($n = 3$).

1.3. Efectos negativos

La representación falsa de ondas en una malla tiene varios efectos negativos. Consideremos ahora algunos de ellos.

a) Errores al determinar el período de un proceso discreto. Los resultados de las primeras dos secciones nos muestran que la determinación del período de oscilación de un campo geofísico usando los datos de medición en puntos de una malla es un problema mal planteado, ya que no tiene solución única.

b) Errores en las series de Fourier. Cada función suave impar se representa por una serie de Fourier:

$$u(x) = \sum_{k=1}^{\infty} a_k \operatorname{sen} kx \quad (4)$$

en el intervalo $[-\pi, \pi]$. Cuando la misma función está definida sólo en puntos de una malla x_i , tenemos

$$u(x_i) = \sum_{k=1}^{\infty} a_k \operatorname{sen} kx_i. \quad (5)$$

En este caso, por la representación falsa de ondas (véase secciones 1.1 y 1.2), es difícil encontrar correctamente las amplitudes a_k y, en particular, las ondas energéticamente dominantes (con las amplitudes a_k más grandes).

c) Inestabilidad no lineal. Consideremos ahora la ecuación no lineal

$$u_t + uu_x = 0. \quad (6)$$

Demostramos ahora que la presencia del término no lineal en (6) genera inestabilidad. En efecto, es posible plantear la solución aproximada antisimétrica del problema en la forma de una suma finita de Fourier (método espectral [2,3]):

$$u(t, x) = \sum_{k=1}^K a_k(t) \operatorname{sen} kx, \quad (7)$$

donde la senoide con el número de onda máximo $K = \pi/h$ es la más corta que puede ser presentada en la malla. Entonces el término uu_x contendrá varios productos de dos armónicos con distintos números de onda: $\operatorname{sen} mx \cdot \cos nx$. Debido a la fórmula

$$\operatorname{sen} mx \cdot \cos nx = \frac{1}{2} [\operatorname{sen}(n+m)x + \operatorname{sen}(n-m)x], \quad (8)$$

los dos armónicos $\operatorname{sen} mx$ y $\cos nx$ generan otros dos armónicos: uno de escala más pequeña y otro de escala más grande, respecto de las escalas

de los armónicos originales. Así, una perturbación de la solución $u(t, x)$ en una parte del espectro de energía siempre genera las cascadas de energía en ambas direcciones: a las escalas más grandes y a las más pequeñas en comparación con las escalas de la perturbación (Fig.4).

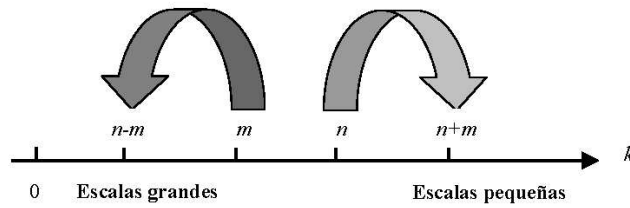


Figura 4: Las dos cascadas de energía.

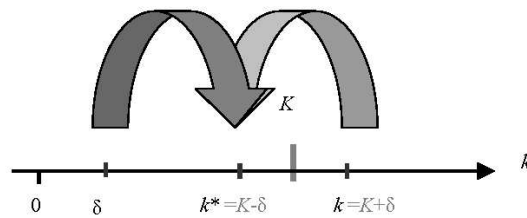


Figura 5: Cascadas de energía cercanas al número de truncación K debido a la representación falsa de ondas con los números $k > K$.

En los métodos espectrales, cuando las series de Fourier están truncadas por el número entero $K = \pi/h$, la solución en el momento inicial

$u(0, x)$ contiene sólo ondas $\text{sen}kx$ con $k \leq K$. Sin embargo, en el proceso de interacción no lineal aparecen los armónicos con números de onda $k > K$, los cuales se representan, según (1) y (2), por las ondas con números $k^* = 2K - k < K$. Así, hay una cascada falsa de energía al intervalo espectral $(0, K)$ desde el intervalo espectral $k > K$. En particular, cuando $n = K$ y m es un número pequeño, la fórmula (8) genera dos ondas $\text{sen}(K + m)x$ y $\text{sen}(K - m)x$. El número $K + m$ de la onda $\text{sen}(K + m)x$ es un poco mayor que K y dicha onda no se toma en cuenta por la fórmula (7). Sin embargo, según (1) y (2), su energía se representa falsamente como la energía de la onda $\text{sen}(K - m)x$ que ya pertenece a la serie (7) (Fig.5). Así, hay cascada falsa de energía a una parte del espectro situado cerca del número límite K (al intervalo $[K - \delta, K]$) de las escalas más pequeñas. Ya que también existe la cascada permanente de energía al mismo intervalo de las escalas más grandes, entonces esas dos cascadas causan la acumulación falsa de energía en los armónicos cuyos números de onda pertenecen al intervalo $[K - \delta, K]$ (Fig.6). Este es un mecanismo de inestabilidad conocido como *inestabilidad no lineal* [4-6].

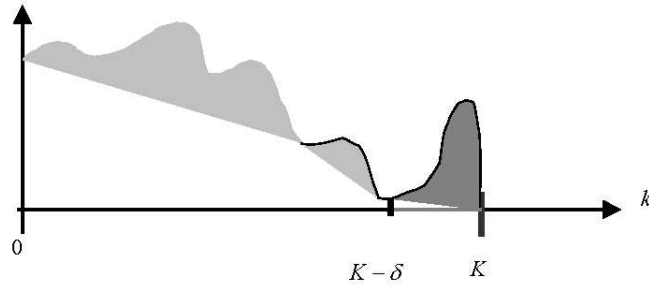


Figura 6: Cascadas de energía cercanas al número de truncación K debido a la representación falsa de ondas con los números $k > K$.

d) Errores de extrapolación. Sea $\{x_0, x_1, \dots, x_n\}$ un conjunto de nodos (diferentes) de una malla y sea $\{f_0, f_1, \dots, f_n\}$ un conjunto de valores de una función $f(x)$ en dichos nodos. Hay que reconstruir $f(x)$ en todo el intervalo $[x_0, x_n]$ de tal manera que

$$f(x_i) = f_i, \quad i = 0, 1, \dots, n. \quad (9)$$

El problema tiene un número infinito de soluciones. Pero si buscamos su solución en la forma de un polinomio algebraico $P_n(x)$ de grado no mayor a n tal que

$$P_n(x_i) = f_i, \quad i = 0, 1, \dots, n \tag{10}$$

(Fig.7), entonces la solución del nuevo problema existe, es única y se obtiene por medio de los polinomios de Lagrange [7]:

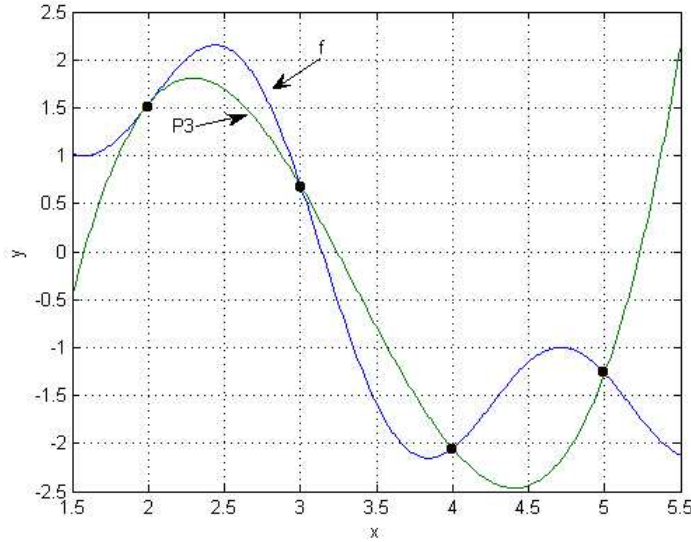


Figura 7: Aproximación de la función $f(x) = 2\text{sen}x + \text{sen}3x$ por el polinomio $P_3(x)$ en los nodos $x_i = 2, 3, 4$ y 5 .

$$P_n(x) = \sum_{k=0}^n L_{nk}(x) \cdot f_k, \tag{11}$$

donde

$$L_{nk}(x) = \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j}$$

es un polinomio algebraico de grado n (Fig.8).

El error de la interpolación (o extrapolación) es

$$R_n(x) = f(x) - P_n(x) = \omega_n(x) \frac{f^{(n+1)}(\xi(x))}{(n+1)!}, \tag{12}$$

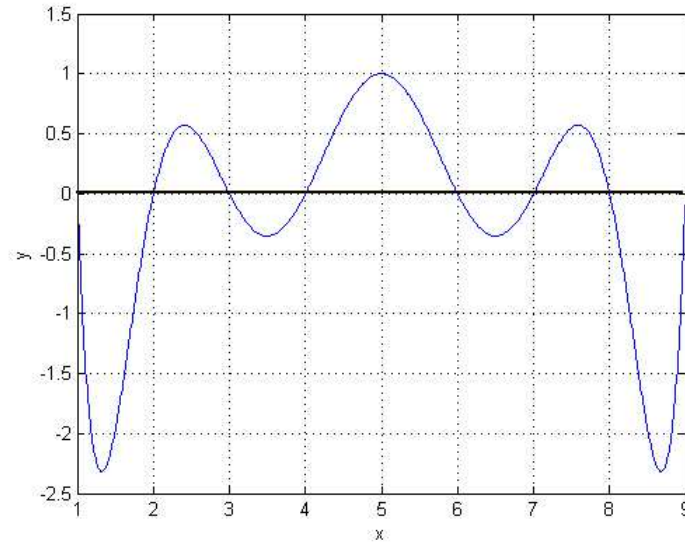


Figura 8: Polinomio de Lagrange $L_{8,4}(x) = \prod_{\substack{j=0 \\ j \neq 4}}^8 \frac{x-x_j}{x_4-x_j}$ para los nodos $x_j = j + 1, j = 0, 1, 2, \dots, 8$.

donde

$$\omega_n(x) = (x - x_0)(x - x_1) \cdots (x - x_n), \quad (13)$$

$f(x)$ tiene $n + 1$ derivadas continuas en el intervalo $[a, b] \equiv [x_0, x_n]$ y $\xi \in [a, b]$ es cierto punto desconocido del mismo intervalo. Hay que notar que el grado de suavidad de la función $f(x)$ es de gran importancia para determinar el error de interpolación. La figura 9 muestra los resultados de varias extrapolaciones (pronósticos) para $n = 1, 2$ y 3 . Se puede ver que los resultados dependen drásticamente del grado n . En geofísica dicho método se usa para extrapolar (hacer un pronóstico). Ya que la suavidad del campo pronosticado es desconocido, hay que ser cuidadoso en las conclusiones, pues los errores pueden ser más que el doble de la variabilidad de los datos (Fig.9).

2. Dispersión numérica

Describimos ahora otro efecto negativo relacionado con la presentación incorrecta de ondas en una malla. En el problema de transporte de un señal, dicho efecto causa su dispersión. Consideramos la ecuación

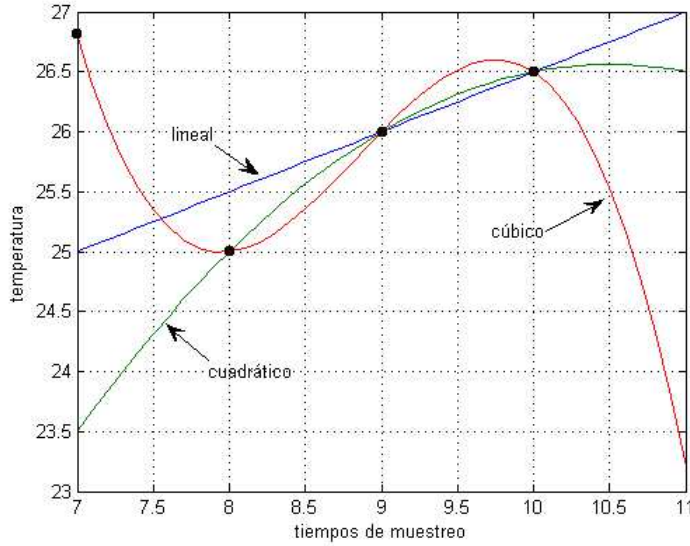


Figura 9: Resultados de la extrapolación en el nodo t_{11} de la temperatura T_{11} usando los polinomios lineal, cuadrático y cúbico ($T_7 = 26,8$, $T_8 = 25,0$, $T_9 = 26,0$, $T_{10} = 26,5$).

de transporte unidimensional

$$\frac{\partial \varphi}{\partial t} + u \frac{\partial \varphi}{\partial x} = 0, \quad (14)$$

con una velocidad constante u en el canal $D = \{0 < x < 1; 0 < t < 1\}$ periódico por x , es decir, $\varphi(1, t) = \varphi(0, t)$. Supongamos que en el momento inicial $t = 0$, la solución satisface la condición

$$\varphi(x, 0) = f(x), \quad (15)$$

donde $f(x)$ es una función diferenciable dada, también periódica por x . Es fácil verificar que la solución exacta del problema (14)-(15) es

$$\varphi(x, t) = f(x - ut) \quad (16)$$

[3-6]. Así, a lo largo de cada *línea característica*

$$x - ut = x_0$$

la solución $\varphi(x, t)$ es invariable, a saber,

$$\varphi(x, t) = f(x_0) = \text{Const.}$$

Considerando $f(x)$ como un señal inicial, obtenemos que, debido a (16), la señal no cambia su forma inicial en el proceso de propagación a lo largo del eje x . Si $f(x)$ es una onda, entonces (16) describe la propagación de la onda con la velocidad u .

Ya que

$$\frac{\partial}{\partial x} e^{\mathbf{i}kx} = \mathbf{i}k e^{\mathbf{i}kx}, \quad (17)$$

es decir, la onda $\exp(\mathbf{i}kx)$ no cambia su forma al diferenciarla, la ecuación (14) tiene la solución en la forma de una onda

$$\varphi(x, t) = \text{Re} \{ \Phi(t) \exp(\mathbf{i}kx) \},$$

donde $\Phi(t)$ satisface la ecuación de oscilación

$$\frac{d\Phi}{dt} + \mathbf{i}ku\Phi = 0$$

con la frecuencia $\omega(k) = ku$ y velocidad de fase invariante

$$\nu(k) = \frac{\omega(k)}{k} = u = \text{Const}. \quad (18)$$

Entonces todas las soluciones de onda $\varphi(x, t) = \text{Re} \{ \Phi(t) \exp(\mathbf{i}kx) \}$, independientemente del número de onda k , se propagan con la misma velocidad. Ya que la solución inicial $\varphi(x, 0)$ se representa, según el método de Fourier, por una serie de varios armónicos $\exp(\mathbf{i}kx)$ que se propagan con la misma velocidad, éste explica porque la solución analítica $\varphi(x, t)$ se traslada sin ningún cambio de su forma (véase (16)).

Analizamos ahora la propagación de las soluciones de onda

$$\xi_j(t) = \text{Re} \{ \Phi(t) \exp(\mathbf{i}kj h) \}$$

de la ecuación semidiscreta

$$\frac{d\xi_j}{dt} + u \frac{\xi_{j+1} - \xi_{j-1}}{2h} = 0, \quad (19)$$

que aproxima la ecuación (14) en el punto x_j de la malla espacial $x_i = ih$ con el paso (tamaño) de malla h ($i = 0, 1, \dots, K$, $x_0 = 0$ y $x_K = 1$). Aquí $\xi_j(t)$ es el valor de la solución en x_j . La derivada $\partial\varphi/\partial x$ ha sido aproximada en (19) por una diferencia finita central [1].

A diferencia del caso continuo (véase (17), la derivada de cada onda $\exp(\mathbf{i}kx)$ en el caso discreto ya no cambia su forma al diferenciarla:

$$\frac{\exp(\mathbf{i}k(j+1)h) - \exp(\mathbf{i}k(j-1)h)}{2h} = \mathbf{i} \frac{\text{sen}(kh)}{h} \exp(\mathbf{i}kj h) \quad (20)$$

y, por tanto, $\Phi(t)$ satisface la ecuación de oscilación

$$\frac{d\Phi}{dt} + \mathbf{i}k \left(u \frac{\text{sen}(kh)}{kh} \right) \Phi = 0 \quad (21)$$

con la frecuencia

$$\omega_h(k) = k\nu_h(k) = k \left(u \frac{\text{sen}(kh)}{kh} \right).$$

Entonces, a diferencia del problema continuo cuando todas las ondas tienen la misma velocidad de fase $\nu(k) = u$, cada onda del problema semidiscreto se propaga con la velocidad de fase

$$\nu_h(k) = \frac{\omega_h(k)}{k} = u \frac{\text{sen}(kh)}{kh} \quad (22)$$

que, a diferencia de (18), ya depende del número de onda k . Notamos que $|\nu_h(k)| < |u|$ para cada k , es decir, la velocidad de cada onda numérica $\nu_h(k)$ es más lenta que la velocidad u de las ondas analíticas, además, la velocidad $\nu_h(k)$ disminuya su valor cuando k aumenta (es decir, cuando la longitud de las ondas disminuye). Por ejemplo, si la resolución es buena ($kh \ll 1$ para ondas largas y una malla fina) entonces $\nu_h(k) \approx u[1 - \frac{1}{6}(kh)^2]$, es decir, el error en la velocidad de fase tiene el segundo grado respecto a kh . Sin embargo, para ondas cortas este error es bastante grande, y la onda con la longitud igual a dos veces el tamaño de la malla ($2h$) es inmóvil. En efecto, en este caso, $\xi_{j+1} = \xi_{j-1}$ por periodicidad, y $\frac{d\xi_j}{dt} = 0$. Notemos que el error en la velocidad de grupo es aún peor, ya que $\frac{\partial \omega_h(k)}{\partial k} = u \cos kh$. Este fenómeno se conoce como *dispersión numérica* de las ondas de distintas escalas (Fig.10) [3-6].

3. Aproximación, viscosidad numérica y estabilidad

En esta sección, la ecuación de transporte (14) se usa para mostrar las conexiones que existen entre la aproximación del problema original, la viscosidad numérica y la estabilidad del esquema.

3.1. Aproximación y viscosidad numérica

Consideremos ahora el esquema de Godunov [5,6]

$$\frac{\xi_i^{n+1} - \xi_i^n}{\tau} = -u \frac{\xi_i^n - \xi_{i-1}^n}{h}, \quad (23)$$

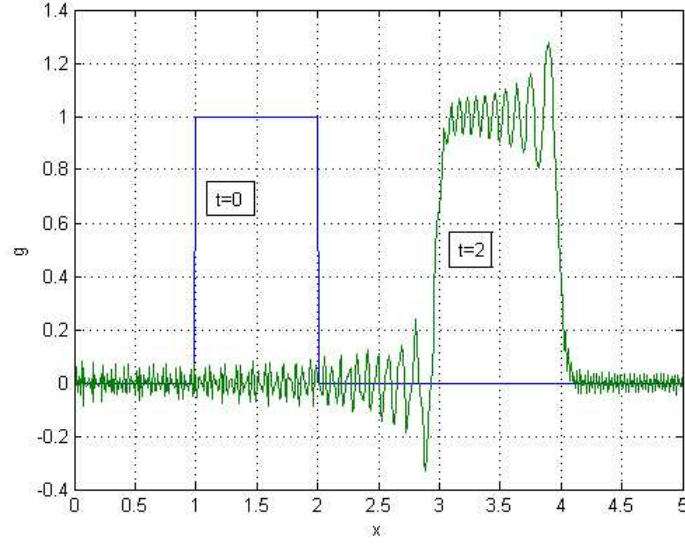


Figura 10: Efecto de la dispersión numérica.

donde $u = \text{Const} > 0$, h es el tamaño de la malla espacial $x_i = ih$ y τ es el tamaño de la malla temporal $t_n = n\tau$ ($n = 0, 1, \dots, N$). El esquema es explícito, ya que hay sólo un término incógnito ξ_i^{n+1} en (23) que se calcula fácilmente por la fórmula

$$\xi_i^{n+1} = \xi_i^n - \frac{\tau u}{h} (\xi_i^n - \xi_{i-1}^n).$$

Supongamos que en el momento $t_n = \tau n$ la solución numérica representa una señal localizada en el punto x_m de la malla espacial:

$$\xi_m^n = \delta, \quad \xi_j^n = 0, \quad j \neq m.$$

Entonces los valores de la solución numérica en el momento $t_{n+1} = \tau(n+1)$ son

$$\xi_{m+1}^{n+1} = \frac{\tau u}{h} \delta, \quad \xi_m^{n+1} = \left(1 - \frac{\tau u}{h}\right) \delta, \quad \xi_j^{n+1} = 0, \quad j \neq m, m+1. \quad (24)$$

Las ecuaciones indican que el esquema de Godunov es monótono y la señal se transporta desde su posición inicial de una manera correcta.

El esquema aproxima la ecuación de transporte (14) con el primer grado respecto a τ y h . En efecto, usando la serie de Taylor en la vecindad del punto (x_i, t_n) , obtenemos

$$\frac{\partial \varphi}{\partial t} + u \frac{\partial \varphi}{\partial x} = \frac{uh}{2} \frac{\partial^2 \varphi}{\partial x^2} - \frac{\tau}{2} \frac{\partial^2 \varphi}{\partial t^2} + o(\tau^2 + h^2). \quad (25)$$

Debido a (14),

$$\frac{\partial^2 \varphi}{\partial t^2} = u^2 \frac{\partial^2 \varphi}{\partial x^2}$$

y, por lo tanto, en lugar de aproximar la ecuación (14), el esquema (23), aproxima la ecuación de transporte y difusión

$$\frac{\partial \varphi}{\partial t} + u \frac{\partial \varphi}{\partial x} = \mu \frac{\partial^2 \varphi}{\partial x^2}, \quad (26)$$

donde

$$\mu = \frac{u}{2} (h - u\tau) = \frac{uh}{2} \left(1 - \frac{\tau u}{h}\right) \quad (27)$$

es el coeficiente de la viscosidad numérica (viscosidad artificial) [8]. Así, el esquema (23) introduce un fenómeno artificial conocido como *viscosidad numérica*.

Si $\tau u/h < 1$, entonces $\mu > 0$, y el problema no estacionario para la ecuación (26) está bien planteado según Hadamard [9], es decir, su solución es única y depende continuamente de los errores iniciales. La dependencia continua de una solución con respecto a los errores iniciales es la manifestación de su estabilidad. Precisamente la viscosidad numérica causa la disminución de la magnitud de la señal en el esquema (23) (véase, (24)).

Si $\tau u/h = 1$ entonces $\mu = 0$, y en la ecuación (25), cada uno de los términos incluidos en $o(\tau^2 + h^2)$ también es nulo. Así, el esquema de Godunov (23) tiene grado infinito de aproximación en h y τ cuando $\tau u/h = 1$, es decir, su solución (numérica) coincide con la solución exacta (analítica). Aparentemente, tal precisión es realizable sólo cuando $u = \text{Const}$. En la teoría de los métodos numéricos, la condición $\tau u/h \leq 1$ se conoce como *condición de Courant* o de *Courant-Fredrichs-Lewy* [3,10], y representa una relación entre los tamaños h y τ de las dos mallas.

Finalmente, si $\tau u/h > 1$, entonces en lugar de aproximar (14), el esquema (23) aproxima la ecuación

$$\frac{\partial \varphi}{\partial t} + u \frac{\partial \varphi}{\partial x} = -|\mu| \frac{\partial^2 \varphi}{\partial x^2} \quad (28)$$

con una viscosidad negativa. Es fácil demostrar que el problema no estacionario para la ecuación (28) está mal planteado en el sentido de Hadamard [9]. Efectivamente, su solución es inestable respecto de errores iniciales.

Así pues, existe una conexión lógica entre una buena aproximación del problema y la condición de Courant, es decir, para resolver bien la

ecuación de transporte (14), es necesario escoger los tamaños de malla τ y h de tal manera que se satisfice la condición $\tau u/h \leq 1$. Mostramos ahora que la misma condición garantiza la estabilidad del esquema de Godunov.

3.2. La condición de Courant y estabilidad

Investigamos la estabilidad espectral del esquema (23). El problema espectral

$$u \frac{\omega_k^p - \omega_{k-1}^p}{h} = \lambda_p \omega_k^p \quad (29)$$

para el operador del esquema de Godunov tiene las funciones propias

$$\omega_k^p = e^{ikph}, \quad (30)$$

y los valores propios

$$\lambda_p = \frac{u}{h} \left(2\text{sen}^2 \frac{ph}{2} + i\text{sen}ph \right). \quad (31)$$

Buscaremos la solución aproximada de la ecuación (14) de la siguiente forma

$$\xi_k^j = \sum_{p=-\infty}^{\infty} \zeta_p^j e^{ikph}, \quad (32)$$

donde k y j son los índices para las mallas en espacio y tiempo, respectivamente, y el coeficiente de Fourier ζ_p^j de la solución satisface la ecuación

$$\frac{\zeta_p^{j+1} - \zeta_p^j}{\tau} + \lambda_p \zeta_p^j = 0 \quad (33)$$

y, por consiguiente,

$$\zeta_p^{j+1} = T_p \zeta_p^j = (1 - \tau \lambda_p) \zeta_p^j. \quad (34)$$

Demostremos ahora que el esquema de Godunov es condicionalmente estable, a saber, es estable bajo la condición de Courant: $C = \frac{\tau u}{h} \leq 1$. En efecto, según (34), es necesario demostrar que

$$|1 - \tau \lambda_p| \leq 1. \quad (35)$$

Denotando $\alpha = ph/2$ en (31) obtenemos

$$|1 - \tau \lambda_p|^2 = (1 - 2C \text{sen}^2 \alpha)^2 + C^2 \text{sen}^2 2\alpha =$$

$$\begin{aligned}
 &= 1 - 4C\text{sen}^2\alpha + C^2 (4\text{sen}^4\alpha + \text{sen}^2 2\alpha) = \\
 &= 1 - 4C\text{sen}^2\alpha + 4C^2\text{sen}^2\alpha (\text{sen}^2\alpha + \cos^2\alpha) = \\
 &= 1 - 4C(1 - C)\text{sen}^2\alpha.
 \end{aligned}$$

El término $1 - 4C(1 - C)\text{sen}^2\alpha$ es no negativo, ya que

$$\max_{0 < x < 1} \{x(1 - x)\} = 1/4$$

y es menor que uno si $C = \frac{\tau u}{h} \leq 1$. Así, el esquema de Godunov es estable si $C = \frac{\tau u}{h} \leq 1$.

El siguiente ejemplo ilustra la conexión lógica que existe entre la aproximación de un problema continuo por un esquema y la estabilidad del esquema.

3.3. Viscosidad numérica y estabilidad

Consideremos ahora el esquema implícito

$$\frac{\xi_i^{n+1} - \xi_i^n}{\tau} + u \frac{\xi_i^{n+1} - \xi_{i-1}^{n+1}}{h} = 0. \quad (36)$$

A diferencia del esquema de Godunov (23), (36) se llama “implícito” ya que contiene más de un término incógnito (a saber, ξ_i^{n+1} y ξ_{i-1}^{n+1}) y, para hallar su solución en el momento t_{n+1} , se requiere resolver el problema de algebra lineal $A\vec{x} = \vec{b}$ con una matriz A y un vector \vec{b} dados. Es fácil demostrar que el esquema también aproxima la ecuación de transporte (14) con el primer grado respecto a τ y h . No obstante, en cada punto de malla (x_i, t_{n+1}) , (36) aproxima con un grado más alto la ecuación de transporte y difusión

$$\frac{\partial \varphi}{\partial t} + u \frac{\partial \varphi}{\partial x} = \mu \frac{\partial^2 \varphi}{\partial x^2}, \quad (37)$$

donde

$$\mu = \frac{uh + u^2\tau}{2} > 0 \quad (38)$$

es el coeficiente de viscosidad positivo para cualquier τ y h (Fig.11). Por lo tanto, el problema (37) siempre está bien planteado en el sentido de Hadamard, y el esquema (36) es incondicionalmente estable (para cualquier paso de tiempo τ). A diferencia de (36), el esquema explícito de Godunov es condicionalmente estable (sólo a condición de

Courant $u\tau/h \leq 1$). Para ilustrar efectos de discretización, consideramos aquí sólo los problemas unidimensionales. Sin embargo, los problemas prácticos (por ejemplo, de dispersión y control de contaminantes [11]) son multidimensionales y en este caso la realización del esquema (36) es más difícil que la del esquema de Godunov, el cual no requiere resolver el problema de algebra lineal $A\vec{x} = \vec{b}$ en cada paso temporal.

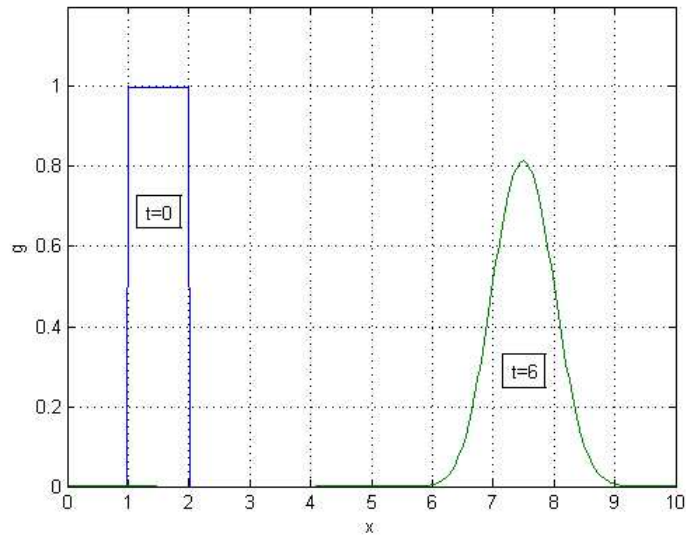


Figura 11: Efecto de la viscosidad numérica.

Agradecimientos

Este trabajo fue apoyado por los proyectos: PAPIIT IN105005 (UNAM) y FOSEMARNAT-2004-01-160 (CONACyT), y el Sistema Nacional de Investigadores (México).

Referencias

- [1] F. Mesinger & A. Arakawa, Numerical Methods Used in Atmospheric Models. V.1, GARP Publ. Ser., 17, WMO/ICSU, 1976.
- [2] B. Machenauer, Spectral Methods. In: Numerical Methods Used in Atmospheric Models., V.2, 1977.

- [3] D.R. Durrant, Numerical Methods for Wave Equations in Geophysical Fluid Dynamics. New York, Springer, 1999.
- [4] K. Morton & D. Mayers, Numerical Solution of Partial Differential Equations. Cambridge, 1994.
- [5] Yu. N. Skiba, Introducción a los Métodos Numéricos. UNAM, D.G.P y F.E., México, 2001.
- [6] Yu. N. Skiba, Métodos y Esquemas Numéricos: Un Análisis Computacional. D.G.P. y F.E., UNAM, México, 2005.
- [7] J.D. Faires & Burden R., Métodos Numéricos, Thomson, 2004.
- [8] G.I. Marchuk, Methods of Numerical Mathematics. Berlin, Springer-Verlag, 1982.
- [9] J. Hadamard, Lectures on Cauchy's Problem in Linear Partial Differential Equations, Dover, 1923.
- [10] R. Courant, & K. Friedrichs, Supersonic Flow and Shock Waves. Nueva York, Wiley-Interscience, 1948.
- [11] D. Parra-Guevara & Skiba Yu.N., Elements of the Mathematical Modelling in the Control of Pollutants Emissions. *Ecological Modelling*, 167 (3), 263-275,2003.