

UN METODO DE PSEUDOGRADIENTE PARA LA RESOLUCION NUMERICA DE UN PROBLEMA DE IDENTIFICACION DE PARAMETROS EN SISMICA

SANTIAGO DE VICENTE

*Departamento de Cálculo Numérico e Informática
E.T.S. de Ingenieros de Minas
Universidad Politécnica de Madrid. España*

RESUMEN

Se justifica el planteamiento de un problema de identificación de parámetros en sísmica de reflexión relativo a la función reflectividad y se propone un método eficiente para su resolución numérica. Para ello, se aplican las técnicas habituales de control óptimo para ecuaciones en derivadas parciales hiperbólicas de segundo orden. Las pruebas numéricas demuestran que la combinación de un método de tipo gradiente, especialmente diseñado teniendo en cuenta la estructura específica del problema, junto con una técnica de regularización de la no linealidad, proporciona excelentes resultados en comparación con otros métodos más clásicos.

SUMMARY

In this paper a parameter identification problem in seismic reflection for the reflectivity function is justified, and an efficient numerical method for its resolution is proposed. To achieve that, we apply the standard optimal control technics to second order hyperbolic partial derivatives equations. The numerical test confirm the performance of a combination of a special type of gradient method, conditioned by the problem's structure, with a regularization technique for the non-linearity.

INTRODUCCION

El objeto del presente trabajo es la aproximación numérica rápida y eficaz de un problema inverso relativo a experiencias de prospección geofísica mediante métodos sísmicos de reflexión. Las hipótesis del modelo físico escogido son las siguientes: Se supone que el medio (subsuelo) es horizontalmente estratificado de modo que una excitación en la frontera (explosión o vibración en la superficie) da origen a la propagación de ondas planas que se transmiten y se reflejan a través de las interfases que separan las capas. El problema consiste entonces en recuperar los parámetros que caracterizan el subsuelo en función de la profundidad partiendo, únicamente, de un registro (sismograma) del estado vibratorio del medio en un punto cercano a la excitación.

Entre las distintas aproximaciones a este problema cabe destacar las realizadas en dimensión uno de espacio por Kunetz¹, que pese a su rapidez de cálculo, adolecen de inestabilidad por excesiva sensibilidad a errores de medida y de redondeo. Otra aproxi-

Recibido: Mayo 1986

mación posible, que es la que nosotros tomamos aquí, es la que formula el problema inverso como un problema de control óptimo de control distribuido (parámetros del subsuelo) con observación en la frontera (sismograma) relativo a una ecuación hiperbólica de segundo orden, tanto en dimensión uno de espacio^{2, 3, 4, 5, 6, 7}, como en dimensión dos^{8, 9, 10}. Esta formulación permite asegurar la estabilidad del algoritmo frente al ruido, pero a cambio, requiere un gran tiempo de cálculo encareciendo la búsqueda de soluciones.

La utilización de métodos de "pseudogradiante" aquí propuestos para la aproximación del problema de control, salvaguarda por un lado la estabilidad, abaratando sustancialmente el cálculo efectivo de la solución.

Otras aproximaciones al problema inverso en sismica desde un punto de vista ligeramente diferente (métodos espectrales, ecuaciones integrales, etc.)^{11, 12, 13} no serán tomadas en cuenta aquí.

PLANTEAMIENTO DE UN PROBLEMA INVERSO

Considérese el problema hiperbólico de segundo orden en una dimensión de espacio:

$$\left. \begin{array}{l}
 \text{Hallar } u(z, t) \text{ para } (z, t) \in [0, Z] \times [0, T] \text{ solución de:} \\
 \rho \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial z} \left(\mu \frac{\partial u}{\partial z} \right) = 0 \quad \text{en } (0, Z) \times (0, T) \\
 -\mu \frac{\partial u}{\partial z} \Big|_{z=0} = g(t) \quad \text{en } [0, T] \\
 u \Big|_{t=0} = \frac{\partial u}{\partial t} \Big|_{t=0} = 0 \quad \text{en } [0, Z]
 \end{array} \right\} \quad (1)$$

Con las notaciones habituales:

$z \in [0, Z]$ representa la profundidad

$t \in [0, T]$ representa el tiempo

$\rho(z)$ es la densidad

$\mu(z)$ es el coeficiente de Lamé

$u(z, t)$ designa el desplazamiento.

el problema (1) modeliza la propagación de ondas planas en un medio unidimensional producidas por la excitación $g(t)$ en la frontera $z=0$.

Definiendo el espacio de funciones:

$$P = \left\{ (\rho, \mu) \in (L^\infty(0, Z))^2 / 0 < \rho \leq \rho(z) \leq \rho^+ < \infty, 0 < \mu \leq \mu(z) \leq \mu^+ < \infty, \forall z \in [0, Z] \right\}$$

puede caracterizarse la traza de la solución de (1) sobre la frontera $z=0$ mediante el resultado siguiente^{2, 5}:

Teorema 1: Si $(\rho, \mu) \in P$ y $g \in W^{1,1}(0,T)$, entonces el problema (1) admite una solución única u tal que $u(0, \cdot) \in L^\infty(0,T)$.

Si además $g \in W^{2,1}(0,T)$ y $g(0) = 0$ entonces la solución única de (1) es tal que $u(0, \cdot) \in W^{1,\infty}(0,T)$.

Por otra parte, mediante el cambio de variable:

$$\frac{dx}{dz} = \left(\frac{\rho(z)}{\mu(z)} \right)^{1/2} \tag{2}$$

y llamando $\sigma(x) = (\rho(x(z)) \cdot \mu(x(z)))^{1/2}$ el problema (1) se transforma en:

Hallar $u(x, t)$ con $(x, t) \in [0, X] \times [0, T]$ solución de:

$$\begin{cases} \sigma \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x} \left(\sigma \frac{\partial u}{\partial x} \right) = 0 & \text{en } (0, X) \times (0, T) \\ -\rho \frac{\partial u}{\partial x} \Big|_{x=0} = g(t) & \text{en } [0, T] \\ u \Big|_{t=0} = \frac{\partial u}{\partial t} \Big|_{t=0} = 0 & \text{en } [0, X] \end{cases} \tag{3}$$

con $X = \int_0^z \left(\frac{\rho(z)}{\mu(z)} \right)^{1/2} dz$

Bajo esta nueva forma, donde la notación empleada es:

x tiempo de viaje (“travel time”)

$\sigma(x)$ impedancia acústica

e introduciendo el espacio funcional:

$$\Sigma = \left\{ \sigma \in L^\infty(0, X) / 0 < \sigma_- \leq \sigma(x) \leq \sigma^+ < \infty \quad \forall x \in [0, X] \right\}$$

puede demostrarse⁵ el:

Teorema 2: Supuesto que $g \in W^{1,1}(0,T)$ y que $(\rho, \mu) \in P$ (resp. $\sigma \in \Sigma$), la solución única del problema (1) (resp. del problema (3)) es tal que su traza en la

frontera $z=0$ (resp. $x=0$) $w(t) = u(0, t) \quad \forall t \in [0, T]$ depende únicamente de:

- (i) $\sigma(x)$ con $x \in [0, X]$ siendo $X = T/2$
- (ii) $g(t)$ con $t \in [0, T]$.

Este resultado permite plantear un problema inverso (de identificación) sobre $\sigma \in \Sigma$ y no sobre $(\rho, \mu) \in P$, ya que la traza w depende únicamente de σ y no de (ρ, μ) como tal, pareja de parámetros independientes. En el apartado siguiente se verá otra de las ventajas de tomar las ecuaciones (3) en vez de las (1) como ecuaciones de estado.

Es bien sabido que una de las formas de abordar los problemas de identificación de coeficientes para ecuaciones en derivadas parciales es utilizar las técnicas clásicas del control óptimo en ecuaciones en derivadas parciales¹⁶ que permiten formular el problema de minimización de una funcional relativa a alguna de las funciones conocidas experimentalmente.

En el caso que nos ocupa, es posible, en la práctica, medir la traza en la frontera $x=z=0$, $\bar{u}(t)$, e intentar hallar la variable de control (coeficiente σ) que produce una traza $w(t)$ lo más cercana posible, en un sentido de mínimos cuadrados, a la traza experimental $\bar{u}(t)$.

Se puede entonces enunciar un problema inverso bajo la forma del problema de control óptimo siguiente:

Hallar $\sigma^* \in \Sigma^* \subset \Sigma$ tal que:

$$J(\sigma^*) \leq J(\sigma) \quad \forall \sigma \in \Sigma^* \quad (4)$$

$$\text{con } J(\sigma) = \frac{1}{2} \|w(t) - \bar{u}(t)\|_{L^2(0, T)}^2$$

siendo $w(t) = u(0, t)$ la traza sobre $x=0$ de la solución u de las ecuaciones de estado (3) y $\bar{u} \in C^0(0, T)$ la observación sobre la frontera $x=z=0$.

de manera que con las hipótesis

$$g \in W^{2,1}(0, T) \text{ y } g(0) = 0$$

no es difícil demostrar^{2, 5} la existencia de al menos una solución para el problema (4) en el espacio funcional Σ , es decir, se puede asegurar la existencia de la aplicación inversa:

$$C^0(0, T) \longrightarrow \Sigma$$

$$\bar{u} \longrightarrow \sigma$$

Por otra parte, puede demostrarse^{2, 5} que la aplicación directa:

$$\Sigma \rightarrow C^0(0, T)$$

$$\sigma \rightarrow w$$

es continua para una "distancia" definida sobre Σ que no es equivalente a la distancia inducida por la norma de $L^2(0, X)$. En cambio, esa misma "distancia" sí es equivalente a la definida por la norma de $L^2(0, X)$ sobre el espacio de funciones:

$$\Sigma_k = \Sigma \cap K$$

siendo K un subespacio compacto de $L^2(0, X)$.

Además, el espacio Σ es compacto para la "distancia" anterior de manera que, de todo lo dicho, puede afirmarse que la aplicación inversa es continua sobre Σ_k .

Esta propiedad de continuidad, o dicho de otro modo, de estabilidad frente a las perturbaciones sobre el sismograma \bar{u} (ruidos o errores de medida), sugiere que los buenos espacios en el sentido de Hadamard para buscar soluciones de (4) son del tipo $\Sigma^* = \Sigma_k$ sobre los cuales la "distancia" para la que la aplicación inversa es continua, es equivalente a la distancia asociada a $L^2(0, X)$.

UNA APROXIMACION DEL PROBLEMA DE CONTROL

Se describe a continuación la aproximación más simple posible del problema de control (4) mediante funciones σ constantes por trozos.

Sean:

- $T > 0$ y $N \in \mathbb{N}$

- $X = T/2$ y $h = X/(N+1)$

- El espacio de funciones constantes por trozos:

$$\Sigma^h = \left\{ (\sigma_1, \dots, \sigma_{N+1}) \in \mathbb{R}^{N+1} / \sigma_- \leq \sigma_k \leq \sigma^+, k=1, \dots, N+1 \right\}$$

- $x_k = kh$, $t_k = 2kh$, $k=0, 1, \dots, N$

- $\Gamma_0 = 0$, $\Gamma_k = (\sigma_k - \sigma_{k+1}) / (\sigma_k + \sigma_{k+1})$, $k=1, \dots, N$

Al considerar el espacio Σ^h se está considerando de hecho, un medio unidimensional horizontalmente estratificado en $N+1$ capas homogéneas de impedancia σ_k ($k=1, \dots, N+1$) y de espesor (en tiempo de viaje) h .

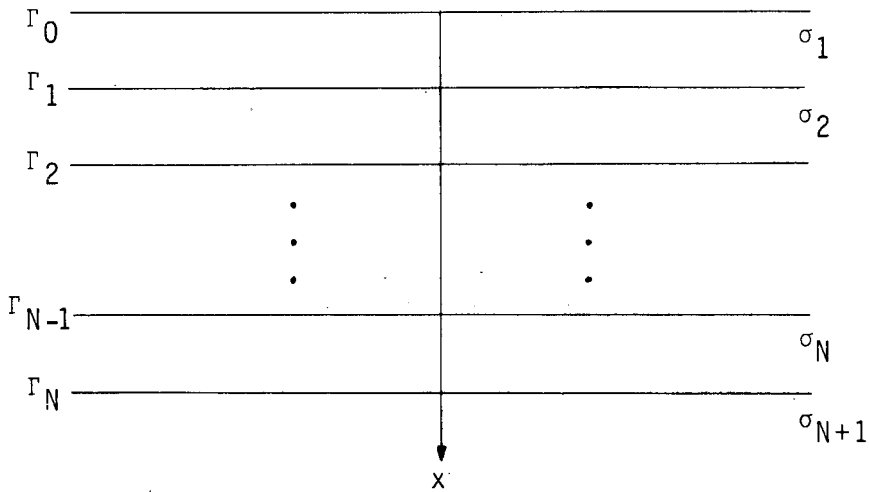


Figura 1

El paso de discretización en tiempo de viaje es, pues h y el paso de discretización en tiempo es $2h$.

Los valores discretos Γ_k ($k=0, \dots, N$) corresponden a lo que en Geofísica se conocen como coeficientes de reflexión en cada una de las interfaces entre capas y no son más que los valores discretos correspondientes a la función:

$$\Gamma(x) = -\frac{1}{2} \frac{\sigma'}{\sigma} = -\frac{1}{2} \frac{d}{dx} \log(\sigma) \quad (5)$$

donde σ' es la derivada de σ con respecto a x .

Aprovechando el hecho de que la ecuación diferencial de (3) es una ecuación hiperbólica lineal de 2.º orden cuyo sistema de características en el plano (x, t) viene dado por:

$$\frac{dx}{dt} = \pm 1 \quad (6)$$

puede aplicarse de manera sencilla y eficiente el método de las características de forma que, para funciones $\sigma \in \Sigma$ se tiene²:

$$\frac{\partial u}{\partial t}(0, t) = \frac{1}{\sigma_1} \sum_{k=0}^N \lambda_k g(t-t_k) \quad (7)$$

donde:

$$\lambda_i = \frac{1}{2}(\beta_{ii} + \gamma_{ii}) \quad i = 0, \dots, N \quad (8a)$$

estando dados β_{ii} y λ_{ii} por la resolución de

$$\left\{ \begin{array}{l} \beta_{i,j} = (1-\Gamma_{i,j}) \beta_{i,j-1} + \Gamma_{i,j} \gamma_{i-1,j} \\ \gamma_{i,j} = -\Gamma_{i,j} \beta_{i,j-1} + (1+\Gamma_{i,j}) \gamma_{i-1,j} \\ \text{para } i = 1, \dots, N ; j = 0, \dots, i \end{array} \right. \quad (8b)$$

con las condiciones de contorno:

$$\left\{ \begin{array}{l} \beta_{i-1} = 0 \\ \gamma_{i-1,i} = \beta_{i,i} \end{array} \right. \quad i = 1, \dots, N \quad (8c)$$

y las condiciones iniciales:

$$\left\{ \begin{array}{l} \beta_{0,0} = 0 \\ \gamma_{0,0} = 2 \end{array} \right. \quad (8d)$$

La ventaja de considerar la formulacion (3) con respecto a la (1) además de la reseñada en el apartado anterior, es la facilidad del cálculo en la aplicación del método de las características ya que el sistema característico viene dado por (6), es decir, las características son rectas con pendiente ± 1 , mientras que en la formulación (1) las características son curvas de pendiente $\pm(\mu/\rho)^{1/2}$, es decir, de pendiente variable con la profundidad z .

Por otra parte, en una experiencia de sísmica de reflexión no es realista suponer conocida la función g , pero en cambio y bajo ciertas condiciones², es posible medir la función impulso:

$$\Psi(t) = \frac{1}{\sigma_1} \int_0^t g(s) ds$$

por lo que, entrando en (7) puede escribirse:

$$u(0,t) = \sum_{k=0}^N \lambda_k \Psi(t-t_k) \quad (9)$$

lo que proporciona ya un método de cálculo de la traza de la solución de (3) sobre la frontera $x=z=0$.

Así planteada la aproximación del problema (3), es posible formular el problema de control (4) como un problema de optimización en dimensión finita:

Hallar el vector $\underline{g}^* \in \Sigma^N = \{ \underline{g} \in \mathbb{R}^N / \sigma_- \leq \sigma_j \leq \sigma^+, j=2, \dots, N+1 \}$

solución de:

$$J(\underline{g}^*) \leq J(\underline{g}) \quad \forall \underline{g} \in \Sigma^N \quad (10)$$

donde:

$$J(\underline{g}) = \frac{1}{2} \sum_{k=2}^{N+1} (w_k - \bar{u}_k)^2 = \frac{1}{2} \|\mathbf{w} - \bar{\mathbf{u}}\|_{\mathbb{R}^N}^2$$

siendo $w_k = u(0, t_k)$ ($t_k = 2kh$; $k=2, \dots, N+1$) la traza de la solución discreta de (3) en la frontera $x=z=0$ dada por (9) y \bar{u}_k la serie de observaciones discretas.

Si la función impulso $\Psi(t)$ es de soporte compacto en $[0, T]$ y cumple además las condiciones fijadas para g en los teoremas 1 y 2, que $\Psi(0) = 0$ y $\Psi(t_1) \neq 0$, no es complicado demostrar⁵ un resultado de existencia y unicidad para el problema de dimensión finita (10) donde se supone conocido el valor de σ_1 . Además, este resultado también es de estabilidad (en el sentido de la continuidad de la aplicación inversa) ya que Σ^N es un compacto de $L^2(0, X)$.

SENSIBILIDAD DE LA TRAZA RESPECTO A LOS COEFICIENTES

Manipulando convenientemente las expresiones (8) que resultan de la aplicación del método de las características se tiene que para $k = 1, \dots, N$:

$$\lambda_k = 2\Gamma_k \prod_{j=1}^{k-1} (1 - \Gamma_j^2) + P_k(\Gamma_1, \dots, \Gamma_{k-1}) \quad (11)$$

donde P_k es un polinomio de forma que:

$$\frac{\partial \lambda_k}{\partial \Gamma_k} = 2 \prod_{j=1}^{k-1} (1 - \Gamma_j^2) \quad (12a)$$

$$\frac{\partial \lambda_k}{\partial \Gamma_m} = \begin{cases} 0 & \text{si } m > k \\ -\frac{2\Gamma_k \Gamma_m}{(1 - \Gamma_m^2)} \frac{\partial \lambda_k}{\partial \lambda_k} + \frac{\partial}{\partial \Gamma_m} P_k(\Gamma_1, \dots, \Gamma_{k-1}) & \text{si } m < k \end{cases} \quad (12b)$$

Llamando:

$$\lambda(t) = \sum_{k=0}^N \lambda_k \delta(t - t_k)$$

siendo $\delta(\cdot)$ la delta de Dirac, puede escribirse la traza (9) como:

$$w = u(0, \cdot) = \lambda^* \psi$$

donde el signo * indica convolución; diferenciando:

$$\delta w = \delta \lambda^* \psi$$

se tiene la estimación siguiente¹⁴:

$$\|\delta w\|_{L^p(0,T)} \leq \|\delta \lambda\|_{L^p(0,T)} \|\psi\|_{L^1(0,T)} \quad (13)$$

con $1 \leq p \leq \infty$.

Por otra parte:

$$\delta \lambda = \sum_{j=1}^N \frac{\partial \lambda}{\partial \Gamma_j} \delta \Gamma_j = \sum_{j=1}^N \sum_{k=0}^N \frac{\partial \lambda_k}{\partial \Gamma_j} \delta(t-t_k) \delta \Gamma_j$$

Como por (12b):

$$\delta \lambda = \sum_{j=1}^N \frac{\partial \lambda_j}{\partial \Gamma_j} \delta(t-t_j) \delta \Gamma_j + \sum_{j=1}^N \delta(t-t_j) \delta \Gamma_j \sum_{k=j+1}^N \frac{\partial \lambda_k}{\partial \Gamma_j}$$

se puede escribir entonces:

$$\delta \lambda(t_i) = \left(\frac{\partial \lambda_i}{\partial \Gamma_i} + \sum_{k=i+1}^N \frac{\partial \lambda_k}{\partial \Gamma_i} \right) \delta \Gamma_i \quad (14)$$

Definiendo el espacio de coeficientes de reflexión admisibles:

$$R = \left\{ \underline{\Gamma} \in \mathbb{R}^N, |\Gamma_j| \leq \Gamma_{max} \right\}$$

se deduce de (12) que para $\underline{\Gamma} \in R$:

$$\left| \frac{\partial \lambda_k}{\partial \Gamma_k} \right| \leq 2 \quad (15a)$$

$$\left| \frac{\partial \lambda_k}{\partial \Gamma_j} \right| \leq \frac{4 \Gamma_{max}^2}{(1-\Gamma_{max}^2)} + C_1 \Gamma_{max}^m \text{ con } m \geq 0, k > j \quad (15b)$$

Así para $\underline{\Gamma} \in R$ se tiene en (14):

$$|\delta\lambda(t_i)| \leq C(\Gamma_{max}) |\delta\Gamma_i| \text{ con } 0 < C < \infty$$

y entrando en (13):

$$\|\delta w\|_{L^p(0,T)} \leq C(\Gamma_{max}) \|\psi\|_{L^1(0,T)} \|\delta\Gamma\|_{L^p(0,T)} \quad (16)$$

con $1 \leq p \leq \infty$

La estimación (16) permite enunciar el:

Teorema 3: La aplicación no lineal:

$$Tr: R \rightarrow L^p(0,T)$$

$$\underline{\Gamma} \rightarrow Tr \underline{\Gamma} = \{w_j\}_{j=1}^N = \{u(0,t_j)\}_{j=1}^N$$

es continua en el sentido de Lipschitz.

Este teorema (o equivalentemente, la expresión (16)), junto con el hecho de que el problema discreto (10) se dispone de una dependencia explícita de la traza $w(t)$ en función de los coeficientes de reflexión $\underline{\Gamma}$, hace aconsejable el plantear un nuevo problema inverso:

Hallar $\underline{\Gamma}^* \in R$ solución de:

$$J(\underline{\Gamma}^*) \leq J(\underline{\Gamma}) \quad \forall \underline{\Gamma} \in R \quad (17)$$

donde:

$$J(\underline{\Gamma}) = \frac{1}{2} \|w(\underline{\Gamma}) - \bar{u}\|_{\mathbb{R}^N}^2$$

siendo $\|\cdot\|_{\mathbb{R}^N}$ la norma euclídea sobre \mathbb{R}^N .

La figura 2 pretende ser una ilustración de la estimación (16); en ella se aprecia que dos funciones impedancia $\sigma(x)$ definidas sobre $[0,X]$ muy parecidas entre sí, producen, ante una misma impulsión, unas trazas $u(0,t)$ muy diferentes. La razón es que, si bien una impedancia puede obtenerse como una pequeña perturbación de la otra, los coefi-

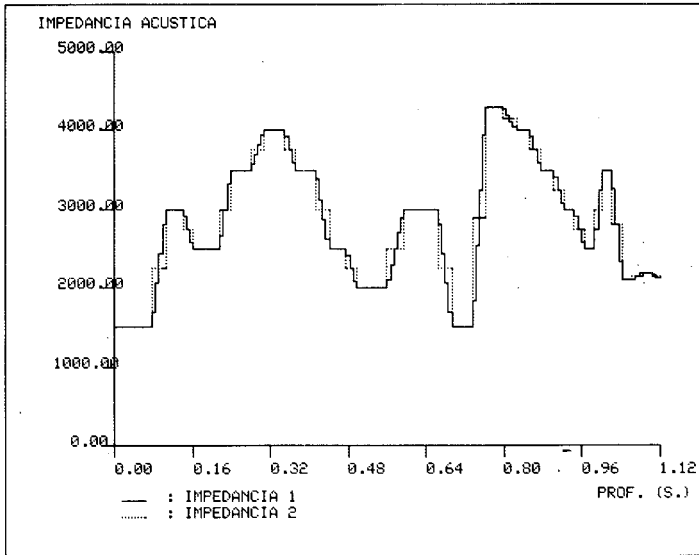


Figura 2.a

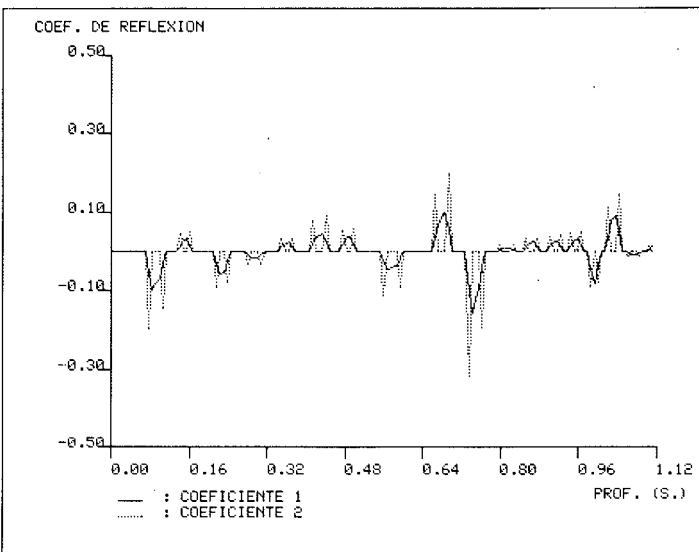


Figura 2.b

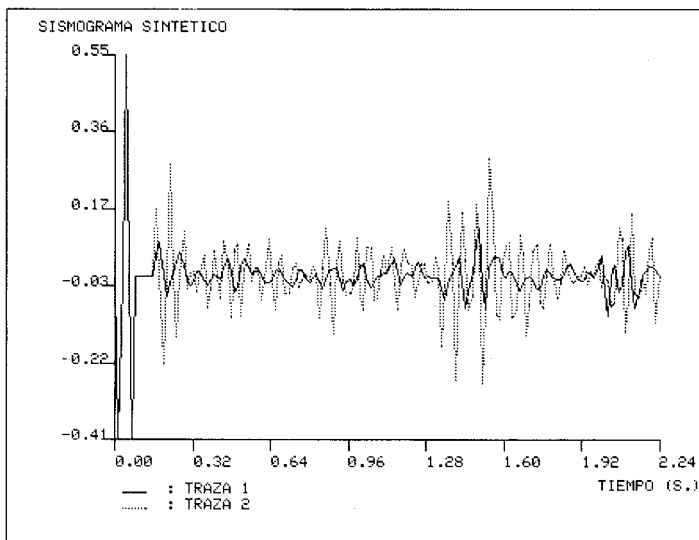


Figura 2.c

cientes de reflexión correspondientes son muy diferentes entre sí (advírtase este hecho en la expresión (5) donde el término σ' es el que produce aquí la diferencia entre coeficientes de reflexión). Dado que, como se deduce de (16), la distancia entre las trazas depende linealmente de la distancia entre los coeficientes de reflexión, no son de extrañar los resultados numéricos de la figura 2 que ponen de manifiesto que dos impedancias muy parecidas pueden originar trazas (sismogramas) muy distantes unas de otras.

METODO DE PSEUDOGRADIENTE CON REGULARIZACION

Considérese el problema de minimizar sobre el espacio:

$$R = \{ \underline{\Gamma} \in \mathbb{R}^N, |\Gamma_j| \leq \Gamma_{max}, j = 1, \dots, N \}$$

la funcional del problema (17):

$$J(\underline{\Gamma}) = \frac{1}{2} \| \mathbf{w}(\underline{\Gamma}) - \bar{\mathbf{u}} \|_{\mathbb{R}^N}^2 = \frac{1}{2} (\underline{\delta} \mathbf{w}, \underline{\delta} \mathbf{w})_{\mathbb{R}^N} \quad (18)$$

donde se ha utilizado la notación $\underline{\delta} \mathbf{w} = \mathbf{w}(\underline{\Gamma}) - \bar{\mathbf{u}}$.

Se trata pues, de hallar la solución de un problema clásico de mínimos cuadrados donde los "residuos" $\underline{\delta} \mathbf{w}$ dependen de forma no lineal de las incógnitas $\underline{\Gamma}$ (coeficientes de reflexión). Además la minimización debe ser efectuada sobre un subconjunto R , convexo de \mathbb{R}^N .

La idea natural para la minimización de la funcional (18) es utilizar los métodos iterativos tipo gradiente en los cuales la solución se obtiene descendiendo sucesivamente según las direcciones locales del gradiente de $J(\cdot)$, o bien según direcciones que de una u otra forma vengan determinadas por éste (gradiente conjugado, relajación, etc.)

Así, las derivadas (gradiente) de $J(\cdot)$ vienen dadas por:

$$\mathbf{g}(\underline{\Gamma}) = \nabla J(\underline{\Gamma}) = \mathbf{A} \underline{\delta} \mathbf{w} \quad (19)$$

donde el jacobiano \mathbf{A} es una matriz $N \times N$ tal que:

$$A_{ij} = \frac{\partial \delta w_j}{\partial \Gamma_i} = \frac{\delta w_j}{\delta \Gamma_i}, \quad i = 1, \dots, N; \quad j = 2, \dots, N+1 \quad (20)$$

es decir, las columnas de \mathbf{A} son los vectores gradientes de las trazas: ∇w_j .

Análogamente, las derivadas segundas de $J(\cdot)$ vienen determinadas por el hessiano \mathbf{G} , matriz $N \times N$ dada por:

$$\mathbf{G}(\underline{\Gamma}) = \mathbf{A} \mathbf{A}^T + \sum_{j=1}^N \delta w_j \nabla^2(\delta w_j) \quad (21)$$

siendo $\nabla^2(\delta w_j)$ las matrices hessianas definidas mediante:

$$(\nabla^2(\delta w_j))_{ik} = \frac{\partial^2(\delta w_j)}{\partial \Gamma_i \partial \Gamma_k} = \frac{\partial^2 w_j}{\partial \Gamma_i \partial \Gamma_k} \quad (22)$$

Entonces, la solución iterativa del problema de minimización (18) puede escribirse en cada iteración (método de Newton) como:

$$\begin{aligned} (i) \quad & \text{Resolución del sistema } \mathbf{G}^n \underline{\delta \Gamma}^n = \mathbf{g}^n \\ (ii) \quad & \underline{\Gamma}^{n+1} = P_R(\underline{\Gamma}^n - \underline{\delta \Gamma}^n) \end{aligned} \quad (23)$$

donde el superíndice k indica la k -ésima iteración y P_R designa el operador de proyección ortogonal sobre el convexo $R \subset \mathbb{R}^N$.

Los métodos aquí denominados de “pseudogradiante” consisten en elegir, mediante criterios de simplicidad, direcciones de descenso parecidas a las del gradiente calculado en (19) pero que no coinciden exactamente con él, es decir, en aproximar el gradiente por un vector suficientemente próximo a él pero más simple de obtener, que permita un ahorro sustancial de tiempo de cálculo salvaguardando, por otra parte, su calidad como dirección de descenso.

Teniendo en cuenta por un lado la expresión de la matriz jacobiana \mathbf{A} dada por (20), y por otro, las acotaciones (15) puede definirse un jacobiano aproximado \mathbf{A}_h por:

$$\mathbf{A}_h = \text{diag}(\mathbf{A}) \quad (24)$$

de manera que la dirección de descenso aproximada o “pseudogradiante” viene dada por:

$$\mathbf{g}_h = \mathbf{A}_h \underline{\delta \mathbf{w}} \quad (25)$$

La elección de (25) permite asegurar que:

$$\|\mathbf{g} - \mathbf{g}_h\|_2 \leq \|\mathbf{A} - \mathbf{A}_h\|_2 \|\underline{\delta \mathbf{w}}\|_2 \leq 0$$

ya que \mathbf{A} y \mathbf{A}_h tienen el mismo espectro, es decir, los mismos valores propios. La conclusión es que la diferencia entre los vectores de gradiente y “pseudogradiante” es únicamente una diferencia de dirección. Definiendo θ como el ángulo que forman

\mathbf{g} y \mathbf{g}_h y teniendo en cuenta la estructura particular de \mathbf{A} y \mathbf{A}_h así como las acotaciones (15) puede demostrarse¹⁵ que:

$$\cos \theta = \frac{(\mathbf{g}, \mathbf{g}_h)_{\mathbb{R}^N}}{\|\mathbf{g}\|_{\mathbb{R}^N}^2} = 1 - \epsilon(\Gamma_{max})$$

donde $\epsilon = \epsilon(\Gamma_{max})$ es una función creciente.

$$\mathbf{A} = \begin{pmatrix} \bullet & & & & \\ \times & \bullet & & & \\ \vdots & & & & \\ \times & \times \dots \times & \bullet & & \\ \times & \times \dots \times & & & \bullet \end{pmatrix} \quad \mathbf{A}_h = \begin{pmatrix} \bullet & & & & \\ & \bullet & & & \\ & & \ddots & & \\ & & & \bullet & \\ & & & & \bullet \end{pmatrix} \quad \begin{matrix} \bullet & \frac{\partial \lambda_j}{\partial \lambda_j} \\ \times & \frac{\partial \lambda_j}{\partial \Gamma_k} \quad k < j \end{matrix}$$

Figura 3. Jacobiano y Jacobiano aproximado

En la práctica, la sustitución del gradiente \mathbf{g} por el “pseudogradiente” \mathbf{g}_h permite simplificar muchísimo los cálculos ya que los elementos diagonales de \mathbf{A} son conocidos explícitamente, (12a), mientras que para el cálculo de los demás elementos distintos de cero del jacobiano es necesario plantear y resolver (¡en cada iteración!) las ecuaciones del estado adjunto que proporciona la teoría del Control Óptimo^{5, 8, 9, 10, 16} (ecuaciones en derivadas parciales hiperbólicas de segundo orden, en este caso).

Una vez definido un “pseudogradiente” (25) parece natural definir también un pseudo-hessiano de la forma siguiente:

$$\mathbf{G}_{h\alpha} = \mathbf{A}_h \mathbf{A}_h^T + \alpha \mathbf{I} \quad \text{con } \alpha \geq 0 \tag{26}$$

denominado “pseudo-hessiano regularizado”, en el cual el primer término corresponde a la aproximación del primer término del hessiano (21) y donde el segundo término es un término de regularización en el sentido de Tikhonov¹⁷ de la no linealidad correspondiente a la suma

$$\sum_{j=1}^N \delta w_j \nabla^2(\delta w_j)$$

de (21).

La elección $\alpha = 0$ (caso sin regularización) corresponde al método de Gauss-Newton (o Newton-Raphson en este caso particular) en el que se hace una aproximación lineal de los “residuos” $\tilde{\delta w}(\Gamma)$ despreciando la no linealidad del problema inicial.

De hecho, la regularización puede interpretarse de dos modos:

- a) Como una regularización en el sentido de Tikhonov considerando la no linealidad como un “ruido” que afecta al operador linealizado del operador no lineal considerado en el teorema 3.

- b) Como un método de paso restringido del tipo Levenverg-Marquardt¹⁸ que, como es bien sabido, estabiliza la solución de un problema no lineal mediante un método iterativo.

De esta forma es trivial deducir a partir de (23) con las modificaciones (25) y (26) que una iteración por el método propuesto queda reducida a:

$$(\Gamma_h^{n+1})_j = P_R \left[(\Gamma_h^n)_j - \frac{A_{jj}^n}{(A_{jj}^n)^2 + \alpha} \delta w_j \right] ; n \geq 0 ; j=1, \dots, N \quad (27)$$

donde el operador de proyección P_R sobre el convexo R viene dado por:

$$P_R(v_j) = \text{Min} \left\{ |v_j|, v_{max} \right\} \quad \forall v \in \mathbb{R}^N \quad (28)$$

Llamando:

$$\tilde{\mathbf{B}} = \sum_{j=1}^N \delta w_j \mathbf{V}^2(\delta w_j)$$

se demuestra¹⁵ que la convergencia del algoritmo (27)-(28) depende de la elección del parámetro $\alpha = \emptyset(\Gamma_{max}, \delta)$ siendo

$$\delta = \|(\mathbf{A} \mathbf{A}^T)^{-1} \tilde{\mathbf{B}}\|_2$$

y \emptyset una función creciente.

Dado que lo que se pretende minimizar no es más que la funcional $J(\cdot)$ (18), el criterio de convergencia más coherente con dicho problema de minimización es un criterio del tipo:

$$\frac{J(\tilde{\Gamma}^n)}{J(\tilde{\Gamma}^0)} \leq \epsilon \begin{cases} \text{sí: Parar el algoritmo} \\ \text{no: Hacer } n=n+1 \text{ y continuar} \end{cases} \quad (29)$$

donde ϵ es un parámetro prefijado a priori, cuyo valor, en la práctica, puede ser del orden de 10^{-6} .

RESULTADOS NUMERICOS

El proceso de experimentación numérica para testear el algoritmo propuesto ha sido el siguiente:

Etapa de simulación

En esta etapa se suponen conocidos:

- (i) Una función $\sigma(x)$ de repartición de impedancias (coeficientes de la ecuación en derivadas parciales (3)), como la de la figura (4a).
- (ii) El impulso $\psi(t)$ que en este caso será de los denominados en Geofísica de tipo Ricker.

Seguidamente, elegido un $T > 0$, se procede a discretizar el intervalo $[0, X] = [0, T/2]$ en $N + 1$ capas de longitud h , y el intervalo $[0, T]$ en subintervalos de longitud $2h$.

A continuación se calcula la traza en la frontera (sismograma) correspondiente al impulso $\psi(t)$ y a la impedancia $\sigma(x)$ mediante el método de las características (9).

En la figura (4.b) se representa la traza correspondiente a la impedancia σ de la figura (4.a) con $T = 2.24$ segundos y $N + 1 = 160$.

El impulso $\psi(t)$ (figura 4.b) aplicado es:

$$\psi(t) = \begin{cases} w(t) = u(0, t) & \text{para } t \in [0, T/14] \\ 0 & \text{para } t \in [T/14, T] \end{cases}$$

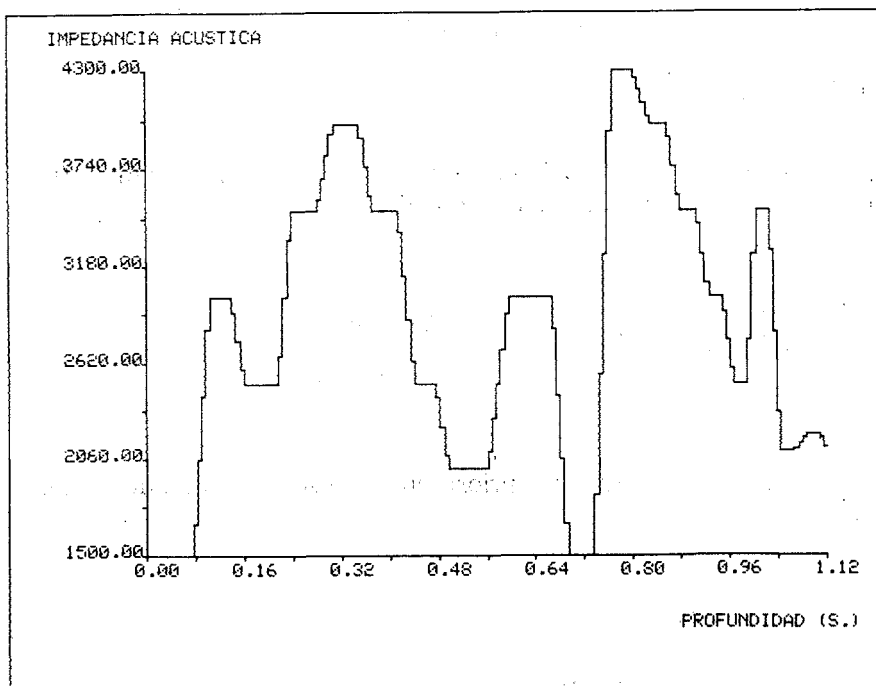


Figura 4.a. Test de simulación: Coeficientes de partida

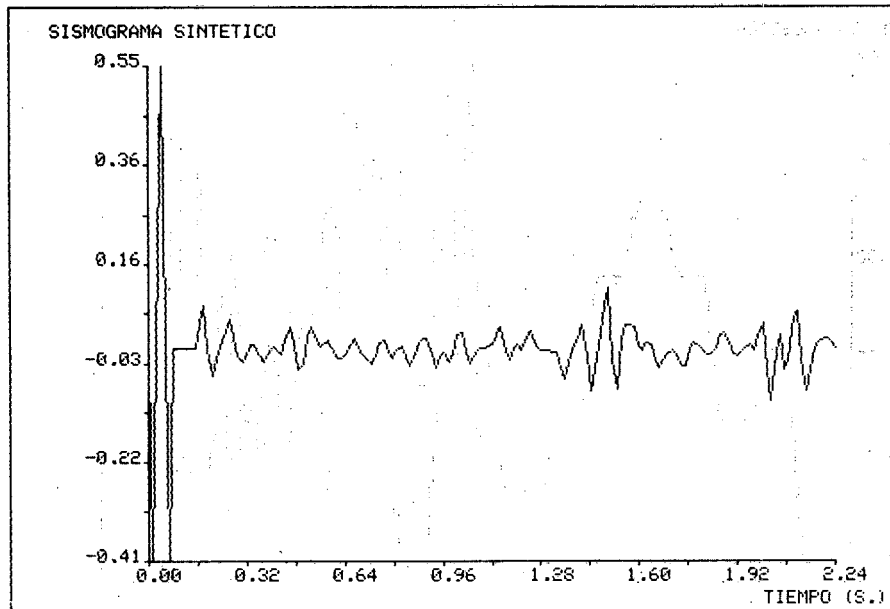


Figura 4.b. Test de simulación: traza correspondiente a $z=0$ para los coeficientes de la figura 4.a.

Etapa de identificación

El objeto de la identificación es la recuperación de la impedancia σ a partir del conocimiento del sismograma experimental (la traza $w(t) = u(0,t)$) y de la función impulso $\psi(t) \forall t \in [0, T]$. Además se supondrá conocido el valor $\sigma(0)$ de la impedancia acústica en el origen, lo cual es una hipótesis realista en la mayoría de los casos.

En los ensayos numéricos realizados se han proporcionado una impedancia inicial σ^0 :

$$\sigma^0(x) = \sigma(0) \quad \forall x \in [0, X]$$

de manera que no se han tenido en cuenta posibles informaciones "a priori" sobre el valor medio de σ .

El número de incógnitas a identificar es 160 (las impedancias de cada capa σ_j , $j=2, \dots, N+1$) y se ha procedido resolviendo el problema en dimensión finita (10) mediante el algoritmo de "pseudogradiante con regularización" (27) - (28).

La versión del algoritmo implementada es una variante de (27) en la que se ha preferido cambiar el parámetro α de regularización por un parámetro θ de relajación de forma que (27) queda:

$$(r_{h\alpha}^{n+1})_j = P_R \left[(r_{h\alpha}^n)_j - \frac{\theta}{A_{jj}^n} \delta w_j^n \right]$$

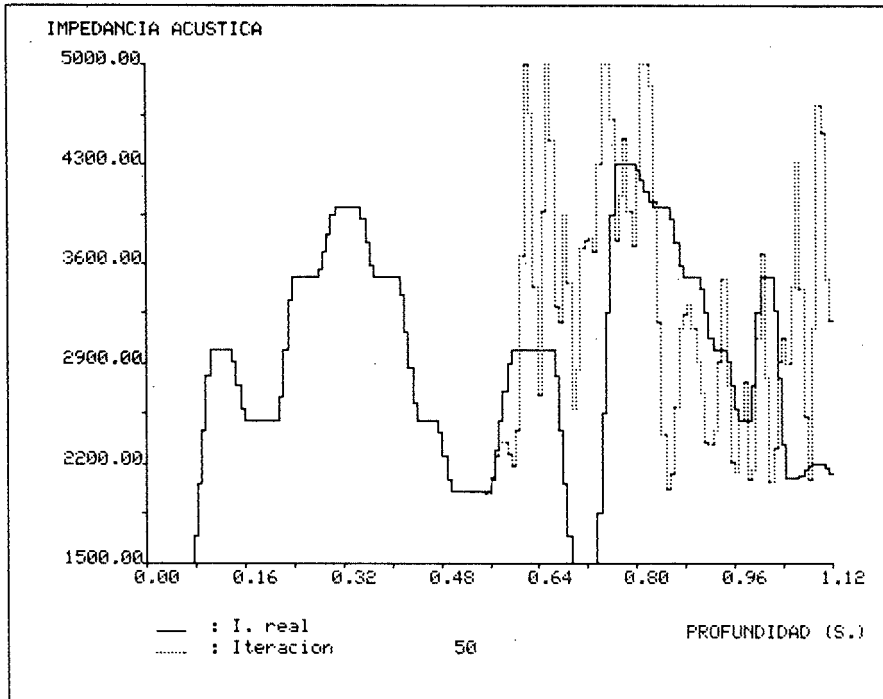


Figura 5.a. Test de identificación: Coeficientes recuperados en la iteración 50.

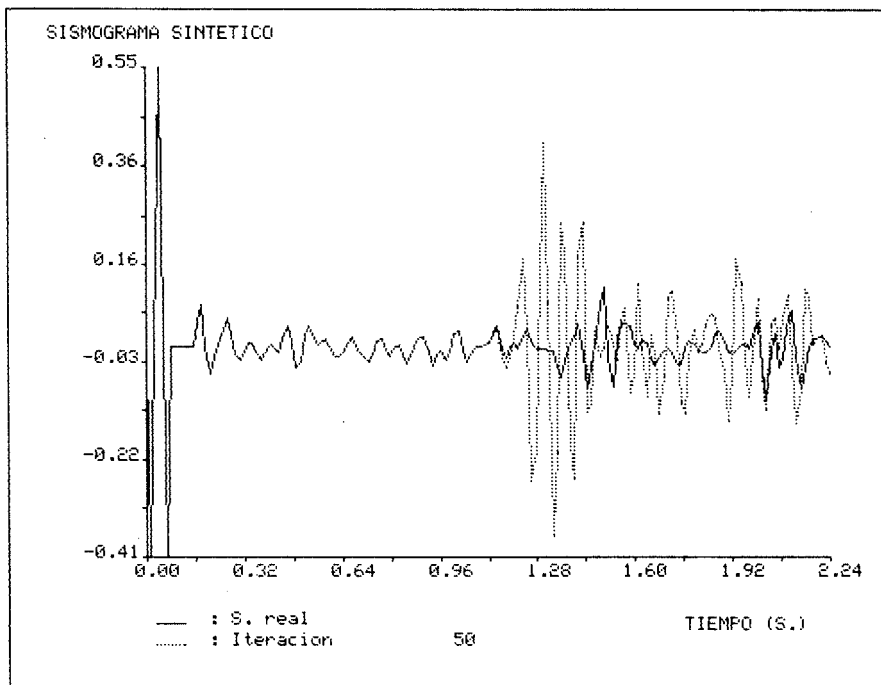


Figura 5.b. Test de identificación: Traza recuperada en la iteración 50.

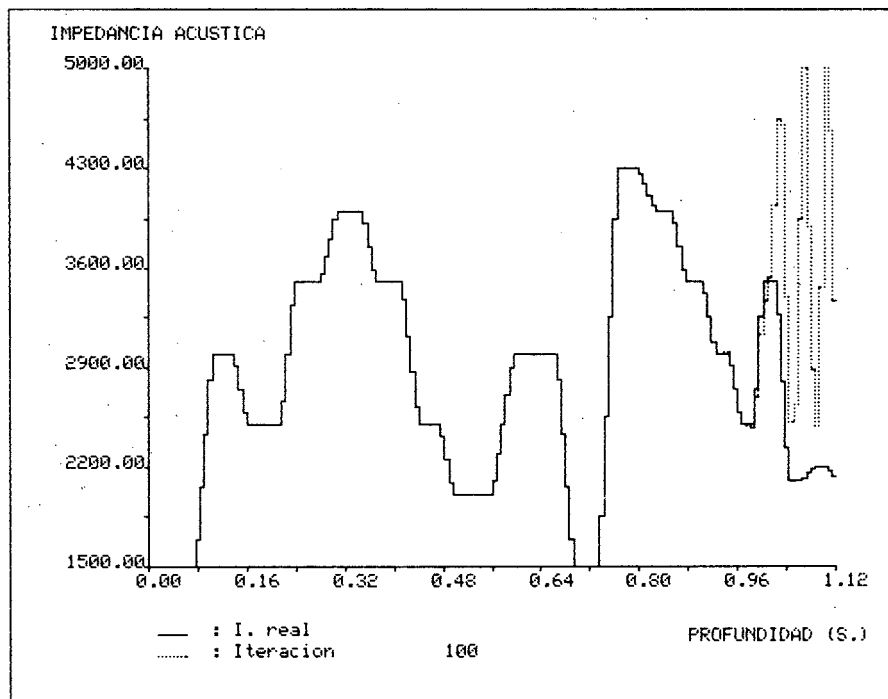


Figura 6.a. Test de identificación: Coeficientes recuperados en la iteración 100.

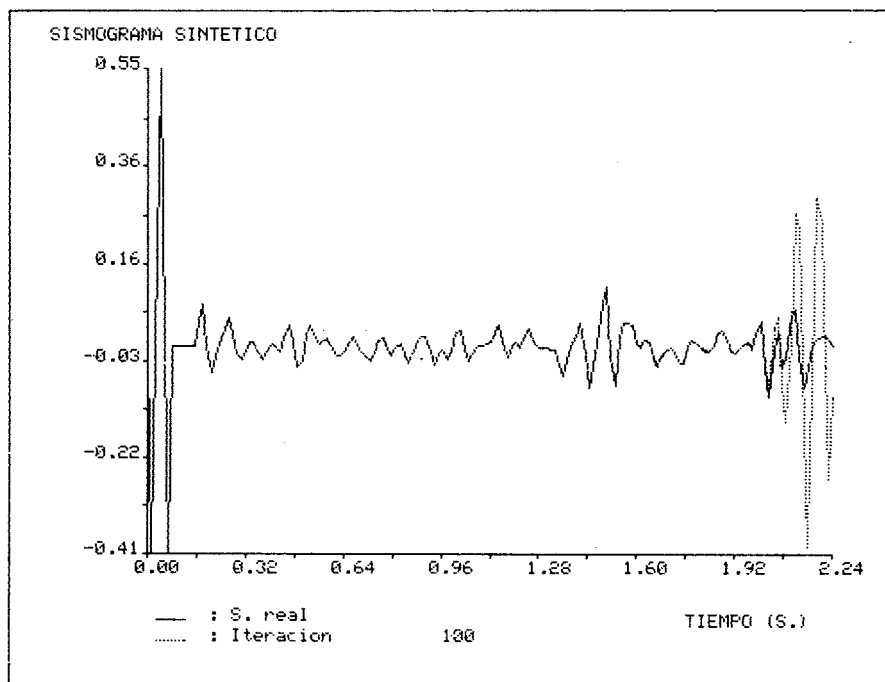


Figura 6.b. Test de identificación: Traza recuperada en la iteración 100.

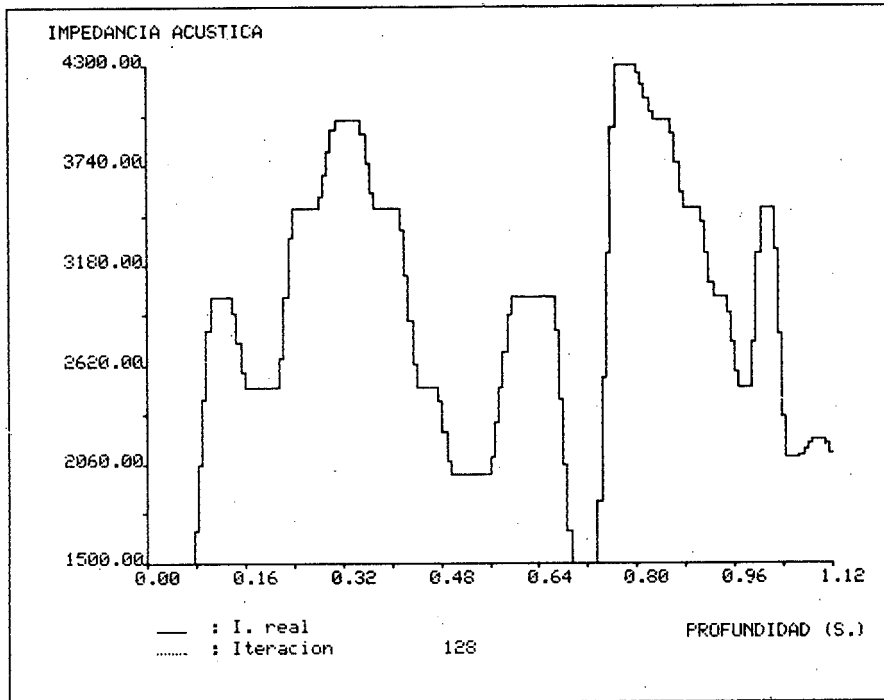


Figura 7.a. Test de identificación: Coeficientes recuperados en la iteración 128 (iteración final).

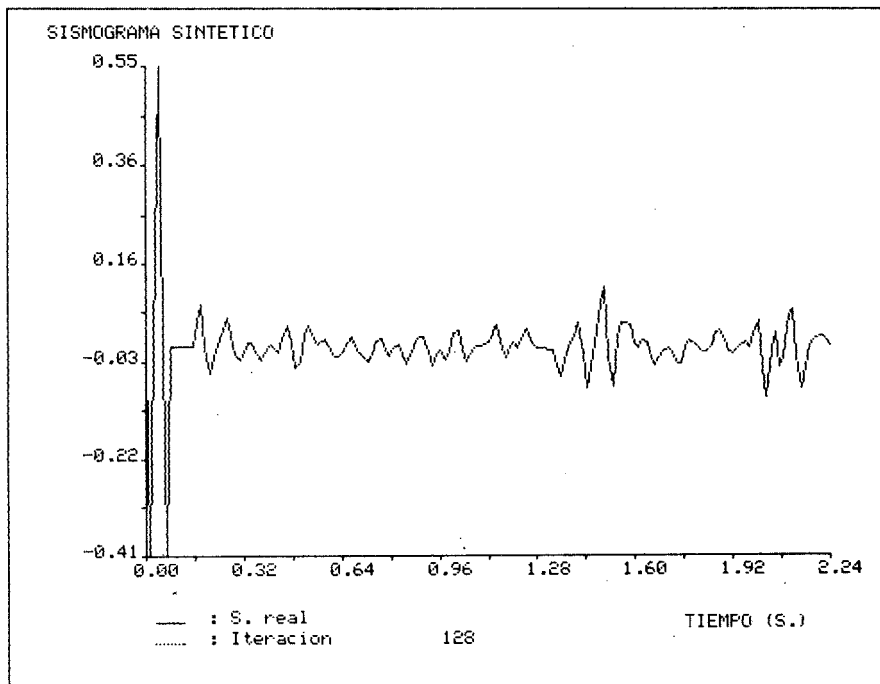


Figura 7.b. Test de identificación: Trazo recuperada en la iteración 128 (iteración final).

siendo, por tanto $\theta \in (0,1]$:

$$\theta = \frac{(A_{jj}^n)^2}{(A_{jj}^n)^2 + \alpha} \quad \text{con } \alpha \geq 0 \quad \text{y } j=1, \dots, N$$

El criterio de parada empleado para el algoritmo descrito es del tipo (29) donde $\epsilon = 10^{-6}$. Este criterio asegura la distribución de impedancias recuperada en la iteración final corresponde en efecto a la distribución de impedancias que ha generado el sismograma experimental (traza experimental) $\bar{u}(t)$.

Las figuras 5 a 7 muestran las impedancias recuperadas así como las trazas correspondientes en las iteraciones 50, 100 y 128 respectivamente, con un valor de $\theta = 0.6$ y $r_{max} = 0.2$. En la iteración 128 ($J(q^{128}) < 10^{-6}$) la traza calculada y la traza real, así como las impedancias respectivas son indistinguibles entre sí. En el test numérico se ha supuesto una restricción (r_{max}) sobre r no muy fuerte ya que en la realidad física los valores de los coeficientes de reflexión son, con mucho, inferiores al valor supuesto para r_{max} .

La elección del parámetro $\theta \in (0,1]$ de relajación (o, equivalentemente, del parámetro α de regularización) ha sido hecha de modo experimental. Para un gran número de problemas testeados el parámetro óptimo, en cuanto a número de iteraciones necesarias para la convergencia, está comprendido entre 0.5 y 0.6. En la figura 8.a se representan, en función de θ , las iteraciones necesarias para resolver un problema con una sola superficie de discontinuidad y donde se ha efectuado una discretización en 100 capas. Para el mismo problema, la figura 8.b muestra claramente el efecto regularizante de

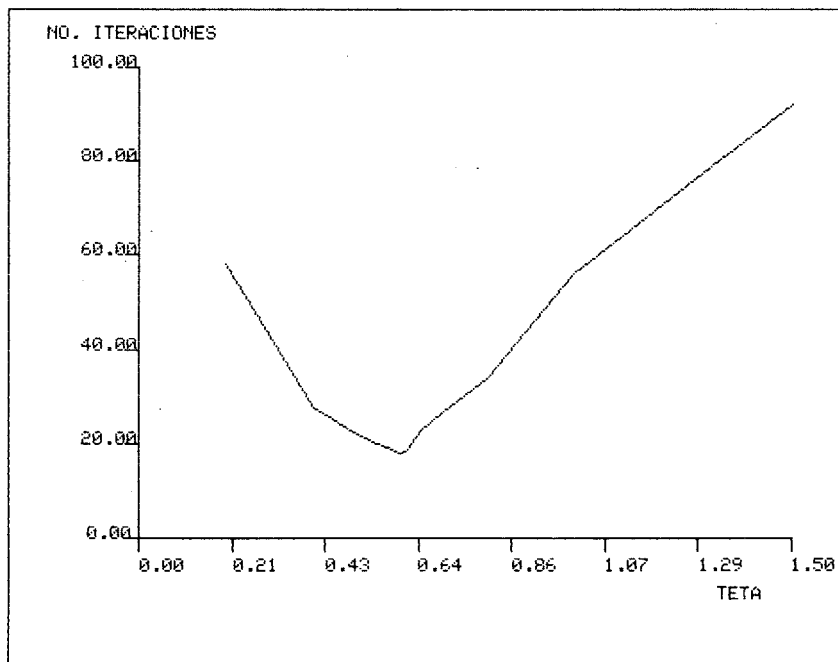


Figura 8.a. Número de iteraciones necesarias para la convergencia en función del parámetro θ de relajación. Caso de una discontinuidad y 100 capas.

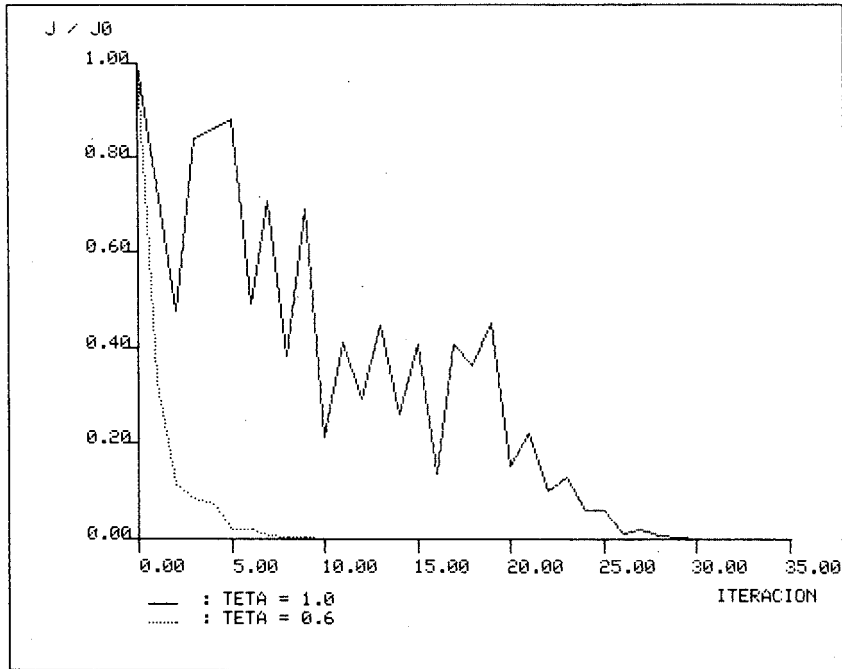


Figura 8.b. Evolución de la funcional normalizada J^k/J^0 a lo largo de las iteraciones. Nótese el efecto regularizante para $\theta=0.6$. Caso de una discontinuidad y 100 capas.

un $\theta=0.6$ sobre la funcional normalizada (J^k/J^0) en las sucesivas iteraciones k . Con $\theta=1$ (caso sin regularización) la evolución del criterio (18) normalizado es muy irregular debido al carácter fuertemente no lineal de los residuos δw con respecto a las incógnitas r . Este fenómeno de zig-zag, que se presenta cuando se linealiza demasiado el hessiano, puede llegar, en muchos casos, a estancar el algoritmo impidiendo la convergencia. Por tanto, en la mayoría de las ocasiones para tener convergencia será necesario regularizar, o lo que es lo mismo, restringir el paso en cada linealización (en cada iteración tipo Newton o tipo gradiente). En lenguaje geométrico, el no regularizar implica descender demasiado en un sólo paso según la dirección de g_h^n cruzando las líneas de nivel de $J(r^n)$ hasta "la otra ladera de valle".

Los tiempo de cálculo son del orden de 4.3 segundos de C.P.U. por iteración en un miniordenador VAX-11/730 sin acelerador de coma flotante para el caso de 160 capas ilustrado en las figuras 5 a 7. Los test numéricos realizados han puesto de manifiesto que el tiempo de C.P.U. necesario por iteración crece con el cuadrado del número de capas:

$$\text{Tiempo C.P.U./iteración} \approx C(N+1)^2$$

Sobre el procesador antes mencionado $C \approx 1.7 \times 10^{-4}$.

CONCLUSIONES Y COMENTARIOS FINALES

Frente a los planteamientos clásicos¹ del problema sísmico inverso 1-D y a los más recientes^{2, 3, 4, 5, 6} formulándolo como un problema de control óptimo relativo a los coeficientes (impedancias) de la ecuación de ondas escrita en variables de tiempo de llegada, justificamos una formulación del problema de control en términos de la función de reflectividad (5) (derivada logarítmica de los coeficientes de la E.D.P. (3)).

La utilización de los métodos clásicos de la teoría de optimización se ha llevado a cabo con una serie de modificaciones importantes:

- (i) La dirección de descenso marcada por el gradiente de la funcional no cuadrática (18) se ha sustituido por una dirección denominada "pseudogradiente". Esta dirección de descenso es mucho más sencilla de obtener directamente a partir de la expresión explícita del operador traza (9) a través del método de las características, evitando la resolución en cada iteración de las ecuaciones adjuntas proporcionadas por la teoría del Control Óptimo.
- (ii) La adopción de una estrategia de regularización de tipo Tikhonov para la no linealidad del problema. Se puede considerar que este método de control del paso en cada iteración proviene de una perturbación de la funcional (18), sustituyéndola por una funcional de Tikhonov:

$$J_{\alpha}(\mathbf{r}) = \frac{1}{2} \|\mathbf{w} - \bar{\mathbf{u}}\|_{\mathbb{R}^N}^2 + \frac{\alpha}{2} \|\mathbf{r}\|_{\mathbb{R}^N}^2$$

donde la minimización del término $\frac{\alpha}{2} \|\mathbf{r}\|_{\mathbb{R}^N}^2$ proporciona un modelo de "mínima estructura"¹⁹.

- (iii) El planteamiento del problema de optimización en dimensión finita (17) sobre un convexo R de \mathbb{R}^N . La proyección realizada de las soluciones sobre el convexo R (del mismo tipo que los llamados conjuntos regularizantes⁸) hace que el problema inverso sea continuo, es decir, que la solución \mathbf{r} del problema no lineal de mínimos cuadrados (17) dependa continua y uniformemente de las trazas experimentales $\bar{\mathbf{u}}$ en el sentido de L^2 . Esta proyección se realiza de forma trivial sobre cada incógnita por separado.

Los test numéricos revelan (para el caso en el que la medida experimental $\bar{\mathbf{u}}$ esté exenta de ruido) una convergencia excelente, obtenido, además, a unos costos computacionales muy bajos. En efecto, si se compara el algoritmo propuesto con las soluciones al problema de control (10) resuelto por Bamberger et al.^{2, 3, 4, 5} penalizando la funcional mediante un término no diferenciable (para asegurar la continuidad de la aplicación inversa) se advierte que, para un ejemplo prácticamente igual al de las figuras 7 a 9, se requieren al menos 1800 iteraciones de un algoritmo de subgradiente frente a las 128 iteraciones que necesita nuestro algoritmo. Además cada iteración de subgradiente resulta mucho más cara que una iteración de "pseudogradiente" con regularización. Los tiempos de cálculo son enormemente expresivos a este respecto (Tabla I).

TABLA I
COMPARACION METODOS DE RESOLUCION

Referencia	Bamberger et al. ⁴	Bamberger et al. ²	Presente
Método de resolución del problema directo	Elementos finitos de Lagrange de grado 3	Método de las características	Método de las características
Incógnitas	σ (funcional (10))	σ (funcional (10))	r (funcional (17))
Nº INCOGNITAS	160	160	160
Método de Optimización	Gradiente conjugado con subgradiente	Gradiente conjugado con subgradiente	Pseudogradient con regulación
Nº ITERACIONES	>> 400	1800	128
Procesador	CDC 7600	CDC 7600	VAX-11/730 sin procesador de coma flotante
C.P.U.	>> 4 h. 30 min.	≈ 20 min.	9 min. 20 seg.

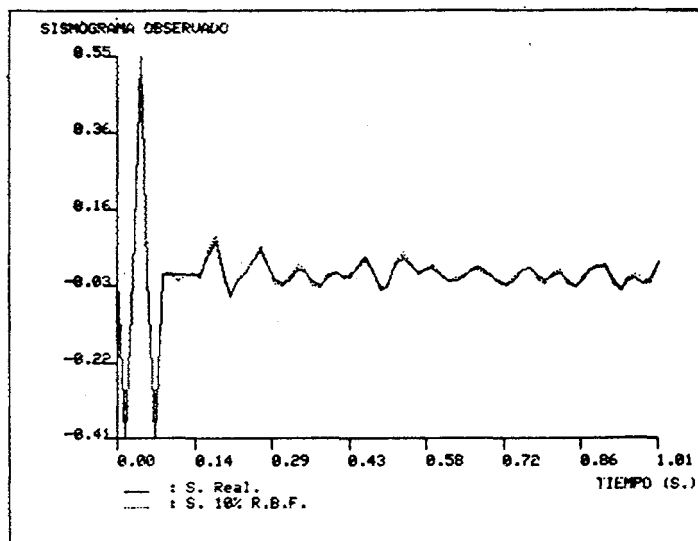


Figura 9.a. Test de simulación: Traza en $z=0$ y ruido baja frecuencia de 10%.

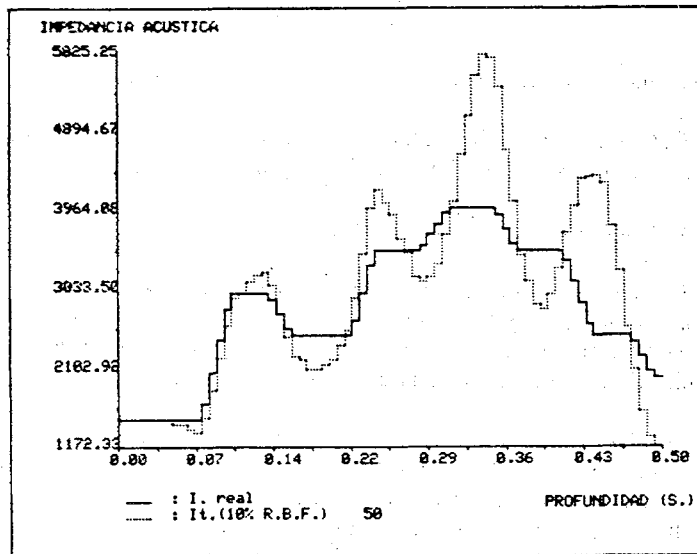


Figura 9.b. Test de identificación: Coeficientes recuperados en la iteración 50. $\theta = 0.5$, $r_{max} = 0.5$. Ruido baja frecuencia del 10%.

El estudio preciso de los resultados de estabilidad frente a perturbaciones debidas al ruido en las medidas \bar{u} está actualmente en período de desarrollo, siendo los resultados parciales muy esperanzadores a este respecto.

En este sentido, el autor¹⁵ ha realizado un estudio teórico de la continuidad en el sentido de Lipschitz de la aplicación inversa cuyo planteamiento funcional riguroso excede los límites de este trabajo. Como ilustración del comportamiento de los algoritmos aquí propuestos frente al ruido, se presenta, en la figura 9, un test realizado añadiendo a la traza de partida, figura 9.a un ruido gaussiano de media cero y de un nivel determinado (10% en media cuadrática, para este caso) que se ha convolucionado con el impulso $\psi(t)$ para generar un contenido en frecuencia del mismo tipo que el contenido en frecuencia del sismograma \bar{u} . Como se advierte en la figura 9.b un pequeño nivel de ruido baja frecuencia en la traza induce variaciones mucho mayores en la distribución de coeficientes $\alpha(x)$, que son reproducidos perfectamente por el algoritmo aquí expuesto.

Finalmente, también está en fase de estudio la generalización de este tipo de métodos para problemas en 2-D.

REFERENCIAS

1. G. Kunetz. "Quelques exemples d'analyse d'enregistrements sismiques". *Geophysical Prospecting*, Vol. **11**, (1963).
2. A. Bamberger, G. Chavent, P. Lailly. "Etude mathématique et numérique d'un problème inverse pour l'équation des ondes à une dimension". *Rapport de Recherche N.º 226*, IRIA-Laboria, (1977).
3. A. Bamberger, G. Chavent. "Sur le problème inverse en sismique". *Ann. Sc. Math. Québec*, Vol. **1**, no. **2**, 153-174, (1977).
4. A. Bamberger, G. Chavent. "Une application de la théorie du contrôle à un problème inverse en sismique". *Ann. Geophys.*, t. **33**, fasc. **1/2**, 183-200, (1977).
5. A. Bamberger, G. Chavent, P. Lailly. "About the stability of the inverse problem in 1-D wave equations. Application to the interpretation of seismic profiles". *App. Math. Optim.*, Vol. **5**, 1-47, (1979).
6. A. Bamberger, G. Chavent, C. Heman, P. Lailly. "Inversion of normal incidence seismograms". *Geophysics*, Vol. **47**, no. **5**, 757-770, (1982).
7. S. de Vicente. "Algunos problemas de control óptimo con observación en la frontera relativos a E.D.P. hiperbólicas. Aplicaciones a problemas inversos en Geofísica". *Seminario de Análisis Numérico* (Ed. Michavila), Universidad Politécnica de Madrid, (1985/86).
8. Y. H. Pao, F. Santosa, W. Symes. "Inverse problems of acoustic and elastic waves". *International Conference on Inverse Problems of Acoustic and Elastic Waves*. SIAM (Ed. F. Santosa et al.), 274-302, (1984).
9. A. Tarantola. "The seismic inverse problem" *International Conference on Inverse Problems of Acoustic and Elastic Waves*. SIAM (Ed. F. Santosa et al.), 104-181, (1984).
10. P. Kolb. "Least-squares 2-D inversion of horizontally stratified media". *Seventh International Conference on Computing Methods in Applied Sciences and Engineering*. Versailles, Francia (Preprints), 333-347, (1985).
11. N. Bleistein "Mathematical methods for wave phenomena" Ed. Academic Press, Orlando (Florida), (1984).
12. S. de Vicente. "Métodos iterativos para la resolución de un problema inverso relativo a la ecuación de ondas 2-D y 3-D". A aparecer en *Proceedings of International Symposium of Numerical Analysis* (ISNA), Madrid, España, (1985).
13. Y. M. Chen. "Generalized pulse-spectrum technique". *Geophysics*, Vol. **50**, no. **11**, 1664-1675, (1985).
14. H. Brézis. "Analyse Fonctionnelle. Théorie et applications". Ed. Masson, París, Francia, (1983).
15. S. de Vicente. *Análisis Numérico de las clases de funciones velocidad para el problema sísmico inverso 2-D y 3-D*. Tesis Doctoral. Universidad Politécnica de Madrid. Septiembre, (1986).
16. J. L. Lions. "Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles". Ed. Dunod, París, Francia, (1968).
17. C. W. Groetsch. "The theory of Tikhonov regularization for Fredholm equations of the first Kind". Ed. Pitman (Research Notes in Mathematics, 105), Boston, U.S.A., (1984).
18. R. Fletcher. "Practical methods of optimization. Vol 1: Unconstrained optimization". Ed. John Wiley & Sons. New York, U.S.A., (1980).
19. D. W. Oldenburg. "Inversion of band limited reflection seismograms" *International Conference on Inverse Problems of Acoustic and Elastic Waves*. SIAM (Ed. Santosa et al.), 48-64, (1984).