

REVISTA
DE LA
ACADEMIA
DE
CIENCIAS

Exactas
Físicas
Químicas y
Naturales

DE
ZARAGOZA



Serie 2.^a
Volumen 46

1991

INDICE DE MATERIAS

	<i>Págs.</i>
C. Romo Santos. — «Resolución de singularidades en característica cualquiera» ..	5
I. O. Argyros. — «On some projection methods for solving nonlinear operator equations with a nondifferentiable term»	17
I. O. Argyros and M. M. Losta. — «Integran equations for two-point boundary-value problems»	25
V. Novo. — «Optimización de funciones no derivables»	37
T. Grande, J. I. Montijano, L. Rández. — «Implicit output by continuous runge-kutta methods»	51
M. Membrado, A. F. Pacheco and J. Sañudo. — «A new smoothed-step function»	57
A. Abad, J. F. San Juan. — «Nota sobre el desarrollo de la función de perturbación en el movimiento orbital»	67
R. Cid Palacios and M. C. Fernández Quiñones. — «Motion of a gyroscope in a set of redundant variables»	75
A. Elipe, M. A. Soriano. — «Glaciaciones y astronomía»	83
J. Badal. — «Constraints in the filtering and inversion of long —and short— period seismic data and uncertainty in interpretation»	97
J. Mandado, M. Arce, L. F. Auqué, M. ^a J. Gimeno y J. M. Tena. — «Diseño metodológico de una campaña de prospección geoquímica multielemental en la red de drenaje del paleozoico de la cordillera ibérica»	113
L. F. Auqué, J. Tena, M. J. Gimeno, J. Mandado y J. Fernández. — «Las sugerencias termales de Los Baños (Teruel). I. Caracterización fisicoquímica y aproximación a los procesos de interacción agua-roca»	125
J. Gómez. — «Factores condicionantes de la reflectividad sísmica de la corteza inferior»	137
J. Tena, L. F. Auqué, J. Mandado y M. J. Gimeno. — «Las sugerencias termales de Los Baños (Teruel). II. Cálculos geotermométricos»	161

RESOLUCION DE SINGULARIDADES EN
CARACTERISTICA CUALQUIERA.

por

Concepción Romo Santos.

Departamento de Algebra y Fundamentos.
Facultad de Ciencias Matemáticas
UNIVERSIDAD COMPLUTENSE DE MADRID.

Resolución de singularidades en característica cualquiera

El problema de resolución de singularidades planteado en el lenguaje actual consiste en demostrar o probar la falsedad del teorema siguiente.

Teorema 1 Para toda variedad algebraica X definida sobre un cuerpo algebraicamente cerrado K , existe un morfismo $\pi: Y \rightarrow X$ tal que:

- 1) Y es no singular
- 2) π es propio
- 3) $\pi|_{\pi^{-1}(U)}: \pi^{-1}(U) \rightarrow U$ es un isomorfismo para un cierto abierto U denso en X .

Hironaka prueba el teorema anterior sobre un cuerpo de característica cero. Sus métodos de trabajo son los de Zariski y su herramienta fundamental la teoría de esquemas elaborada por Grothendieck.

En el proceso de resolución de singularidades de Hironaka hay que separar dos ideas distintas:

- i) Proceso de resolución local
- ii) Proceso de inducción y globalización

El proceso local de Hironaka se basa en un estudio de tallado del comportamiento de las bases de ideales por explosiones y la posibilidad de encontrar una base, base standard normalizada, con una cierta estabilidad respecto a dichas explosiones. A la vez se desarrolla un método de control del estado de la singularidad compuesto por dos series de funciones, la función de Hilbert-Samuel y los caracteres ν^* . Pero al pasar a la resolución de singularidades en característica positiva este proceso de resolución de Hironaka falla en los puntos infinitamente próximos, es decir en los que se estabiliza la función de Hilbert-Samuel. Este problema - nosotros lo superaremos mediante el contacto maximal.

El trabajo de Hironaka completa entonces el problema en el caso de ser X una variedad algebraica sobre un cuerpo de característica cero.

Sobrevive aún el problema de desingularización en los casos en que X no es una variedad algebraica sobre un cuerpo de característica cero, veremos brevemente qué hay hecho por el momento según las distintas posibilidades:

X variedad algebraica sobre un cuerpo de característica positiva

Hay multitud de resultados parciales entre los que -/ destacaremos los siguientes:

- 1) Caso de dim $X = 1$.- El problema está totalmente - resuelto por multitud de matemáticos clásicos como Noether y Walker.
- 2) Caso de dim $X = 2$.- Abhyankar en su Tesis de Harvard de 1954 demuestra el teorema de resolución su- mergida. Hironaka demuestra también este teorema - aplicando métodos derivados de una generalización del polígono de Newton.
- 3) Caso de dim $X = 3$.- El resultado más importante - es el de Abhyankar que prueba el teorema de resolu- ción para X variedad algebraica de dimensión 3 so- bre un cuerpo de característica distinta de 2, 3 y 5.
- 4) Caso general.- Ekloff, Hironaka y Giraud han con- seguido algunos resultados parciales. Analizaremos el trabajo de Ekloff.

Trabajo de Ekloff.- Ekloff utilizando los teoremas de Hiro- naka y las técnicas de ultraproductos consigue demostrar un teorema que permite resolver las singularidades de las va- riedades algebraicas sobre cuerpos de característica prima. La resolución es para todos los primos excepto un número fi- nito que depende de ciertos parámetros numéricos de las va- riedades.

Analizaremos brevemente este trabajo.

Definición 2.- Una relación finita y en particular una re- lación de rango n sobre un conjunto A es un subconjunto de A^n .

Definición 3.- Un sistema relacional es una sucesión $\mathcal{A} = \langle A, R_1, \dots, R_p, \dots \rangle$ donde A es un conjunto no vacío y cada R_p es una relación finita sobre A .

Nota 4.- La importancia capital de los sistemas relacionales estriba en que las estructuras matemáticas se pueden considerar como sistemas relacionales y así se manejarán conjuntos con relaciones en vez de conjuntos con operaciones.

Definición 5.- Sea $\{\mathcal{A}_i \mid i \in I\}$ una familia de sistemas relacionales, $\mathcal{A}_i = \langle A_i, R_1^i, \dots, R_p^i, \dots \rangle$. Sea D un ultrafiltro definido en el conjunto I .

Se llamará ultraproducto de la familia $\{\mathcal{A}_i \mid i \in I\}$ relativo al ultrafiltro D al siguiente sistema relacional

$\prod_{i \in I} \mathcal{A}_i / D = \langle \prod_{i \in I} A_i / D, R_1, \dots, R_p, \dots \rangle$ donde el dominio $\prod_{i \in I} A_i / D$ es el conjunto de las clases de equivalencia f / D ,

del producto cartesiano

$$\prod_{i \in I} A_i = \{ f \mid f : I \rightarrow \bigcup_{i \in I} A_i, f(i) \in A_i \}$$

para la relación de equivalencia \equiv_D definida

$$f \equiv_D g \leftrightarrow \{ i \mid f(i) = g(i) \} \in D$$

y la relación R_p se define por la condición siguiente

$$\langle f_1 / D, \dots, f_n / D \rangle \in R_p \leftrightarrow \{ i \mid \langle f_1(i), \dots, f_n(i) \rangle \in R_p^i \} \in D$$

Aplicando la teoría de los ultraproductos a la estructura de cuerpo se obtiene:

Teorema 6.- Sea $P = \{p \mid p \text{ n}^\circ \text{ primo}\}$, \mathcal{A}_p un cuerpo de característica p para cada primo p y D ultrafiltro no principal en el conjunto de todos los primos. Entonces con estos datos se verifica que $\mathcal{A} = \prod_{p \in P} \mathcal{A}_p / D$ es un cuerpo de característica cero.

Gracias a este teorema Ekloff resuelve las singularidades en característica prima basandose en el teorema de Hironaka de resolución en característica cero.

Pasaremos ya al estudio de la resolución de singularidades de variedades algebroides definidas sobre un cuerpo de característica cualquiera.

Resolución de singularidades de variedades algebroides

En este caso el problema de resolución de singularidades, se puede enfocar de la manera siguiente:

Sea X una variedad algebroides, entonces resolver las singularidades de X será lo mismo que conseguir que disminuya su función de Hilbert-Samuel.

X será una variedad algebroides sumergida en K^n cuando $X = \text{Spec}(K[[z_1, \dots, z_n]]/I)$ con I radical y K un cuerpo algebraicamente cerrado de característica cualquiera.

De lo dicho anteriormente se desprende que la resolución de singularidades de X nos vendrá dada al demostrar el siguiente teorema.

Teorema 7.- Sea X una variedad algebroides. Existe un número finito de transformaciones monoidales $\pi^{(i)} : X^{(i)} \rightarrow X^{(i-1)}$ con $X^{(0)} = X$, $0 < i \leq p$ tal que $H_{X^{(p)}} < H_X$ donde H_X es la función de Hilbert-Samuel de X .

Analizaremos brevemente los pasos que hemos seguido para la demostración de este teorema. En primer lugar daremos otra definición de variedad algebroides.

Definición 8.- Llamamos variedad algebroides intrínseca o simplemente variedad algebroides a $\text{Spec}(\hat{A})$ siendo \hat{A} un anillo local noetheriano, completo, equicaracterístico y reducido, de dimensión de inmersión finita y cuerpo de coefi

cientes algebraicamente cerrado.

Aplicando el teorema de estructura de los anillos locales completos se demuestra que esta definición coincide con la definición de variedad algebroide dada anteriormente.

Definiremos ahora los transformados monoïdales y cuadráticos del anillo de coordenadas de una variedad algebroide.

Definición 9.- Sean \square el anillo de coordenadas de una variedad algebroide, M el ideal maximal de \square y P un ideal primo de \square . Diremos que una \square -álgebra local \square' es un transformado monoïdal de \square con centro P si y solo si $\square' = A_N$ donde:

i) $A = \square [P/f_1]$ con $f_1 \in P$

ii) N es un ideal primo de A que contiene a MA .

Si $P = M$ diremos que \square' es un transformado cuadrático de \square .

Se verifica que si \square es el anillo de coordenadas de una variedad algebroide y \square' un transformado monoïdal de \square , entonces \square' no es el anillo de coordenadas de una variedad algebroide pues no es completo. Debido a ello daremos la definición siguiente:

Definición 10.- Sean \square el anillo de una variedad algebroide, \square' un transformado monoïdal de \square con centro P . Llamaremos transformado monoïdal formal de \square con centro P a $\hat{\square}'$ -complección de \square' respecto de la topología M' -ádica con M' ideal maximal de \square' . En el caso de que $P = M$ con M ideal maximal de \square diremos que $\hat{\square}'$ es un transformado cuadrático formal de \square .

Se verifica que si \square es el anillo de coordenadas de una variedad algebroide y $\hat{\square}$ un transformado monoidal formal de \square , entonces $\hat{\square}$ es también el anillo de una variedad algebroide.

Después de estos preliminares pasamos a un aspecto particular del problema que nos ocupa: la resolución de singularidades en las hipersuperficies algebroides. Exponemos brevemente algunos detalles.

Sea $R = K[[Z, W_1, \dots, W_{n-1}]]$,
 $H = \text{Spec}(R/(f(Z, W_1, \dots, W_{n-1})))$ una hipersuperficie algebroide, $f(Z, W_1, \dots, W_{n-1}) \in R$.

El teorema preparatorio de Weierstrass permite escribir la ecuación de la hipersuperficie H en la forma

$$f(Z, W_1, \dots, W_{n-1}) = Z^v + \varphi_1(W_1, \dots, W_{n-1})Z^{v-1} + \dots + \varphi_{v-1}(W_1, \dots, W_{n-1})Z + \varphi_v(W_1, \dots, W_{n-1}) = 0,$$

$$\varphi_i(W_1, \dots, W_{n-1}) \in K[[W_1, \dots, W_{n-1}]], \quad v \geq i.$$

Si la característica del cuerpo K no divide a v siendo v la multiplicidad de la hipersuperficie entonces con la transformación

$$Z = Z' + \frac{\varphi_1(W_1, \dots, W_{n-1})}{v}$$

se puede considerar que la ecuación de la hipersuperficie es de la forma

$$f(Z', W_1, \dots, W_{n-1}) = Z'^v + \varphi'_2(W_1, \dots, W_{n-1})Z'^{v-2} + \dots + \varphi'_{v-1}(W_1, \dots, W_{n-1})Z' + \varphi'_v(W_1, \dots, W_{n-1}) = 0$$

Así en el estudio que hacemos de hipersuperficies distinguimos dos casos:

1° caso. - La característica del cuerpo K no es divisor de la multiplicidad de la hipersuperficie.

2° caso.- La característica del cuerpo K es divisor de la multiplicidad de la hipersuperficie.

En el primer caso podemos suponer que en la ecuación de H no existe término en Z^{v-1} . Después de estudiar con todo detalle la variación del diagrama de Newton de una hipersuperficie mediante una transformación cuadrática formal demostramos los dos teoremas de desingularización en el caso de hipersuperficies.

Teorema 11.- Sea $f(Z, W_1, \dots, W_{n-1}) = 0$ la ecuación de la hipersuperficie H cuya forma inicial no es la potencia v -ésima de una forma lineal. Se verifica entonces que la multiplicidad de H decrece mediante una transformación cuadrática formal.

Teorema 12.- Sea $f(Z, W_1, \dots, W_{n-1}) = 0$ la ecuación de la hipersuperficie H cuya forma inicial $f_v(Z, W_1, \dots, W_{n-1})$ sea la potencia v -ésima de una forma lineal. Se verifica entonces que la multiplicidad de la hipersuperficie H decrece en un número finito de transformaciones cuadráticas formales.

La demostración de estos resultados es muy laboriosa. En la demostración de cada uno de ellos tienen que distinguirse los siguientes posibles casos, atendiendo a la característica del cuerpo base. 1° caso: la característica de K no es divisor de v siendo v la multiplicidad de la hipersuperficie. 2° caso.-la característica de K es divisor de v , en esta caso tenemos a su vez que distinguir dos subcasos: 1°) $v = \ell \cdot p^m$, $\ell > 1$, $m \geq 1$, $p = x(K)$, 2°) $v = p^m$; $m \geq 1$.

Una vez resuelto el problema, en el caso de las hipersuperficies algebroides, pasamos a considerarlo en el -

caso general de una variedad algebroide. Para ello, en nuestro orden de ideas, hemos necesitado utilizar el concepto de contacto maximal.

Sea X una variedad algebroide de anillo coordenadas \square . Designamos

$$H_X = H_{\square}$$

siendo H_{\square} la función de Hilbert-Samuel del anillo local \square . El teorema de Bennett afirma en el caso de ser K algebraicamente cerrado de característica cero que la función de Hilbert-Samuel no crece mediante una transformación monoidal normalmente plana. Hemos generalizado el teorema de Bennett a nuestro caso de ser K algebraicamente cerrado de característica arbitraria. De acuerdo con el teorema de Bennett, H_X es no creciente, por tanto ó se mantiene constante ó decrece. Si decrece, el problema queda resuelto. En el caso de que H_X se mantenga constante, para resolverlo utilizaremos el contacto maximal pues sabemos que si X posee contacto maximal la función de Hilbert-Samuel baja en un número finito de transformaciones monoidales. El problema, en nuestro caso de ser K de característica arbitraria es que el contacto maximal no existe siempre.

Analizaremos las condiciones de existencia de dicho contacto maximal.

Si consideramos K un cuerpo algebraicamente cerrado de característica p y H una hipersuperficie algebroide sumergida en K^n de multiplicidad v , con $v = p^m$, $m \geq 1$. En estas condiciones no existe siempre contacto maximal, es decir se puede encontrar una hipersuperficie H de manera que para cualquier hipersuperficie regular \bar{W} , se verifica

que \bar{W} no tiene contacto maximal con H.

Mostramos sin embargo un teorema que nos da la existencia del contacto maximal para hipersuperficies algebroides en el caso de que $v \neq p^m$, $m \geq 1$.

En el caso $v = p^m$, hemos obtenido los siguientes resultados:

Indicamos previamente las notaciones que utilizamos.

Sea H una hipersuperficie sumergida en K^n de ecuación $f(Z, W_1, \dots, W_{n-1}) = 0$. Sea W la hipersuperficie algebroides regular sumergida en K^n , de ecuación $Z = 0$. Si h es una función de prueba de W, es decir una aplicación

$$K[[W_1, \dots, W_{n-1}]] \rightarrow K[[t]] \text{ con}$$

$$h(W_1) = W_1(t), \dots, h(W_{n-1}) = W_{n-1}(t) \text{ y}$$

$$W_1(0) = \dots = W_{n-1}(0) = 0 \text{ y } K[[Z, t]] \text{ es el anillo de series}$$

de potencias formales en las indeterminadas Z, t sobre el cuerpo K, entonces llamaremos H_h a la hipersuperficie algebroides sumergida en K^2 de ecuación

$$f(Z, W_1(t), \dots, W_{n-1}(t)) = 0.$$

Hemos demostrado las siguientes proposiciones que nos dicen en qué condiciones existe contacto maximal cuando $v = p^m$.

Proposición 13.— Sean K un cuerpo de característica p y H una hipersuperficie algebroides sumergida en K^n de multiplicidad v. La hipersuperficie H no tiene contacto maximal si y solo si se verifican las dos condiciones siguientes:

i) $v = p^m$

ii) Existe h una función de prueba de W, siendo W la hipersuperficie de ecuación $Z = 0$, tal que el primer segmento del polígono de Newton de H_h tiene -

solo dos puntos y en los ejes, puntos de coordenadas $(0, v)$ y $(\ell v, 0)$, $\ell \geq 1$.

Proposición 14.- Sea K un cuerpo de característica p . Sea H una hipersuperficie algebroide sumergida en K^n de multiplicidad v , con

i) $v = p^m$

ii) el primer segmento del polígono de Newton de H tiene solo dos puntos y en los ejes. Entonces la hipersuperficie algebroide H no tiene contacto maximal.

Con esto, podemos ya afrontar el resultado general de la resolución de singularidades en variedades algebroides. Para ello estudiamos el contacto maximal para variedades algebroides en general. El método que seguiremos será considerar la variedad como una intersección tangencial adecuada de hipersuperficies o lo que es lo mismo trabajar con bases standard normalizadas. Sabemos que si X es una variedad algebroide sumergida en K^n , f_1, \dots, f_p una base standard normalizada de la variedad, H_1, \dots, H_p las hipersuperficies algebroides de ecuaciones $f_1 = 0, \dots, f_p = 0$.

Entonces si las hipersuperficies H_1, \dots, H_p tienen contacto maximal, es posible encontrar una variedad algebroide regular que tenga contacto maximal con la variedad X . Luego el problema del contacto maximal en variedades algebroides y por lo tanto el de resolución de singularidades queda resuelto con el teorema siguiente que hemos demostrado.

Teorema 15.- Sea X una variedad algebroide y $\{f_1, \dots, f_p\}$ una base standard normalizada de la variedad. Supongamos -

que las hipersuperficies H_i de ecuaciones $f_i = 0$, $-/$
 $1 \leq i \leq t$, $t < p$ tienen contacto maximal y sin embargo las
hipersuperficies de ecuaciones $f_i = 0$, $t + 1 \leq i \leq p$ no lo
tienen. En estas condiciones se verifica que existe un nú-
mero finito de transformaciones cuadráticas formales que -
transforman la base $\{f_1, \dots, f_p\}$ en $\{f'_1, \dots, f'_p\}$ de manera
que todas la hipersuperficies H'_i de ecuación $f'_i = 0$, $-/$
 $i = 1 \dots p$ tienen contacto maximal.

ON SOME PROJECTION METHODS FOR SOLVING NONLINEAR
OPERATOR EQUATIONS WITH A NONDIFFERENTIABLE TERM

IOANNIS K. ARGYROS
Cameron University
Department of Mathematics
Lawton, OK 73505-6377, U. S. A.

Abstract. We provide sufficient conditions for the convergence of a certain modified Newton-Kantorovich method to the solution of an equation with a nondifferentiable term.

Keywords and phrases. Banach space, Newton-Kantorovich method, majorant method.

(1980) A. M. S. Classification codes: 49D15, 47H17, 65J15, 65B05.

I. Introduction. Consider the fixed point problem

$$T(x) = x \text{ with } T(x) = F(x) + G(x) \quad (1)$$

where F, G are nonlinear operators defined on some convex subset D of a Banach space E with values in a Banach space \hat{E} . We assume that F is Frechet-differentiable on D , whereas G is not. Let $x_0 \in D$ and choose $R > 0$ such that the closed ball with center x_0 radius R , denoted by $\bar{U}(x_0, R)$ is included in D . Zabrejko-Nguen in [10] and others [9], [11] have proposed the modified Newton-Kantorovich iteration

$$z_{n+1} = z_n - (F'(z_n) - I)^{-1} (F(z_n) + G(z_n) - z_n), \quad z_0 = x_0, \quad n \geq 0, \quad (2)$$

for approximating a fixed point x^* of equation (1).

The above authors showed that under certain conditions, iteration (2) generates a sequence which converges to x^* . for $G = 0$, iteration (2) reduces to the classical Newton-Kantorovich method which has been studied

extensively by several authors [1] - [7].

However the iterates $\{z_n\}, n \geq 0$ can rarely be computed in infinite-dimensional spaces, since it may be difficult or even impossible to compute the inverses of the linear operators $F'(z_n) - I, n \geq 0$.

In this paper we will make practical use of iteration (2), by considering the iteration

$$x_{n+1} = x_n - (I - PF'(x_n))^{-1} (x_n - T(x_n)), n = 0, 1, 2, \dots (3)$$

where P is a projection operator ($P^2 = P$) on D .

Let us assume that the inverse of the operator $I - PF'(x_0)$ exists and

$$\|(I - PF'(x_0))^{-1} [PF'(x_1) - PF'(x_2)]\| \leq K_1(r) \|x_1 - x_2\|, (4)$$

$$\|(I - PF'(x_0))^{-1} [(QF(x_1) + G(x_1)) - (QF(x_2) + G(x_2))]\| \leq K_2(r) \|x_1 - x_2\| (5)$$

for all $x_1, x_2 \in \bar{U}(x_0, r) \subset \bar{U}(x_0, R)$ with $Q = I - P$, where $K_1(r)$ and $K_2(r)$ are nonnegative, nondecreasing functions on $[0, R]$. We note that for $P = I$ the conditions (4)-(5) reduce to the Zabrejko-Nguen conditions given in [9], [10].

It is easy to see that the solution of iteration (3) reduces to solving certain operator equations in the space E_P . If moreover E_P is a finite dimensional space of dimension N , we obtain a system of linear algebraic equations of order at most N .

We will provide sufficient conditions for the convergence of iteration (3) to x^* as well as error bounds on the distances $\|x_{n+1} - x_n\|$ and $\|x_n - x^*\|, n \geq 0$.

Finally we illustrate our results by considering a nondifferentiable nonlinear integral equation.

II. Convergence results.

We will need to introduce the constant

$$a = \|(I - PF'(x_0))^{-1} (x_0 - T(x_0))\|$$

and the functions

$$\omega(r) = \int_0^r K_1(t) dt, \quad \varphi(r) = a + \int_0^r w(t) dt - r,$$

$$\psi(r) = \int_0^r K_2(t) dt, \quad \chi(r) = \varphi(r) + \psi(r).$$

We can now prove the main theorem.

Theorem 1. Suppose that the function $\chi(r)$ has a unique zero s^* in the interval $[0, R]$ and $\chi(R) \leq 0$.

Then

(a) the equation (1) has a fixed point $x^* \in \bar{U}(x_0, s^*)$,

which is unique in $\bar{U}(x_0, R)$;

(b) the iterates generated by (3) are well defined, remain in $\bar{U}(x_0, s^*)$ for all $n \geq 0$ and satisfy

$$\|x_{n+1} - x_n\| \leq s_{n+1} - s_n, \quad n \geq 0 \quad (6)$$

and

$$\|x_n - x^*\| \leq s^* - s_n, \quad n \geq 0 \quad (7)$$

where the sequence $\{s_n\}$, $n \geq 0$ given by

$$s_{n+1} = s_n - \frac{x(s_n)}{\varphi'(s_n)} = s_n + u(s_n), \quad \text{with } u(r) = -\frac{\chi(r)}{\varphi'(r)}, \quad s_0 = 0, \quad n \geq 0$$

is monotonically increasing and converges to s^* .

Proof. The function $\chi(r)$ is positive on $[0, s^*]$, since s^* is the unique zero of $\chi(r)$. Exactly as in proposition 3 in [10 p. 677] we can show that the function $\varphi'(r)$ is negative on $[0, s^*]$ and that the sequence $\{s_n\}$, $n \geq 0$ is monotonically increasing and converges to s^* .

We will only show (6), since (7) will follow then immediately. We must show that the iterates $\{x_n\}$, $n \geq 0$ belong to $\bar{U}(x_0, s_n) \subset \bar{U}(x_0, s^*)$ and that the inverses $I - P\varphi'(x_n)$, $n \geq 0$ exist. For $n = 0$, (6) becomes

$$a = \|x_1 - x_0\| = s_1 - s_0 = a.$$

Hence, (6) is true for $n = 0$. Suppose the (6) is true for $N < k$;

then

$$\|x_k - x_0\| \leq \sum_{j=1}^k \|x_j - x_{j-1}\| \leq \sum_{j=1}^k (s_j - s_{j-1}) = s_k.$$

but by (4) and the result on p. 676 in [10]

$$\|(I - PF'(x_0))^{-1} (PF'(x_k) - PF'(x_0))\| \leq w(s_k) < w(s^*) = \varphi'(s^*) + 1 \leq 1.$$

By the Banach lemma on invertible operators $(I - PF'(x_k))^{-1}$ exists and

$$\|(I - PF'(x_k))^{-1} (I - PF'(x_0))\| \leq \frac{1}{\varphi'(s_k)}. \quad (8)$$

Using the identity

$$\begin{aligned} x_{k+1} - x_k &= [(I - PF'(x_k))^{-1} (I - PF'(x_0))] \{(I - PF'(x_0))^{-1} \\ &\quad [(PF(x_k) - PF(x_{k-1}) - PF'(x_{k-1})(x_k - x_{k-1})) \\ &\quad + ((G(x_k) + QF(x_k)) - (G(x_{k-1}) + QF(x_{k-1})))]\}, \end{aligned} \quad (9)$$

(4), (5), and (8), we get

$$\begin{aligned} &\|(I - PF'(x_0))^{-1} [PF(x_k) - PF(x_{k-1}) - PF'(x_{k-1})(x_k - x_{k-1})]\| \\ &\leq \int_0^1 \|(I - PF'(x_0))^{-1} (PF'((1-t)x_{k-1} + tx_k) - PF'(x_{k-1}))\| \|x_k - x_{k-1}\| dt \\ &\leq \int_0^1 (w((1-t)s_{k-1} + ts_k) - w(s_{k-1})) (s_k - s_{k-1}) dt \\ &= \int_{s_{k-1}}^{s_k} w(t) dt - w(s_{k-1})(s_k - s_{k-1}) \end{aligned}$$

and

$$\begin{aligned} &\|(I - PF'(x_0))^{-1} [(G(x_k) + QF(x_k)) - (G(x_{k-1}) + QF(x_{k-1}))]\| \\ &\leq \int_{s_{k-1}}^{s_k} K_2(t) dt = \psi(s_k) - \psi(s_{k-1}). \end{aligned}$$

With these majorizations, (9) becomes

$$\|x_{k+1} - x_k\| \leq \frac{\varphi(s_k) - \varphi(s_{k-1}) - \varphi'(s_{k-1})(s_k - s_{k-1}) + \psi(s_k) - \psi(s_{k-1})}{\varphi'(s_k)} = s_{k+1} - s_k. \quad (10)$$

Hence (6) is true for $n = k$. Since, the sequence $\{s_n\}$, $n \geq 0$

majorizes the sequence $\{x_n\}$, $n \geq 0$, there exists $x^* \in \bar{U}(x_0, s^*)$ such

that $x^* = \lim_{n \rightarrow \infty} x_n$. By taking the limit in (3) we get $x^* = T(x^*)$.

To show uniqueness we consider the sequences

$$\begin{aligned} y_{n+1} &= y_n - (I - PF'(x_0))^{-1} (y_n - T(y_n)), \quad n \geq 0, \\ v_{n+1} &= v_n - (I - PF'(v_0))^{-1} (v_n - t(v_n)), \quad n \geq 0, \quad v_0 \in \bar{U}(x_0, R), \\ q_{n+1} &= d(q_n), \quad n \geq 0, \quad q_0 = 0, \quad d(r) = r + \chi(r) \end{aligned}$$

and

$$p_{n+1} = d(p_n), \quad n \geq 0, \quad p_0 = R. \quad (11)$$

Then it is simple calculus to show that the iteration $\{q_n\}$, $n \geq 0$ is monotonically increasing and converges to s^* , whereas the sequence $\{p_n\}$, $n \geq 0$ is monotonically decreasing and converges to s^* also.

Exactly as we derived (10), we get

$$\|y_{n+1} - y_n\| \leq q_{n+1} - q_n, \quad n \geq 0 \quad (12)$$

and

$$\|y_n - v_n\| \leq p_n - q_n, \quad n \geq 0. \quad (13)$$

From (12) we get $\lim_{n \rightarrow \infty} y_n = x^* \in \bar{U}(x_0, s^*)$. If for v_0 we choose the second solution $x_1^* \in \bar{U}(x_0, s^*)$ of equation (1), we get, by (13) that $\|x^* - x_1^*\| \leq p_n - q_n$ and hence $x^* = x_1^*$.

That completes the proof of the theorem. For completion we will now obtain some further error bounds. Let $r_n = \|x_n - x_0\|$, $K_n(r) = K_1(r_n + r)$ and $E_n(r) = K_2(r_n + r)$ for $r \in [0, R - r_n]$ and set $a_n = \|x_{n+1} - x_n\|$, $b_n = (1 - w(r_n))^{-1}$. Without loss of generality, we may assume that $a_n > 0$. Then by following exactly the same steps as the proof of theorem 2 in [9 p. 989] and theorem 2 in [10, p.680] we can easily prove the theorem.

Theorem 2. Suppose that the hypotheses of theorem 1 are true.

Then

(a) the equation

$$r = a_n + b_n \int_0^r \{(r-t) K_n(t) + E_n(t)\} dt$$

has a unique positive zero s_n^* in the interval $[0, R - r_n]$ and

$$\|x_n - x^*\| \leq s_n^*, \quad n \geq 0, \quad \text{with } s_0^* = s^*.$$

(b) Moreover, the following estimates are true:

$$\|x_{n+1} - x_n\| \leq \Delta^{(n)}(a) = s_{n+1} - s_n, \quad n \geq 0$$

and

$$\|x_n - x^*\| \leq w(\Delta^{(n)}(a)) = s - s_n, \quad n \geq 0$$

where

$$\Delta^{(0)}(r) = r, \quad \Delta^{(n+1)}(r) = \Delta(\Delta^{(n)}(r)), \quad n \geq 0$$

and

$$w(r) = \sum_{n=0}^{\infty} \Delta^{(n)}(r).$$

(c) Furthermore, the following estimates are true:

$$\begin{aligned} \|x_n - x^*\| &\leq s_n^* \\ &\leq (s^* - s_n) a_n / \Delta s_n, \quad n \geq 0 \\ &\leq (s^* - s_n) a_{n-1} / \Delta s_{n-1}, \quad n \geq 0, \\ &\leq s^* - s_n, \quad n \geq 0. \end{aligned}$$

We now complete this paper with an application.

III. Applications.

Consider the integral equation

$$x(t) = \int_0^1 K(t, s, x(s)) ds,$$

where the kernel $K(t, s, x(s))$ is nondifferentiable on some convex subset

$D \subset E = C[0,1]$. We set $T(x) = \int_0^1 K(t, s, x(s)) ds$ and $F(x) = \int_0^1 \bar{K}(t, s, x(s)) ds$,

where $\bar{K}(t, s, x(s))$ is differentiable on D . Then

$$PF'(x) = \int_0^1 \bar{K}_x(t, s, x(s)) ds,$$

where

$$\bar{K}(t, s, x(s)) = \sum_{i=1}^{\infty} A_i(t) B_i(s, x(s))$$

is a degenerate kernel approximating the functions $\bar{K}(t, s, x)$, e.g., a portion of the Taylor or Fourier series for the function $\bar{K}(t, s, x)$ if we consider it as a function of t . The modified Newton-Kantorovich iteration (3) can now be written as

$$x_{n+1}(t) = \int_0^1 K(t, s, x_n(s)) ds - \int_0^1 \bar{K}'_x(t, s, x_n(s)) x_n(s) ds + \int_0^1 \bar{K}'_x(t, s, x_n(s)) x_{n+1}(s) ds. \quad (14)$$

$$\text{Let } f_n(t) = \int_0^1 K(t, s, x_n(s)) ds - \int_0^1 \bar{K}'_x(t, s, x_n(s)) x_n(s) ds,$$

then iteration (14) can be written as

$$x_{n+1}(t) = f_n(t) + \sum_{i=1}^m A_i(t) \int_0^1 B'_i(s, x_n(s)) x_{n+1}(s) ds,$$

which can be solved to give a system of linear algebraic equations

$$\int_0^1 B'_i(s, x_n(s)) x_{n+1}(s) ds - \sum_{i=1}^m \int_0^1 B'_i(s, x_n(s)) A_j(s) ds \int_0^1 B'_i(s, x_n(s)) x_{n+1}(s) ds = \int_0^1 B'_i(s, x_n(s)) f_n(s) ds.$$

Denote by $D(x_n)$ the determinant of the above system and assume $D(x_n) \neq 0, n \geq 0$. Then,

$$\int_0^1 B'_i(s, x_n(s)) x_{n+1}(s) ds = \frac{1}{D(x_n)} \int_0^1 \sum_{k=1}^n D_{ki}(x_n) B'_k(s, x_n(s)) f_n(s) ds$$

and

$$x_{n+1}(t) = f_n(t) + \int_0^1 \sum_{i=1}^m \sum_{k=1}^m \frac{A_i(t) D_{ki}(x_n) B'_k(s, x_n(s))}{D(x_n)} f_n(s) ds$$

where $D_{ki}(x_n)$ is the cofactor of the element in the i -th row and k -th column of the determinant $D(x_n)$.

Suppose now that the operators $\bar{K}_x(t, s, x), Q(t, s, x), G(t, s, x)$ and $L(t, s, x)$, where

$Q(t, s, x) = \bar{K}(t, s, x) - \bar{K}(t, s, x), G(t, s, x) = K(t, s, x) - \bar{K}(t, s, x)$
 and $L(t, s, x) = \frac{1}{D(x)} \sum_{i=1}^m \sum_{k=1}^m A_i(t) D_{ki}(x) B_k'(s, x)$, satisfy the conditions

$$|\bar{K}'_x(t, s, x) - \bar{K}'_x(t, s, y)| \leq c_1(t, s) |x - y|$$

$$|(Q(t, s, x) - Q(t, s, y)) + (G(t, s, x) - G(t, s, y))| \leq c_2(t, s) |x - y|,$$

and $|L(t, s; x)| \leq r(t, s)$ on D .

For simplicity set $K_1(r) = K_1$ and $K_2(r) = K_2$ for all $r = [0, R]$ in (4)-(5). Then the constants K_1 and K_2 can be computed as follows:

$$K_1 \leq w \sup_{t \in [0,1]} \int_0^1 c_1(y, s) ds, \quad K_2 \leq w \sup_{t \in [0,1]} \int_0^1 c_2(t, s) ds$$

where

$$w = 1 + \sup_{t \in [0,1]} \int_0^1 r(t, s) ds.$$

REFERENCES

- [1] I.K Argyros. On Newton's method and nondiscrete Matheemtical induction. Bull. Austral. Math. Soc. Vol. 38, (1988), pp. 131-140.
- [2] M. Balazs and G. Goldner. On the method of the cord and on a modification of it for the solution of nonlinear operator equations. Stud. Cerc. Mat. 20, (1968), pp. 981-990.
- [3] W.B. Gragg and R.A. Tapia. Optimal error bounds for the Newton-Kantorovich Theorem. S.I.A.M. J. Numer. Anal. 1 (1974), pp. 10-13.
- [4] F.M. Migovich. On the convergence of projection-iterative methods for solving nonlinear operator equations. Dopov. Akad. Nauk. Ukr. RSR, ser. A, No. 1, (1970), pp. 20-23.
- [5] F.A. Potra and V. Pták. Sharp error bounds for Newton's process. Numer. Math. 34, (1980), pp. 63-72.
- [6] W.C. Rheinboldt. A unified convergence theory for a class of iterative processes. S.I.A.M. J. Number. Anal. 5 (1968), pp. 42-63.
- [7] J.W. Schmidt. Unter Fehrschranke für regular-falsi-verfahren. Period. Math. Hung. 9, (1978), pp. 241-247.
- [8] T. Yamamoto. A method for finding sharp error bounds for Newton's method under the kantorovich assumptions. Numer. Math. 44, (1986), pp. 203-220.
- [9] _____. A note on a posteriori error bound of Zabrieiko and Nguen for Zincenko's iteration. Nuber. Funct. Anal. and Optimiz., 9, (9 and 10), (1987), pp. 987-994.
- [10] P.P. Zabrejko and D.F. Nguen. The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates. Numer. Funct. Anal. Optimiz. 9, (1987), pp. 671-684.
- [11] A.I. Zincenko. Some approximate methods of solving equations with nondifferentiable operators. (Ukrainian). Dopovidi Akad. Navk. Ukrain. RSR (1963), pp. 156-161.

INTEGRAL EQUATIONS FOR TWO-POINT
BOUNDARY-VALUE PROBLEMS.

by

Ioannis K. Argyros

and

Mansor M. Losta

IOANNIS K. ARGYROS
Cameron University
Department of Mathematics
Lawton, OK 73505-6377, U. S. A.

Abstract We present some collocation methods for the numerical solution of two-point boundary value problems as well as quadrature methods for the numerical solution of Fredholm integral equations.

(1980) A.M.S. classification codes: 65J15, 65H10, 65B05.

Key words and phrases: Collocation method, Quadrature method, Fredholm integral equation.

I Introduction In this paper we define and analyze numerical methods for the solution of the integral equations describing linear two-point boundary value problems. The integral equation approach is due to Nystrom, who approximates the integral equations by numerical quadratures. Integral equations appear in many important fields of science. For example they appear in astrophysics, kinetic theory of gases. Relevant work has been done in [7],[8] and [10].

II. Collocation method for the numerical solution of two-point value problems.

In this section we provide some collocation methods for the numerical solution of two-point boundary value problems involving operators with symmetric separable kernels. This leads to a system of algebraic equations whose solution is an approximation of the unknown function over the given domain.

Symmetric separable kernels appear frequently as Green's functions for the two-point boundary value problems for ordinary differential operators. More precisely, suppose that

$$L[u(t)] = \varphi(t) \quad (1.1)$$

which, together with derivatives of lower order than the order of the operator, constitute a linear ordinary differential operator where the solution $u(t)$ satisfies

$$u(a) = \alpha, \quad u(b) = \beta. \quad (1.2)$$

We assume that the boundary-value problem has a unique solution $u(t)$ which can be represented as

$$u(t) = \int_a^b G(t, \tau) \varphi(\tau) d\tau \quad (1.3)$$

for all functions $\varphi(t)$ from some class such as the continuous functions. Then $G(t, \tau)$ is said to be the Green's function for (1.1) corresponding to the given boundary conditions. Let us assume that $\varphi(t)$ is of the form

$$\varphi(t) = \lambda u(t) + q(t). \quad (1.4)$$

Then the boundary value problem (1.1) is equivalent to the Fredholm integral equation given by

$$u(t) - \lambda \int_a^b G(t, \tau) u(\tau) d\tau = f(t), \quad a \leq t \leq b \quad (1.5)$$

where

$$f(t) = \int_a^b G(t, \tau) q(\tau) d\tau, \quad a \leq t \leq b. \quad (1.6)$$

We replace (1.5) by

$$u(t) \left[1 - \lambda \int_a^b G(t, \tau) d\tau \right] - \lambda \int_a^b G(t, \tau) [u(\tau) - u(t)] d\tau = f(t). \quad (1.7)$$

Now, since $G(t, \tau)$ is known, the function $\int_a^b G(t, \tau) d\tau$ can be determined as a function of t , say $\psi(t)$. The integral

$$\int_a^b G(t, \tau) [u(\tau) - u(t)] d\tau \quad (1.8)$$

gives a better approximation than the integral in (1.5) when using a quadrature formula (see [1]).

III. Quadrature Methods for the Numerical Solution of Fredholm Integral Equations.

The method will be described for

$$u(t) - \lambda \int_a^b G(t, \tau) u(\tau) d\tau = f(t), \quad a \leq t \leq b \quad (2.1)$$

where λ is a regular value of the kernel, $f(t)$ is a continuous function and $u(t)$ is the unknown function. We will assume that G is sufficiently smooth. Many different quadrature rules can be used to approximate the integral in (2.1). Thus for example a composite trapezoid rule defined for $t_j = a + jh$, $h = (b-a)/n$ and $j \leq n$ as

$$\int_a^b u(t)dt \cong \frac{h}{2} \sum_{j=1}^n [u(t_{j-1}) + u(t_j)], \quad (2.2)$$

can be used. If we use a quadrature formula with weights ω_j and points t_j , $j = 1, 2, \dots, n$, to evaluate (1.8), then approximate values of the function $u(t_j)$ may be determined by solving the linear system

$$u(t_j)[1 - \psi(t_j)] - \sum_{j=1}^n \omega_j G(t_i, t_j)[u(t_j) - u(t_i)] = f(t_i), \quad (2.3)$$

$$i = 1, 2, \dots, n.$$

After obtaining the values at the grid points, several different interpolation schemes can be used to approximate the solution.

We now list some well-known results. Let X be the Banach space $C[a, b]$ together with the Chebyshev norm

$$\|u\| = \max \{|u(t_j)| : a \leq t_j \leq b\}.$$

We wish to approximate the solution of the integral equation (2.1), expressed in the operator form

$$(I + G)u = f, \quad (2.4)$$

where G is a compact linear operator in $BL(X)$ for which $(I + G)$ possess a bounded inverse $(I + G)^{-1}$ in X . Here $BL(X)$ defines the set of bounded linear operator from X into X with the norm

$$\|G\| = \sup\{\|Gx\| : x \in X \text{ and } \|x\| = 1\},$$

$BL(X)$ becomes a Banach space.

Theorem 1. (Banach Lemma)

Let $(X, \|\cdot\|)$ be a Banach space and $G \in BL(X)$. If $\|G\| < 1$, then operator $(I - G)$ is nonsingular,

$$(i) (I - G)^{-1} = \sum_{m=0}^{\infty} G^m$$

and

$$(ii) \|(I - G)^{-1}\| \leq (1 - \|G\|)^{-1}.$$

Proof. If we define $L_N = \sum_{m=0}^N G^m$, then $\{L_N\}_0^{\infty}$ can be shown to be a Cauchy sequence in the Banach space $BL(X)$. Therefore, for some $L \in BL(X)$ the sequence $\{L_N\}_0^{\infty}$ converges to L . Furthermore, since

$$L_N(I - G) = I - G^{N+1}$$

we find

$$L(I - G) = \lim_{N \rightarrow \infty} L_N(I - G) = I.$$

Therefore $L = (I - G)^{-1}$. Since

$$\|L_N\| \leq \sum_{m=0}^N \|G\|^m \leq (1 - \|G\|)^{-1}$$

then

$$\|L\| = \lim_{N \rightarrow \infty} \|L_N\| \leq (1 - \|G\|)^{-1}.$$

Theorem 2.

Let $(X, \|\cdot\|)$ be a Banach space. If $G, G^{-1}, L \in BL(X)$ then

for each positive $\epsilon_0 < \|G^{-1}L\|^{-1}$ and all $|\epsilon| \leq \epsilon_0$ the operator $G + \epsilon L$ is nonsingular. Furthermore

$$(1) \quad (G + \epsilon L)^{-1} = (I + \epsilon G^{-1}L)^{-1}G^{-1}$$

and

$$(2) \quad \|(G + \epsilon L)^{-1}\| \leq (1 - \epsilon \|G^{-1}L\|)^{-1} \|G^{-1}\|.$$

Proof. Since G is invertible we deduce

$$G + \epsilon L = G(I + \epsilon G^{-1}L).$$

Therefore (1),(2) follow since $I + \epsilon G^{-1}L$ is nonsingular by the Banach Lemma.

Theorem 3. (Neumann's theorem)

Let T, S and $T^{-1} \in BL(X)$. If $\|T^{-1}\| \|S - T\| < 1$. Then S^{-1} exists in $BL(X)$ and

$$\|S^{-1} - T^{-1}\| \leq \frac{\|T^{-1}\| \|T - S\|}{1 - \|T^{-1}\| \|S - T\|},$$

where $\|T\| = \sup \{ \|Tx\| : \|x\| = 1 \}$.

To approximate the integral (1.5) we let u_n be the approximate solution chosen from the finite dimensional subspaces X_n of X where u_n converges to u and $p_n u$ be the projection of u onto X_n . We then replace (1.5) by

$$(I + p_n G)u_n = p_n \psi. \quad (2.5)$$

Prenter [9] showed that, for all n sufficiently large (2.5) has a unique solution $u_n \in X_n$ for each $\psi \in X$. With a partition $\{t_i\}_{i=0}^n$ of $[a, b]$ the equation (2.5) is equivalent to the system of linear equations

$$(u_n + Gu_n)(t_i^n) = \psi(t_i^n), \quad i = 0, 1, \dots, n,$$

which we will solve with the pivot method. We list the following theorems

Theorem 4.

$$\| p_n G - G \| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

The proof can be found in [9].

Theorem 5.

Let $(I + p_n G)^{-1}$ exist for all $n \geq n_0$. Then $\{(I + p_n G)^{-1} : n \geq n_0\}$ is uniformly bounded.

Proof. We use Theorem 1 and 3.

Let $T = I + G$ and $S = (I + p_n G)^{-1}$. Then

$$\|(I - G)^{-1} - (I - p_n G)^{-1}\| \leq \frac{\|(I + G)^{-1}\| \|G - p_n G\|}{1 - \|(I + G)^{-1}\| \|G - p_n G\|}.$$

As $n \rightarrow \infty$, $\|(I + G)^{-1}\| \|G - p_n G\| \rightarrow 0$. Thus there is a constant

$$(Lw)(s) = \int_0^1 K(s,t)\psi(t)w(t)dt \quad (2.7)$$

and

$$\bar{B}(w,z)(s) = w(s)(Lz)(s) + z(s)(Lw)(s). \quad (2.8)$$

Denote by $\bar{B}'(w)$ the Frechet derivative of \bar{B} . Define the linear functional E on a subset D of $C[0,1]$ by

$$E(w) = \int_0^1 w(t)H(t)\psi(t)dt \quad (2.9)$$

where H is the solution of Equation (2.6).

$$\text{If } U = \int_0^1 H(t)\psi(t)dt \text{ and } R = \int_0^1 \psi(t)dt, \text{ then } U^2 - 2U + R = 0.$$

The proof can be found in Argyros [6]. We now introduce the two additional theorems proved in [3].

Theorem 6.

Let $H = H(\lambda_0)$ be any solution to equation (2.4) with $\lambda = \lambda_0$. If $\int_0^1 \psi(t)dt \neq 0$ and $\ker(L) = \{0\}$ then $\bar{B}'(H(\lambda_0))$ is an invertible linear operator on $C[0,1]$ with bounded inverse.

We consider the quadratic equation

$$x = y + B(x,x) \tag{2.10}$$

in a Banach space X , where $y \in X$ is fixed and B is a bounded bilinear operator on X . Equation (2.6) is a special case of equation (2.10) provided that $\lambda\bar{B} = B$, $H = x$, $y = 1$ and $X = C[0,1]$. We now state the following lemmas.

Lemma 1. Let L_1 and L_2 be bounded linear operators on a Banach space X where L_1 is invertible, and $\|L_1^{-1}\| \cdot \|L_2\| < 1$. Then $(L_1 + L_2)^{-1}$ exists and

$$(L_1 + L_2)^{-1} \leq \frac{\|L_1^{-1}\|}{1 - \|L_2\| \cdot \|L_1^{-1}\|}.$$

Lemma 2. Let $z \neq 0$ be fixed in X . Assume that the linear operator $B(z)$ is invertible, then $B(x)$ is also invertible for all $x \in U(z,r) = \{x \in X : \|x - z\| < r\}$, where $r \in (0, r_0)$ and

$$r_0 = [\|B\| \cdot \|B(z)^{-1}\|]^{-1}.$$

Proof . We have

$$\begin{aligned} \|B(x-z)\| \cdot \|B(z)^{-1}\| &\leq \|B\| \cdot \|x-z\| \cdot \|B(z)^{-1}\| \\ &\leq \|B\| \cdot \|B(z)^{-1}\| \cdot r \end{aligned}$$

for $r \in (0, r_0)$. The result now follows from Lemma 1 for $L_1 = B(z)$, $L_2 = B(x-z)$ and $x \in U(z, r)$.

Definition 1. Let $z \neq 0$ be fixed in X . Assume that the linear operator $B(z)$ is invertible. Define the operators P, T on $U(z, r)$ by

$$\begin{aligned} P(x) &= B(x, x) + y - x, \\ T(x) &= (B(x))^{-1}(x - y) \end{aligned}$$

and the real polynomials $f(r), g(r)$ on R by

$$f(r) = a'r^2 + b'r + c',$$

$$g(r) = ar^2 + br + c,$$

$$a' = \|B\| \cdot \|B(z)^{-1}\|$$

$$b' = -2\|B\| \cdot \|B(z)^{-1}\|$$

$$c' = 1 - \|B(z)^{-1}\| - \|B\| \cdot \|B(z)^{-1}\|^2 \cdot \|z-y\|$$

$$a = r_0^{-1}$$

$$b = \|B(z)^{-1}(I - B(z))\| - 1$$

$$c = \|B(z)^{-1}P(z)\|.$$

Theorem 7.

Let $z \in X$ be such that $B(z)$ is invertible and that the following are true:

- (a) $c' > 0$;
- (b) $b < 0$, $b^2 - 4ac > 0$, and
- (c) there exists $r > 0$ such that $f(r) > 0$ and $g(r) \leq 0$.

Then the iteration

$$x_{n+1} = B(x_n)^{-1}(x_n - y), \quad n = 0, 1, 2, \dots \quad (2.11)$$

is well defined and it converges to a unique solution x in $\bar{U}(z, r)$ of (2.10) for any $x_0 \in \bar{U}(z, r)$. The proof can be found in [8].

We now introduce the iteration

$$H_n = H_{n-1} + \bar{B}(H_{n-1})^{-1}(H_{n-1} - 1 - \bar{B}(H_{n-1}, H_{n-1})). \quad (2.12)$$

Convergence of the above iteration is a consequence of Theorems 6 and 7.

IV References

- 1.- Anselone, P.M. "Singularity subtraction in the numerical solution of integral equations," J. Austral. Math. Soc. 22: 408-418, 1981.
- 2.- Argyros, I.K. "Quadratic equation and application to Chandrasekhar's and related equation." Bull. Austral. Math. 32,2: 275-292, 1985.
- 3.- Argyros, I.K. "On the approximation of some nonlinear equations," Aequationes Mathematicae 32: 87-95, 1987.
- 4.- Argyros, I.K. "On a class of nonlinear integral equations arising in neutron transport," Aequationes Mathematicae 36: 99-111, 1988
- 5.- Argyros, I.K. "On a new iteration for solving Chandrasekhar's H-equation," Revista Proyecciones 15: 21-31, 1988.
- 6.- Argyros, I.K. "On the number of solutions of some integral equations arising in radiative transfer," Internat. J. Math. and Math. Sci, 12,2: 297-304, 1989.

- 7.- Atkinson, K.E. "The numerical solution of ferdholm integral equations of the second kind," SIAM J. Numer. Anal. 4: 337-348 1967.
- 8.- Crout, P.D. "An application of polynomial approximation to the solution of integral equations arising in physical problems," J. Math. and Phys. 19: 34-92, 1941.
- 9.- Prenter, P.M. "A collection method for the numerical solution of integral equation," SIAM J. Numer. Anal. 10,4: 570-581, 1973.
- 10.- Robert, W. "The stability of finite rank methods with applications to integral equations," SIAM J. Numer. Anal. 23: 118-134, 1986.

OPTIMIZACION DE FUNCIONES NO DERIVABLES

V. Novo

Departamento de Matemática Aplicada

ETSI. Industriales.

Univ. Nac. Educ. Dist.

28080-MADRID.

Abstract: Conditions both necessary and sufficient for extrema of non-derivable functions are stated using a new generalized derivative.

1. INTRODUCCION Y NOTACIONES

La investigación de técnicas matemáticas adecuadas al estudio de problemas de optimización de funciones no diferenciables en el sentido clásico, ha llevado, en los últimos años, al desarrollo de nuevas teorías de diferenciación generalizada, aplicables a este tipo de problemas. Rockafellar [12] trata el caso de las funciones convexas en un abierto de \mathbb{R}^n e introduce las definiciones de subgradiente y subdiferencial $\partial_R f$. Clarke [2] amplía la clase de funciones utilizada por Rockafellar y considera funciones reales localmente lipschitzianas en un abierto de \mathbb{R}^n , introduciendo la definición de gradiente generalizado $\partial_c f$ que extiende posteriormente a funciones entre espacios de dimensión finita [3] y a funcionales reales sobre un espacio de Banach [4]. Los principales resultados obtenidos pueden verse en [1], [2], [3], [4], [6], [8] y [12].

Con objeto de ampliar la clase de funciones utilizada por Clarke se introduce en [9] una derivada generalizada para funciones reales de variable real que denominaremos G-derivada. Esta derivada generalizada, a diferencia de la de Clarke, se reduce a la derivada clásica si f es derivable. En este trabajo, se dan las definiciones y los principales resultados sobre G-derivadas, así como sus aplicaciones al estudio de extremos de funciones reales de variable real. La utilización de G-derivadas nos permite (teorema 3.6) dar una demostración sencilla del teorema de Stepanoff [5].

En todo lo que sigue utilizaremos las notaciones siguientes. f es una función real de variable real definida en el intervalo I abierto o cerrado con la topología inducida.

$$F(x_0, x) = \frac{f(x) - f(x_0)}{x - x_0} ; l(f, x_0, x_n) = \lim_{n \rightarrow \infty} F(x_0, x_n)$$

en donde $x, x_0, x_n \in I; (x_n) \rightarrow 0$. $Conv(I), L(I), A(I), V(I), C(I)$ denotan respectivamente las clases de funciones convexas, localmente lipschitzianas, absolutamente continuas, de variación acotada y continuas en I . $co A$ denota la envoltura convexa de A .

2. DEFINICIONES Y PROPIEDADES.

Definición 2.1. f es fuertemente continua (f.c.) en c si existen un entorno reducido V de c y una constante $k > 0$ tales que

$$|F(c, x)| \leq k \text{ para todo } x \in V$$

Se denota por $FC(I)$ la clase de funciones f.c. en cada punto de I . Es claro que

$$L(I) \subset FC(I) \subset C(I)$$

sin embargo no existe ninguna relación de contenido entre $A(I)$ y $FC(I)$ ni entre $V(I)$ y $FC(I)$ siendo I cerrado. Además $(FC(I), +, \cdot)$ es una subálgebra densa en $C(I)$ y la composición de funciones f.c. es f.c.

Definición 2.2. Una sucesión $(x_n) \subset I$ convergente a x_0 , se llama sucesión de G-derivabilidad de f en x_0 , si existe y es finito $l(f, x_0, x_n)$. Se denota por $S(f, x_0)$ el conjunto de sucesiones de G-derivabilidad de f en x_0 .

Definición 2.3. f es G-derivable en x_0 si toda sucesión $(x_n) \rightarrow x_0$ contiene al menos una subsucesión $(x_k) \in S(f, x_0)$. La G-derivada de f en x_0 es el conjunto

$$\partial f(x_0) = \text{co} \{ l(f, x_0, x_n) ; (x_n) \in S(f, x_0) \}.$$

Nótese que si f es G-derivable en x_0 , entonces $\partial f(x_0)$ es la envoltura convexa de los números derivados de Dini. Análogamente se definen l^+, s^+, l^- y s^- y las G-derivadas laterales $\partial^+ f(x_0)$ y $\partial^- f(x_0)$ considerando respectivamente sucesiones convergentes a x_0 por la derecha o por la izquierda. A partir de esta definición se obtienen los resultados siguientes de los que únicamente se demuestra la regla de la cadena.

Proposición 2.4. Si f es G-derivable en x_0 , $\partial f(x_0)$ es un subconjunto convexo y compacto de \mathbb{R} .

Proposición 2.5. Si f es G^+ y G^- derivable en x_0 , entonces f es G-derivable en x_0 y se tiene que

$$\partial f(x_0) = \text{co} (\partial^- f(x_0) \cup \partial^+ f(x_0))$$

Teorema 2.6. f es f.c. en I si y solo si f es G-derivable en cada punto de I .

Como consecuencia inmediata se tiene que las funciones convexas y las localmente lipschitzianas en I son G -derivables en cada punto de I .

Teorema 2.7. Si f y g son $FC(I)$ y k es un número real, entonces se tiene que $kf, f+g, fg$ y f/g son G -derivables en cada $x_0 \in I$ y además se verifica que:

$$\partial(kf)(x_0) = k\partial f(x_0) \quad ; \quad \partial(f+g)(x_0) \subset \partial f(x_0) \cup \partial g(x_0)$$

$$\partial(fg)(x_0) \subset g(x_0)\partial f(x_0) + f(x_0)\partial g(x_0)$$

$$\partial(f/g)(x_0) \subset \frac{g(x_0)\partial f(x_0) - f(x_0)\partial g(x_0)}{[g(x_0)]^2} \quad \text{si } g(x_0) \neq 0$$

Teorema 2.8. Sean I y J intervalos abiertos de números reales, f de I en J y g de J en el conjunto de números reales. Si f es G -derivable en $x_0 \in I$ y g es G -derivable en $y_0 = f(x_0)$, entonces $g \circ f$ es G -derivable en x_0 y

$$\partial(g \circ f)(x_0) \subset \partial g(y_0)\partial f(x_0)$$

Demostración: Como la composición de funciones f.c. es f.c., por 2.6 $g \circ f$ es G -derivable en x_0 . Sea $(x_n) \in S(g \circ f, x_0)$ y para cada número natural n , $k_n = f(x_n) - f(x_0)$, $y_n = f(x_n)$ y $h_n = x_n - x_0$, se tiene que

$$\begin{aligned} \frac{g[f(x_n)] - g[f(x_0)]}{x_n - x_0} &= \frac{g(y_0 + k_n) - g(y_0)}{h_n} = \frac{g(y_0 + k_n) - g(y_0)}{k_n} \frac{k_n}{h_n} = \\ &= \frac{g(y_n) - g(y_0)}{y_n - y_0} \frac{f(x_n) - f(x_0)}{x_n - x_0} \end{aligned}$$

de donde $1(g \circ f, x_0, x_n) \in \partial g(y_0)\partial f(x_0)$ y en consecuencia

$$\partial(g \circ f)(x_0) \subset \partial g(y_0)\partial f(x_0)$$

por ser el último conjunto convexo.

El teorema del valor medio se puede extender a este contexto. Daremos previamente la condición necesaria de extremo relativo y las generalizaciones del teorema de Rolle y de la fórmula de Cauchy.

Teorema 2.9. Dados el intervalo abierto I de números reales y f f.c. en I , si $x_0 \in I$ es un extremo relativo de f , entonces

$$0 \in \partial f(x_0)$$

Demostración: Supongamos que x_0 es un mínimo relativo de f , existe $\delta > 0$ tal que $f(x) - f(x_0) \geq 0$ si $|x - x_0| < \delta$ y dada $(x_n) \in S^+(f, x_0)$ existe $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$, $0 < x_n - x_0 < \delta$ y en consecuencia $f(x_n) - f(x_0) \geq 0$. Por lo tanto $F(x_0, x_n) \geq 0$ y $l^+(f, x_0, x_n) \geq 0$, luego $\partial^+ f(x_0)$ contiene al menos un número no negativo. De igual forma se demuestra que $\partial^- f(x_0)$ contiene al menos un número no positivo y de 2.5 se deduce que $0 \in \partial f(x_0)$. La demostración es análoga si x_0 es un máximo relativo.

Teorema 2.10. Si f es f.c. en $[a, b]$ y $f(a) = f(b)$, entonces existe c de (a, b) tal que $0 \in \partial f(c)$.

Demostración: Como f es f.c. en $[a, b]$, es continua en el compacto $[a, b]$ por lo que alcanza su máximo y su mínimo absolutos en $[a, b]$. Aplicando 2.9 se tiene el resultado.

Teorema 2.11. Si f y g son f.c. en $[a, b]$, entonces existen $c \in (a, b)$, $A \in \partial f(c)$ y $B \in \partial g(c)$ tales que

$$B[f(b) - f(a)] = A[g(b) - g(a)]$$

Demostración: Sea h la función real definida en $[a, b]$ por

$$h(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)].$$

Aplicando el teorema anterior, por 2.7 se tiene el resultado.

Teorema 2.12. Si f es f.c. en $[a,b]$, entonces existen un punto c de (a,b) y $A \in \partial f(c)$ tales que $f(b) - f(a) = A(b-a)$.

Demostración: Basta considerar en el teorema anterior la función $g(x) = x$.

Teorema 2.13. Dada $f \in FC(a,b)$ y $(h,k) \subset (a,b)$, si para todo $x \in (h,k)$, $\partial f(x) \subset \mathbb{R}^+$ (resp. $\partial f(x) \subset \mathbb{R}^-$), entonces f es estrictamente creciente (resp. decreciente) en (h,k) .

Demostración: Supongamos que para todo $x \in (h,k)$, $\partial f(x) \subset \mathbb{R}^+$. Sean $x_1, x_2 \in (h,k)$, $x_1 < x_2$; aplicando a f el teorema 2.12 en $[x_1, x_2]$, existen $c \in (x_1, x_2)$ y $A \in \partial f(c)$ de tal forma que $f(x_2) - f(x_1) = A(x_2 - x_1) > 0$, con lo que f es estrictamente creciente en (h,k) . Si $\partial f(x) \subset \mathbb{R}^-$ el razonamiento es análogo.

3. RELACION CON OTRAS TEORIAS DE DERIVACION

Supondremos que f es fuertemente continua en I y pondremos por brevedad $F(x,y) = [f(y) - f(x)] / [y - x]$ en donde x e y son elementos de I con x distinto de y .

Proposición 3.1. La G -derivada de f en c se reduce a un punto si y solo si f es derivable en c .

Demostración: Si existe $f'(c)$ es claro que la G -derivada de f en c se reduce a $f'(c)$. Recíprocamente, si la G -derivada se reduce a un punto, entonces coinciden los números derivados de Dini y en consecuencia f es derivable en c .

Ejemplo 3.2. Sea $f(x) = x^2 \text{sen}[1/x]$ si $x \neq 0$; $f(0) = 0$. Existe la derivada en el sentido clásico de f en cero y vale cero con lo que por 3.1 la G -derivada de f en cero vale cero mientras que la derivada en el sentido de Clarke de f en cero es el intervalo

$[-1,1]$.

Definición 3.3. Se llama derivada superior (resp. inferior) de f a la función \bar{F} (resp. \underline{f}') de I en \mathbb{R} definida por:

$$\bar{F}(x) = \max \partial f(x) \quad (\text{resp. } \underline{f}'(x) = \min \partial f(x))$$

Lema 3.4. Si A, B y C son los conjuntos dados por

$$A = \{\bar{F}(x); x \in I\}, \quad B = \{\underline{f}'(x); x \in I\}, \quad C = \{F(x, y); x, y \in I, x \neq y\}$$

entonces se tiene

(i) $\sup A = \sup B = \sup C$, (ii) $\inf A = \inf B = \inf C$.

Demostración: Veamos que $\sup A = \sup C$, análogamente se demuestra que $\sup B = \sup C$. En efecto, supongamos que $\sup A > M$, existen $x_0 \in I$, $(x_n) \in S(f, x_0)$ tales que

$$\lim_{n \rightarrow \infty} F(x_0, x_n) > M$$

y por lo tanto

$$\exists n_0 \in \mathbb{N} \text{ tal que } F(x_0, x_n) > M \quad \forall n > n_0$$

luego $\sup A \leq \sup C$. Recíprocamente, sea $F(x_1, x_2) = P$ para algún par de puntos de I . La función $g(x) = f(x) - Px$ definida en I está en las condiciones del teorema 2.10, luego

$$\exists x_3 \in (x_1, x_2) \text{ tal que } 0 \in \partial g(x_3)$$

Por 2.7, $P \in \partial f(x_3)$, $\sup A \geq P$ y en consecuencia $\sup A \geq \sup B$. La demostración de (ii) es análoga.

Teorema 3.5.

(i) $\bar{F}(x) = \underline{f}'(x)$ si y solo si f es derivable en x .

(ii) Si una de las funciones derivada superior o derivada inferior es continua en x_0 , entonces también es continua la otra y además f es derivable en x_0 .

Demostración: La demostración de (i) es inmediata. Para (ii) supongamos que \bar{F} es continua en x_0 , entonces para todo $\epsilon > 0$, existe $\delta > 0$ tal que para todo $x \in (x_0 - \delta, x_0 + \delta)$,

$$\bar{F}(x) \in (\bar{F}(x_0) - \epsilon, \bar{F}(x_0) + \epsilon).$$

Del lema 3.2 se deduce que $F(x_2, x_1) \in (\bar{F}(x_0) - \epsilon, \bar{F}(x_0) + \epsilon)$ para cualquier par de puntos de cualquier intervalo cerrado contenido en $(x_0 - \delta, x_0 + \delta)$ y por lo tanto

$$\underline{f}'(x_0) \in (\bar{F}(x_0) - \epsilon, \bar{F}(x_0) + \epsilon)$$

Si \underline{f}' no es continua en x_0 , existen $\epsilon_1 > 0$ y puntos x de cualquier entorno de x_0 tales que

$$\underline{f}'(x) \notin (\bar{F}(x_0) - \epsilon, \bar{F}(x_0) + \epsilon)$$

lo que contradice el lema anterior. Si $\bar{F}(x_0) \neq \underline{f}'(x_0)$ también se contradice el lema, luego $\bar{F}(x_0) = \underline{f}'(x_0)$ y por (i) f es derivable en x_0 .

La utilización de G-derivadas nos permite dar una sencilla demostración del teorema de Stepanoff [5] en el caso de las funciones reales de variable real. Se supone que I es cerrado.

Teorema 3.6. Si f es f.c. en I , entonces f es derivable en casi todo punto de I .

Demostración: Al ser f f.c. en I , para cada $x_0 \in I$, existen $k > 0$ y un entorno $V(x_0)$ tales que $|F(x_0, x)| \leq k$ para todo $x \in V(x_0)$. Se define en $V(x_0)$ la función $g(x) = f(x) + 2kx$. Es claro que $\partial f(x_0) \subset [-k, k]$ y por 2.7 $\partial g(x_0) = \partial f(x_0) + 2k \in \mathbf{R}^+$, luego por 2.13 existe un entorno $U(x_0)$ en donde g es monótona y en consecuencia f es derivable en casi

todo punto de

$$W(x_0) = U(x_0) \cap V(x_0)$$

Como I es compacto, el recubrimiento abierto de I

$$\{W(x_0); x_0 \in I\}$$

contiene un subrecubrimiento finito y por lo tanto f es derivable en casi todo punto de I .

Si f es convexa es conocido [12] que existen las derivadas laterales en el sentido clásico y que además

$$f'_-(x) \leq f'_+(x) \text{ para cada } x \in I$$

en consecuencia

$$\partial f(x) = [f'_-(x), f'_+(x)] = \partial_{\mathbb{R}} f(x)$$

es decir, que para funciones convexas la G -derivada se reduce a la subdiferencial de Rockafellar [12]. Por otra parte, se demuestra fácilmente que si f es localmente lipschitziana en I , entonces $\partial f(x) \subset \partial_c f(x)$ lo que significa que la G -derivada es más reducida que la derivada de Clarke [2]. El contenido anterior puede ser estricto no solo en puntos aislados del dominio de f (ejemplo 3.2) sino incluso en cada punto del dominio como se pone de manifiesto en el ejemplo siguiente.

Ejemplo 3.7. Sea m la medida de Lebesgue unidimensional. El ejemplo se basa en la existencia de un Borel B tal que para todo segmento no vacío I , $0 < m(B \cap I) < m(I)$ (véase [6]). Se demuestra que tanto B como su complementario son subconjuntos densos en \mathbb{R} . Sea B un Borel de este tipo y X la función característica de B . Se define

$$f(x) = \int_{t \in [0, x]} X(t) dm(t)$$

f es continua en \mathbb{R} y además es lipschitziana. En efecto, dados los números reales x e y con $x < y$ se tiene

$$f(y) - f(x) = \int_{t \in [x,y]} X(t) dm(t) = y - x - m(\mathbb{R} \setminus E \cap [x,y]) < y - x$$

Como X es medible, f admite en casi todo punto derivada que coincide en c.t. punto con X , luego $\partial f(x) = \{f'(x)\}$ vale 0 o 1 en c.t. \mathbb{R} , mientras que al ser B y su complementario densos en \mathbb{R} , $\partial_c f(x) = [0,1]$ para todo $x \in \mathbb{R}$.

4. ESTUDIO DE EXTREMOS RELATIVOS.

En toda la sección, supondremos que I es abierto y que f es f.c. en I , $x_0 \in I$.

Definición 4.1. Se llama G -tangente de f en x_0 al haz de rectas

$$GTf(x_0) = \{y - f(x_0) = \xi(x - x_0); \xi \in \partial f(x_0)\}.$$

Se llama tangente superior (resp. inferior) de f en x_0 a la función T_s (resp. T_I) real de variable real definida por

$$T_s(x) = \sup_{\xi \in \partial f(x_0)} \{f(x_0) + \xi(x - x_0)\} \quad (\text{resp. } T_I(x) = \inf_{\xi \in \partial f(x_0)} \{f(x_0) + \xi(x - x_0)\})$$

Es claro que T_s es convexa (por ser el supremo de una familia de funciones afines) y que T_I es cóncava.

Teorema 4.2. Si $0 \in \partial f(x_0)$ y existe $\delta > 0$ tal que $f(x) \geq T_s(x)$ (resp. $f(x) \leq T_I(x)$) para todo $x \in (x_0 - \delta, x_0 + \delta)$, entonces x_0 es un mínimo (resp. máximo) relativo de f .

Demostración: Como $0 \in \partial f(x_0)$, la recta $y = f(x_0)$ pertenece a la G -tangente de f en x_0 y por lo tanto $T_s(x) \geq f(x_0)$ para todo $x \in \mathbb{R}$, luego $f(x) \geq f(x_0)$ para todo $x \in (x_0 - \delta, x_0 + \delta)$ y x_0 es un mínimo relativo de f . La demostración para el máximo es análoga.

Definición 4.3. La multifunción M que a cada elemento de I le hace corresponder un subconjunto de números reales es monótona no decreciente (resp. no creciente) en I si para todo par de

puntos x, y de I con $x > y$ se tiene que $\inf M(x) \geq \sup M(y)$ (resp. $\inf M(y) \geq \sup M(x)$).

Teorema 4.4.

- (i) f es convexa en I si y solo si ∂f es no decreciente en I .
- (ii) f es cóncava en I si y solo si ∂f es no creciente en I .

Demostración: (i) Si f es convexa en I existen derivadas laterales de f en cada punto de I , ambas son no decrecientes en I y tales que si $x, z_1, z_2 \in I$, $z_1 < x < z_2$ se verifica (véase [12]) que

$$f'_-(z_1) \leq f'_-(x) \leq f'_+(x) \leq f'_+(z_2)$$

Como además $\partial f(x) = [f'_-(x), f'_+(x)]$ resulta que ∂f es no decreciente en I .

Recíprocamente, supongamos que ∂f es no decreciente en I . Sean

$$a, b \in I, a < b \text{ y } g(x) = f(a) + \frac{f(b) - f(a)}{b - a} (x - a)$$

la recta secante a f en $(a, f(a)), (b, f(b))$. Cada x de (a, b) se puede expresar como $x = a + t(b - a)$ para algún t de $(0, 1)$, luego $g(x) = tf(b) + (1 - t)f(a)$. f es convexa en I si $f(x) \leq g(x)$ para todo $x \in (a, b)$ con $a, b \in I$, lo que es equivalente a

$$tf(x) + (1 - t)f(a) \leq tf(b) + (1 - t)f(a)$$

o lo que es lo mismo

$$(1 - t)[f(x) - f(a)] \leq t[f(b) - f(x)] \quad (1)$$

Aplicando a f el teorema 2.12 en $[a, x]$ y en $[x, b]$ se tiene respectivamente que

existen $c \in (a, x)$ y $A \in \partial f(c)$ tales que $f(x) - f(a) = A(x - a)$

existen $d \in (x, b)$ y $B \in \partial f(d)$ tales que $f(b) - f(x) = B(b - x)$

luego (1) es equivalente a $(1-t)(x-a)A \leq t(b-x)B$. Como $(1-t)(x-a) = t(b-x)$, la desigualdad anterior es cierta si $A \leq B$. Al ser $c < d$ y ∂f no decreciente, $A \leq B$ y en consecuencia f es convexa en I .

(ii) Basta tener en cuenta que f es cóncava en I , si y solo si $-f$ es convexa en I y que $\partial(-f)(x) = -\partial f(x)$, de donde por (i) se tiene el resultado.

Como consecuencia inmediata se tiene el corolario siguiente.

Corolario 4.5. Si $0 \in \partial f(x_0)$ y ∂f es monótona no decreciente (resp. no creciente) en $(x_0 - \delta, x_0 + \delta)$ para algún $\delta > 0$, entonces x_0 es un mínimo (resp. máximo) relativo de f .

Teorema 4.6.

(i) Si $0 \in \partial f(x_0) \setminus (\partial^+ f(x_0) \cup \partial^- f(x_0))$, $\partial^- f(x_0) \subset \mathbb{R}^-$ (resp. $\partial^+ f(x_0) \subset \mathbb{R}^+$) y $\partial^+ f(x_0) \subset \mathbb{R}^+$ (resp. $\partial^- f(x_0) \subset \mathbb{R}^-$), entonces x_0 es un mínimo (resp. máximo) relativo de f .

(ii) Si $0 \in \text{Int } \partial^+ f(x_0) \cup \text{Int } \partial^- f(x_0)$, entonces x_0 no es extremo relativo de f .

Demostración: (i) Supongamos el primer caso y veamos que existe $\delta^* > 0$ tal que

$$\forall x \in (x_0 - \delta^*, x_0 + \delta^*) \subset I, f(x) > f(x_0).$$

En efecto, dada $(x_n) \in S^-(f, x_0)$, $l^-(f, x_0, x_n) < 0$, luego existe $n^* \in \mathbb{N}$ tal que para todo $n > n^*$, $F(x_0, x_n) < 0$ y por lo tanto $f(x_n) > f(x_0)$.

Sea $\delta_1 = |x_{n^*} - x_0|$. Si para todo $x \in (x_0 - \delta_1, x_0)$, $f(x) > f(x_0)$ hacemos $\delta = \delta_1$; en caso contrario existe $x_1^* \in (x_0 - \delta_1, x_0)$ tal que $f(x_1^*) \leq f(x_0)$. Sea $\delta_2 = |x_1^* - x_0|$, si para todo $x \in (x_0 - \delta_2, x_0)$ se cumple $f(x) > f(x_0)$ hacemos $\delta = \delta_2$; en caso contrario exis-

te $x_2^* \in (x_0 - \delta_2, x_0)$ tal que $f(x_2^*) \leq f(x_0)$. Reiterando este proceso, al cabo de un número finito k de veces, se obtiene

$$\delta = \delta_k \text{ tal que } \forall x \in (x_0 - \delta, x_0), f(x) > f(x_0)$$

o bien no se obtiene ningún δ_k que lo cumpla con lo que se puede construir una sucesión $(x_n^*) \rightarrow 0^-$ tal que para todo $n \in \mathbb{N}$, se verifica que $f(x_n^*) - f(x_0) \leq 0$. En este caso, al ser f G-derivable en x_0 existe una subsucesión $(x_p^*) \subset (x_n^*)$ tal que $l(f, x_0, x_p^*) \geq 0$, lo que contradice la hipótesis.

Análogo razonamiento con $\partial^+ f(x_0)$ nos conduce a la existencia de un $\delta' > 0$ tal que $f(x) > f(x_0)$ para todo $x \in (x_0, x_0 + \delta)$ y considerando $\delta^* = \min(\delta, \delta')$ se tiene el resultado. La demostración para el máximo es análoga.

(ii) Supongamos que $0 \in \text{Int } \partial^+ f(x_0)$, existen $(x_n), (x_p) \in S^+(f, x_0)$ tales que $l^+(f, x_0, x_n) > 0$ y $l^+(f, x_0, x_p) < 0$ y por lo tanto existen $N, P \in \mathbb{N}$ de forma que

$$\frac{f(x_n) - f(x_0)}{x_n - x_0} > 0 \quad \forall n > N \quad \text{y} \quad \frac{f(x_p) - f(x_0)}{x_p - x_0} < 0 \quad \forall p > P$$

de donde se deduce que x_0 no es extremo relativo de f .

REFERENCIAS

- [1] BUTLER, G. J., TIMOURIAN, J. G., VIGER, C. The rank theorem for locally lipschitz continuous functions. *Canad. Math. Bull.* 31 n°2 (1988); 217-226.
- [2] CLARKE, F. Generalized gradients and applications. *Trans. Amer. Math. Soc.* 305 (1975), 246-262.

- [3] CLARKE, F. On the inverse function theorem. *Pacif. Journ. Math.* 64 nº1 (1976), 97-102.
- [4] CLARKE, F. Generalized gradients of Lipschitz functionals. *Adv. in Math.* 40 (1981), 52-67.
- [5] FEDERER, H. *Geometric Measure Theory*. Springer Verlag. New York (1969).
- [6] HIRIART-URRUTY, J. B. Miscellanies of non-smooth analysis and optimization. *Lect. Notes in Econ. and Math. Syst.* 225 (1985), 8-24.
- [7] KIRK, R. B. Sets with split families of measurable sets. *Amer. Math. Mont.* 79 (1972), 884-886.
- [8] LEBOURG, G. Valeur moyenne pour gradient généralisé. *C.R. Acad. Sc. Paris.* 281 10 (1975), 795-797.
- [9] NOVO, V., RODRIGUEZ MARIN, L. Derivada generalizada en \mathbf{R} . Aplicación al estudio de extremos. *Actas XV Jorn. Luso-Esp. de Mat. Univ. de Evora* (1990).
- [10] NOVO, V., RODRIGUEZ MARIN, L. An extension of the inverse function theorem. *Rev. Real Acad. Cien. Exac. Fis. y Nat.* (preprint).
- [11] POURCIAU, B. Global invertibility of nonsmooth mappings. *J. Math. Anal. Appl.* 131 (1988), 170-179.
- [12] ROCKAFELLAR, T. *Convex Analysis*. Princ. Univ. Press (1970).

IMPLICIT OUTPUT BY CONTINUOUS RUNGE-KUTTA METHODS

T. GRANDE, J. I. MONTIJANO, L. RÁNDEZ

Departamento de Matemática Aplicada.
Universidad de Zaragoza.
50009 Zaragoza

Abstract. In this paper, continuous RK methods are applied to obtain points of implicit output. Such points t^* are defined as the roots of

$$(1) \quad g(t^*, y(t^*), y'(t^*)) = 0.$$

for some given function g . These situations are common in problems such as the search for periodic solutions, the estimation of extremum values of the solution of a differential equation, and so on. To solve (1), a damped Newton method is applied and the function $y(t)$ and $y'(t)$ are approximated by continuous RK methods. Finally, some numerical examples show a behaviour according to the theoretical study carried out.

AMS (MOS) Subject Classification: 65L07, CR517

1. Introduction

Let us consider an initial value problem

$$(1.1) \quad \begin{aligned} y'(t) &= f(t, y(t)), \quad t \geq t_0, \quad y \in \mathbb{R}^N, \\ y(t_0) &= y_0, \end{aligned}$$

and let y_n the approximation to $y(t_n)$, solution of (1.1) at t_n , be given by a numerical method. Sometimes the final point of the integration t_{end} where we want to get the solution, is unknown. For example, in the study of cycle limits in the analysis of periodic solutions, it would be desirable that final points were the intersection of the solution with a Poincaré section. The search for these special points can be formulated in a general manner as: to find the points t^* such that

$$(1.2) \quad G(t^*) = g(t^*, y(t^*), y'(t^*)) = 0,$$

is satisfied for a given function g , where $y(t)$ is the solution of (1.1) (*Implicit output problem*).

It takes a great computational effort to locate such points if only a discrete solution y_n is known because some iterative method for searching roots of a function must be used

Partially supported by C.Y.C.I.T. PS90/0121

and it is necessary to prepare the integration code for positive and negative stepsizes. For this class of problems is almost essential to make use of a continuous solution that can be got by some interpolation process. Recently, Horn [4], Shampine [7], Enright et al. [2], have introduced some RK methods that, by means of interpolatory techniques, define a global numerical solution of class $C^{(1)}$ with a suitable order of approximation. These methods are called continuous RK and they seem to be very efficient for the solution of implicit output problems.

In the second paragraph, a general study of implicit output problems is carried out, and it is proved that under some regularity conditions on the function G , the error in the computation of t^* can be bounded.

In the third paragraph several implicit output problems are solved using for the integration of (1.1) a embedded RK pair of orders 5 and 6 given by Calvo, Montijano and Rández [1], which defines a 5th order continuous solution. Finally, it must be pointed out that the nonlinear equation (1.2) is solved by means of a damped Newton method [5].

2. Continuous Runge-Kutta methods for implicit output

The implicit output problem consists basically on locating a zero t^* of the given function g . However, the function $y(t)$ must be obtained integrating the IVP (1.1) and this fact makes very expensive the usual procedures for finding the roots of a function.

Thus, for example, let's suppose that in the interval $[t_n, t_{n+1}]$ a change in the sign of the function g has been detected, and the process to locate the root starts. We can use the bisection procedure which is trustworthy but slow and very expensive. For this method, we need to calculate the sign of g at the middle point $\hat{t} = (t_n + t_{n+1})/2$, but $y(\hat{t})$ must be obtained using the integration code from t_n to \hat{t} with stepsize $h = \hat{t} - t_n$. If a s stages RK method is used, $s - 1$ evaluations of function f are necessary. Depending on the sign of g at \hat{t} , the root is in a semiinterval, and the process is iterated with the middle point of the semiinterval.

It seems clear that this algorithm is not suitable due to the great number of calculations needed to achieve a specified tolerance because $M \times s$ evaluations of the derivative function f must be calculated if we need to apply M times the bisection process.

It is more convenient to use the Newton algorithm to reduce considerably the number of iterations. Starting from the initial approximation $t^{(0)} = t_n$ (or $t^{(0)} = t_{n+1}$) a sequence $t^{(i)}$ is obtained from

$$(2.1) \quad t^{(i+1)} = t^{(i)} - \frac{g(t^{(i)}, y^{(i)})}{g'(t^{(i)}, y^{(i)})} \quad i = 0, 1, \dots$$

$y^{(i)}$ being an approximation to the theoretical solution $y(t^{(i)})$ and with

$$(2.2) \quad g'(t, y) = \frac{\partial g}{\partial t}(t, y) + \frac{\partial g}{\partial y}(t, y) f(t, y).$$

We suppose that sufficient conditions of convergence are satisfied so that the succession $t^{(i)}$ converges towards the root t^* . It is clear that $t^{(1)}$ is obtained easily from the above expression but to calculate the following iteration $t^{(i)}$ we need to evaluate $g(t^{(i)}, y^{(i)})$ and $g'(t^{(i)}, y^{(i)})$ and the value of $y^{(i)}$ is not yet known. This value can be obtained by a Runge-Kutta method from $t^{(i-1)}$ to $t^{(i)}$, and the integration code must be implemented for positive or negative steps. Also, notice that a point $t^{(i)}$ may lie out of the interval

$[t_n, t_{n+1}]$ so that Newton's process instead of being faster than bisection method is more hazardous and doesn't guarantee the convergence. By this reason, it is convenient to use damped Newton's methods which increase the domain of convergence [5].

Due to the fact that in each iteration it is necessary to calculate (2.2), it would be appropriate to deploy of a continuous approximation to the solution $y(t)$ and its derivatives. It is important to emphasize that when we are looking for the root in the interval $[t_n, t_{n+1}]$, as the continuous approximation to the solution has been calculated, the discrete RK method is not used and therefore no additional evaluation of f are needed. In these terms, if $Q(t)$ is an order p continuous approximation to the solution associate to the RK method in the interval $[t_n, t_{n+1}]$ it can be shown (Shampine [6]) that

$$|Q^{(k)}(t) - y^{(k)}(t)| = \mathcal{O}(h^{p-k}), \quad t \in [t_n, t_{n+1}],$$

and the problem (1.2) can be replaced by

$$(2.3) \quad \hat{G}(t) = g(t, Q(t)) = 0.$$

In the same way the more general implicit output problem

$$G(t) = g(t, y(t), y'(t), \dots, y^{(k)}(t)).$$

can be substituted by

$$\hat{G}(t) = g(t, Q(t), Q'(t), \dots, Q^{(k)}(t)).$$

The fact that the problem (1.2) can be approximated by (2.3) is justified by the following:

Proposition 1 Let us suppose that the function $G(t) = g(t, y(t), y'(t), \dots, y^{(k)}(t))$ is of class $C^{(1)}[t_n, t_{n+1}]$, and that $|(G'(t))^{-1}| \leq \beta$, for all $t \in [t_n, t_{n+1}]$ with $\beta > 0$ and g is Lipschitz with respect to the variables $y, y', \dots, y^{(k)}$ uniformly on t . If $G(t^*) = 0$ and $\hat{G}(\hat{t}) = 0$ for $t^*, \hat{t} \in [t_n, t_{n+1}]$ then

$$|t^* - \hat{t}| = \mathcal{O}(h^{p-k}).$$

In view of the above proposition, we conclude that the higher is the order of the derivatives, the lower is the approximation order in the case that the derivatives of the function are approximated by the derivatives of the interpolating polynomial. However, it is possible to maintain the order p of approximation defined by the continuous RK even though the derivatives are present. This is shown in the:

Proposition 2 If the functions g and G satisfy the hypothesis of the proposition 1, and if we consider the function

$$\bar{G}(t) = g(t, Q(t), Q_1(t), \dots, Q_k(t)),$$

where

$$Q_1(t) = f|_{y=Q(t)},$$

$$Q_2(t) = f'f|_{y=Q(t)},$$

$$Q_3(t) = (f'f'f + f''(f, f))|_{y=Q(t)},$$

$$\vdots \quad \vdots$$

and $G(t^*) = 0$, $\bar{G}(\bar{t}) = 0$ with $t^*, \bar{t} \in [t_n, t_{n+1}]$ then

$$|t^* - \bar{t}| = \mathcal{O}(h^p).$$

According to this proposition, to obtain order p , the derivatives of $y(t)$ up to order k must be evaluated at $Q(t)$ for every iteration.

3. Numerical Results

In the numerical experiments a RK pair of orders 6 and 5 with 8 stages [1] has been used as integrator. This pair has a RK interpolant of order 5 costing one extra evaluation per step. To solve the non linear equation, a damped Newton method, with automatic selection of the damping parameters λ_i has been chosen. Then, equation (2.1) is replaced by

$$t^{(i+1)} = t^{(i)} - \lambda_i \frac{g(t^{(i)}, y^{(i)})}{g'(t^{(i)}, y^{(i)})}, \quad \lambda_i \in (0, 1], \quad i = 0, 1, \dots$$

As a first problem we consider A3 from the DETEST package [3],

$$\begin{aligned} y' &= y \cos(t), & t \in [0, 20] \\ y(0) &= 1, \end{aligned}$$

for the functions $G(t)$ given by $G_1(t) = g(t, y(t)) = y(t) - 1$, $G_2(t) = g(t, y(t), y'(t)) = y'(t)$ and $G_3(t) = g(t, y(t), y'(t), y''(t)) = y''(t)$.

In table 1 we show the maximum of the errors $|t^* - \hat{t}|$ obtained in the location of the roots for the functions $G(t)$ above specified, and the maximum number of Newton iterations for different values of tolerance. A test of pure absolute error has been applied.

Table 1

	$G_1(t)$		$G_2(t)$		$G_3(t)$	
	max. err.	ite. num.	max. err.	ite. num.	max. err.	ite. num.
$TOL = 10^{-4}$	1.01×10^{-4}	3	1.17×10^{-4}	3	1.20×10^{-3}	3
$TOL = 10^{-5}$	2.96×10^{-5}	3	2.38×10^{-5}	2	3.35×10^{-4}	3
$TOL = 10^{-6}$	7.10×10^{-7}	3	2.92×10^{-6}	2	6.78×10^{-5}	3
$TOL = 10^{-7}$	1.62×10^{-8}	3	2.68×10^{-7}	2	1.70×10^{-5}	3
$TOL = 10^{-8}$	4.00×10^{-9}	3	6.73×10^{-8}	2	3.23×10^{-6}	3

It must be noticed that the numerical results show a behaviour according to proposition 1. From table 1, it is clear that the best results are obtained for the function G_1 , which only depends on y , and the errors are bigger if higher derivatives of $y(t)$ are used.

The second test problem is the well known two body problem that belongs to class D (orbital problems) of DETEST package. We consider in the interval $[0, 20]$ the initial value problem

$$\begin{aligned} y_1' &= y_3, & y_1(0) &= 1 - e, \\ y_2' &= y_4, & y_2(0) &= 0, \\ y_3' &= -y_1/(y_1^2 + y_2^2)^{3/2}, & y_3(0) &= 0, \\ y_4' &= -y_2/(y_1^2 + y_2^2)^{3/2}, & y_4(0) &= \sqrt{(1+e)/(1-e)}, \end{aligned}$$

where e is the orbital eccentricity. As the function $G(t)$ we take the second component of the solution, $G(t) = y_2(t)$. It is important to note that $G(t)$ vanishes when the satellite is at its apogee or perigee (which corresponds with $t = k\pi$ for some integer k). We have chosen this problem because the common integrators must carry out a good stepsize control, and therefore the difficulty of finding the roots of the function G is mixed with the degree of difficulty of the numerical integration of this problem. In Table 2 is shown the maximum error in the process of locating the roots and the maximum number of iteration by Newton methods for a range of tolerance values and for different eccentricity values. In each case two results are shown; the above one corresponds to the 6(5) integrator in [1], and the below one, which gives worse outcomes corresponds to the widely used DVERK 6(5) with a fifth order continuous method [3].

Table 2

	$e = .5$ max. err.	ite. num.	$e = .7$ max. err.	ite. num.	$e = .9$ max. err.	ite. num.
$TOL = 10^{-4}$	5.21×10^{-3}	2	6.00×10^{-3}	2	5.80×10^{-3}	2
	2.08×10^{-2}	2	4.50×10^{-2}	2	7.02×10^{-2}	2
$TOL = 10^{-5}$	4.77×10^{-4}	2	4.05×10^{-4}	2	7.35×10^{-5}	2
	1.61×10^{-3}	2	1.32×10^{-3}	2	2.73×10^{-3}	2
$TOL = 10^{-6}$	4.90×10^{-5}	2	2.42×10^{-5}	2	2.36×10^{-6}	2
	1.42×10^{-4}	2	6.86×10^{-4}	2	6.41×10^{-4}	2
$TOL = 10^{-7}$	4.34×10^{-6}	2	1.58×10^{-6}	2	6.76×10^{-7}	2
	1.84×10^{-5}	2	1.41×10^{-4}	2	1.66×10^{-4}	2
$TOL = 10^{-8}$	3.37×10^{-7}	2	1.02×10^{-7}	2	6.14×10^{-8}	2
	3.53×10^{-6}	2	2.06×10^{-5}	2	2.94×10^{-5}	2

Acknowledgement

The authors wish to thank Professor J. R. Dormand for his careful reading of the manuscript.

REFERENCES

- [1] Calvo, M., Montijano, J. I., Rández, L.; A New embedded pair of Runge-Kutta formulas of orders 5 and 6. *Computers and Mathematics with Applications*. Vol. 20, N° 1, (1990), pp. 15-24.
- [2] Enright, W. H., Jackson, K. R., Norsett, S. P., Thomsen, P. G.; Interpolants for Runge-Kutta formulas. *ACM transactions on Mathematical Software*, 12, N° 3, (1986), pp. 193-218.
- [3] Enright, W. H., Pryce J. D.; Two FORTRAN packages for assessing initial value methods. *ACM TOMS*, Vol. 13, N° 1, (1987), pp. 1-27.
- [4] Horn, M. K.; Fourth and fifth-order, scaled Runge-Kutta algorithms for treating dense output. *SIAM J. Numer. Anal.* 20, (1983), pp. 558-568.
- [5] Ortega, J. M., Rheinboldt, W. C.; *Iterative Solution of Non-linear Equations in Several Variables*. New York: Academic Press 1970.
- [6] Shampine, L. F.; Interpolation for Runge-Kutta methods. *SIAM J. Numer. Anal.* N° 22, 5, (1985) pp. 1014-1027.
- [7] Shampine, L. F.; Some Practical Runge-Kutta Formulas. *Math. of Comp.* N° 173 (1986), pp. 135-150.

A NEW SMOOTHED-STEP FUNCTION

M. Membrado ¹, A. F. Pacheco ¹ and J. Sañudo ²

¹ Consejo Superior de Investigaciones Científicas (I.C.M.A.) and
Facultad de Ciencias. Universidad de Zaragoza. 50009 Zaragoza. Spain.

² Departamento de Física. Facultad de Ciencias.
Universidad de Extremadura. 06071 Badajoz. Spain.

Abstract

A new smoothed-step function is presented. Several of its integrals such as momenta, form factor and Coulomb energy are given. The delta function obtained after derivation is also discussed; its form in spherical coordinates seems to be particularly useful.

Introduction and Properties

One of the common uses of smoothed-step functions in Physics is the simulation of density profiles of different saturating systems such as classical liquid drops, atomic nuclei, etc. For this job the Fermi function

$$\frac{n(r)}{n_0} = \frac{1}{1 + e^{(r-R)/a}} \quad (1)$$

is perhaps the most widely used. R is the half density radius, a the surface thickness, and, when R is assumed to be much larger than a , n_0 is the approximate density at the origin. As is well known, the momenta and other integrals¹ involving (1), and also its generalizations², admit expansions in powers of a/R .

In this paper we first want to introduce a new type of smoothed-step function which has appeared in the description of electron distributions of small metal clusters, in the jellium approximation for the positive background, and assuming that the intrinsic chemical potential of the electron cloud is proportional to the density³. Later, it has been also applied for the description of small helium aggregates⁴. As it is said, its origin and initial applications are some topics of Condensed Matter but this new function can logically be applied to other physical systems and in this sense can be an alternative to (1). It is given by

$$f(r) \equiv \frac{n(r)}{n_0} = \left[1 - A \frac{\sinh(sr)}{r} \right] \quad r \leq R \quad , \quad (2a)$$

$$f(r) \equiv \frac{n(r)}{n_0} = B \frac{e^{-sr}}{r} \quad r > R \quad , \quad (2b)$$

with

$$A = R \left[1 + \frac{1}{z} \right] e^{-z} \quad , \quad (3a)$$

$$B = R \left[\cosh z - \frac{\sinh z}{z} \right] \quad , \quad (3b)$$

$$z = sR \quad . \quad (3c)$$

Obviously spherical coordinates are always assumed, and the number of independent parameters introduced is three: n_0 , R , and s . The normalization condition is simply given by

$$Z = \int n \, dr = \frac{4}{3} \pi R^3 n_0 \quad . \quad (4)$$

Before addressing several integrals involving f , let us advance that series expansions are not required in a number of common cases; this is a clear advantage with respect to the Fermi function.

With respect to the momenta of f

$$\langle r^m \rangle \equiv \frac{1}{Z} \int n(r) r^m \, dr \quad , \quad (5)$$

one easily obtains

$$\langle r^m \rangle = R^m \left\{ \frac{3}{m+3} + \frac{3}{2z^{m+3}} [e^{-z} (z+1) (G_m + F_m - E_m) + e^z (z-1) G_m] \right\}, \quad (6)$$

where the functions G_m , F_m and E_m fulfill

$$E_m = z^{m+1} e^z - (m+1) E_{m-1}, \quad (7a)$$

$$F_m = -z^{m+1} e^{-z} + (m+1) F_{m-1}, \quad (7b)$$

$$G_m = z^{m+1} e^{-z} + (m+1) G_{m-1}, \quad (7c)$$

and their initial values are

$$E_0 = e^z (z-1) + 1, \quad (8a)$$

$$F_0 = -e^{-z} (z+1) + 1, \quad (8b)$$

$$G_0 = e^{-z} (z+1). \quad (8c)$$

Equations (7) derive from the recursion relations fulfilled by integrals involving exponentials and powers. As an example, in a well characterized case, the mean-square radius is simply given by

$$\langle r^2 \rangle = R^2 \left[\frac{3}{5} + \frac{6}{z^2} \right]. \quad (9)$$

The form factor, i.e. the Fourier transform of the distribution, has also a closed form, namely

$$F(q) \equiv \frac{1}{Z} \int e^{iqr} n(r) dr = \frac{3 (\sin x - x \cos x)}{x^3 (1 + y^2)}, \quad (10)$$

$$x \equiv qR, \quad y \equiv \frac{q}{s}.$$

In eq. (10), the form factor of the step function is easily recovered by putting $y=0$.

Finally, with regard to electrostatic properties of the system, let us compute the Coulomb self-energy of the distribution:

$$E_c = \frac{1}{2} \int \frac{n(\mathbf{r}_1) n(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad (11)$$

Again, it adopts a closed analytical form:

$$E_c = \frac{Z^2}{R} \left\{ \frac{3}{5} - \frac{3}{z^2} + \frac{45}{8z^3} - \frac{63}{8z^5} + e^{-2z} \left[\frac{9}{4z^2} + \frac{81}{8z^3} + \frac{63}{4z^4} + \frac{63}{8z^5} \right] \right\} \quad (12)$$

There is no problem with this formula when z is very small; it tends to

$$\lim_{z \rightarrow 0} E_c = \frac{Z^2}{R} \left\{ \frac{z}{4} - \frac{z^3}{20} + \dots \right\} \quad (13)$$

Representation of the Delta Function

After studying these integrals, let us now analyze the delta function emerging from the derivative of the newly presented smoothed-step function. The usefulness of the Dirac delta function, δ , in Physics and in particular in Quantum Mechanics⁵ is well known. One of the common representations that defines this distribution in Cartesian coordinates is the sequence

$$D_n(x-x') = \frac{\sin n(x-x')}{\pi(x-x')} \quad (n=1,2,\dots) \quad (14)$$

In spherical coordinates, δ is defined as

$$\delta(\mathbf{r}-\mathbf{R}) = \frac{1}{r^2 \sin\theta} \delta(r-R) \delta(\theta-\Theta) \delta(\phi-\Phi) \quad (15)$$

and from (14) we can find a representation for the radial components $\delta(r-R)$; the result is

$$D_n(r-R) = \frac{r}{R} \left[\frac{\sin n(r-R)}{\pi(r-R)} - \frac{\sin n(r+R)}{\pi(r+R)} \right] \quad (16)$$

In eqs.(2), roughly speaking, R represents the radius of the distribution and $1/s$ the "skin" depth. Since at the limit of infinite s , f converts into the radial step function, the derivative of f must lead to δ . We obtain

$$f(r) = \begin{cases} \frac{(R+1/s)}{2r^2} [(sr-1)e^{s(r-R)} + (sr+1)e^{-s(r+R)}] & r \leq R \quad (17a) \\ \frac{(sr+1)}{2r^2} [(R-1/s)e^{-s(r-R)} + (R+1/s)e^{-s(r+R)}] & r > R \quad (17b) \end{cases}$$

We define a new function g , simply to extract the Jacobian factor $1/r^2$:

$$g(r) = f(r) r^2 \quad (18)$$

which satisfies the following normalization condition:

$$\int_0^{\infty} g(r) dr = R^2 - \frac{2}{s^2} + \frac{2}{s^2} \left(1 + \frac{1}{sR} \right) e^{-sR} \quad (19)$$

Hence, after discretizing s , the representation of the radial component of the δ

$$\delta(r-R) = \lim_{n \rightarrow \infty} F_n(r-R) \quad (20)$$

is

$$F_n(r-R) = \frac{1}{N} \begin{cases} \frac{(R+1/n)}{2r^2} [(nr-1)e^{n(r-R)} + (nr+1)e^{-n(r+R)}] & r \leq R \quad (21a) \\ \frac{(nr+1)}{2r^2} [(R-1/n)e^{-n(r-R)} + (R+1/n)e^{-n(r+R)}] & r > R \quad (21b) \end{cases}$$

$$(n=1,2,\dots)$$

with

$$N = R^2 - \frac{2}{n^2} + \frac{2}{n^2} \left(1 + \frac{1}{nR} \right) e^{-nR} \quad (22)$$

The representation of δ in Cartesian coordinates related to eqs. (21) is the sequence

$$F_n(x-x') = \begin{cases} \frac{n}{2} e^{n(x-x')} & x \leq x' \quad (23a) \\ \frac{n}{2} e^{-n(x-x')} & x > x' \quad (23b) \end{cases}$$

which is already normalized. In order to test the performance of these new sequences F_n and \mathbb{F}_n , and to compare for example with D_n and \mathbb{D}_n , we have numerically calculated several examples of integrals of the type

$$\int_{-\infty}^{+\infty} F_n(x-x') y(x) dx \quad (24a)$$

and

$$\int_0^{+\infty} \mathbb{F}_n(r-R) Y(r) dr \quad (24b)$$

to observe how fast they converge to $y(x')$ and $Y(R)$ respectively, as n increases.

Our results, for several characteristic cases such as monotonous or oscillatory functions, indicate that the convergence obtained from D_n is higher than that of \mathbb{D}_n but this is inverted for the polar coordinates case where \mathbb{F}_n is far more efficient than F_n . As an example suppose the following case:

$$y(x) = \frac{2}{|x|^{3/2} + 1} \quad (25a)$$

and

$$Y(r) = \frac{2}{r^{3/2} + 1} ; \quad (25b)$$

the results of eqs. (24a) and (24b), when $x'=R=1$, are plotted in Fig.1 and Fig.2, respectively.

To conclude, we remark that as F_n and \mathbb{F}_n are constructed using exponentials, a number of explicit analytical integrations involving them can be obtained. Besides, the fact that the second derivative of f is discontinuous is a property, not present in other well known smoothed-step functions, which may have significant effects.

Acknowledgements

Discussions with J.G. Esteve and J. Sesma are acknowledged. This work was supported by the DGICYT (PB 90-0916).

References

- 1.- See for example: A. Bohr and B. Mottelson. "Nuclear Structure" Vol. 1. Benjamin (1961), pag. 160.
- 2.- H. Krivine and J. Treiner. J. Math. Phys. 22 (11) 2484 (1981).
D.K. Srivastava. Phys. Lett. 112B 289 (1982).
- 3.- M. Membrado, A.F. Pacheco and J. Sañudo. Phys. Rev. B41, 5643 (1990).
- 4.- L.C. Balbás, A. Mañanes, M. Membrado, A.F. Pacheco and J. Sañudo. J. Chem. Phys. 94, 7335 (1991).
- 5.- See for example: A. Messiah. "Mécanique Quantique". Dunod. Paris (1959).

Figure Captions

Fig. 1.- Convolution of the δ sequence in Cartesian coordinates (continuous line: D_n , broken line: F_n) with (25a), at $x = 1$, as a function of n .

Fig. 2.- Convolution of the radial δ sequence (continuous line: D_n , broken line: F_n) with (25b), at $r = 1$, as a function of n .

FIG. 1

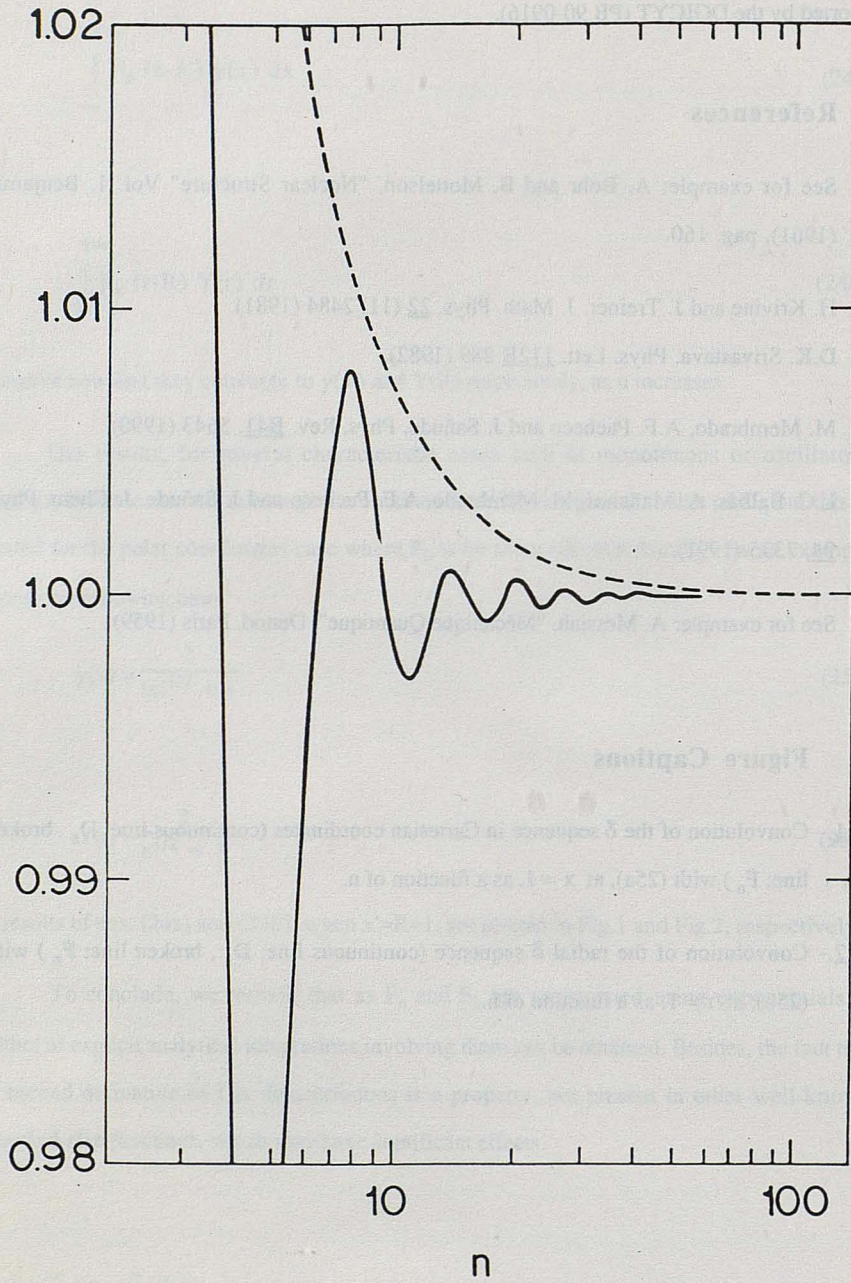
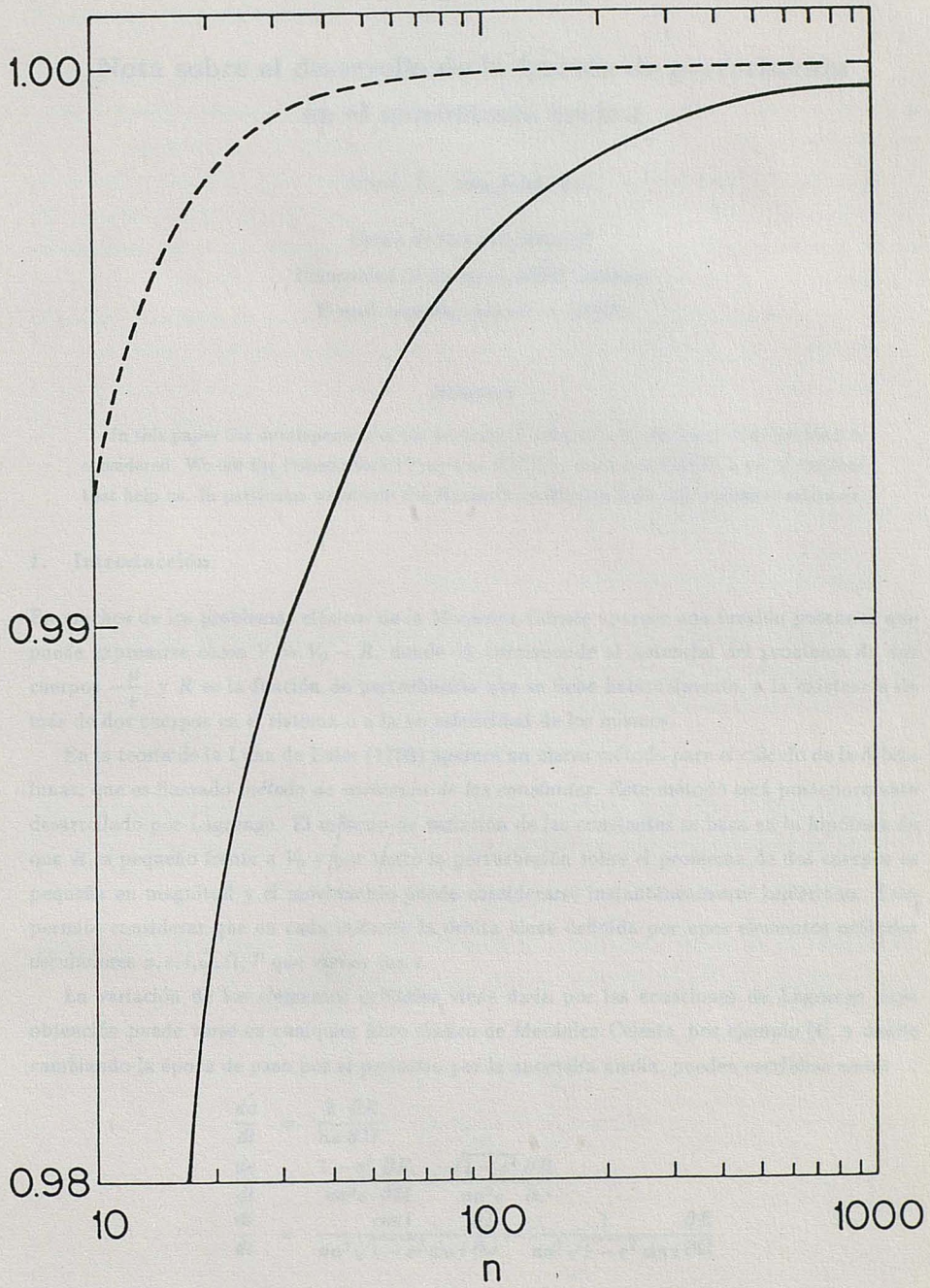


FIG. 2



Nota sobre el desarrollo de la función de perturbación en el movimiento orbital

Abad, A., San Juan, J.F.

Grupo de Mecánica Espacial

Universidad de Zaragoza, 50009 Zaragoza

E-mail: abad@cc.unizar.es (EAN)

Abstract

In this paper the development of the perturbation function in the three body problem is considered. We use the Poisson Series Processor PSPC in order to construct a set of routines that help us. In particular we obtain the Hansen's coefficients with easy recursive relations.

1. Introducción

En muchos de los problemas clásicos de la Mecánica Celeste aparece una función potencial que puede expresarse como $V = V_0 - R$, donde V_0 corresponde al potencial del problema de dos cuerpos $-\frac{\mu}{r}$, y R es la función de perturbación que se debe habitualmente, a la existencia de más de dos cuerpos en el sistema o a la no esfericidad de los mismos.

En la teoría de la Luna de Euler (1753) aparece un nuevo método para el cálculo de la órbita lunar, que es llamado *método de variación de las constantes*. Este método será posteriormente desarrollado por Lagrange. El método de variación de las constantes se basa en la hipótesis de que R es pequeño frente a V_0 y por tanto la perturbación sobre el problema de dos cuerpos es pequeña en magnitud y el movimiento puede considerarse instantáneamente kepleriano. Esto permite considerar que en cada instante la órbita viene definida por unos elementos orbitales osculadores $a, e, i, \omega, \Omega, T$ que varían con t .

La variación de los elementos orbitales viene dada por las ecuaciones de Lagrange cuya obtención puede verse en cualquier libro clásico de Mecánica Celeste, por ejemplo [4], y donde cambiando la época de paso por el periastro por la anomalía media, pueden escribirse como

$$\begin{aligned}\frac{da}{dt} &= \frac{2}{na} \frac{\partial R}{\partial M} \\ \frac{de}{dt} &= \frac{1-e^2}{na^2 e} \frac{\partial R}{\partial M} - \frac{\sqrt{1-e^2}}{na^2 e} \frac{\partial R}{\partial \omega} \\ \frac{di}{dt} &= \frac{\cos i}{na^2 \sqrt{1-e^2} \sin i} \frac{\partial R}{\partial \omega} - \frac{1}{na^2 \sqrt{1-e^2} \sin i} \frac{\partial R}{\partial \Omega}\end{aligned}$$

$$\begin{aligned}
 \frac{d\omega}{dt} &= \frac{\sqrt{1-e^2}}{na^2e} \frac{\partial R}{\partial e} - \frac{\cos i}{na^2\sqrt{1-e^2} \sin i} \frac{\partial R}{\partial i} \\
 \frac{d\Omega}{dt} &= \frac{1}{na^2\sqrt{1-e^2} \sin i} \frac{\partial R}{\partial i} \\
 \frac{dM}{dt} &= n - \frac{2}{na} \frac{\partial R}{\partial a} - \frac{1-e^2}{na^2e} \frac{\partial R}{\partial e}
 \end{aligned} \tag{1}$$

En un reciente artículo [3] se describe como obtener, por ordenador, una integración analítica de dichas ecuaciones utilizando el método estroboscópico. Esta integración se realiza utilizando el procesador de series de Poisson PSPC [2] que está siendo construido por nosotros.

La anterior integración es realizada dando como única entrada en el programa la función de perturbación R expresada en función de los elementos orbitales. Sin embargo, la construcción de esta función R requiere una gran cantidad de cálculos algebraicos, por lo que resulta muy conveniente para su desarrollo el uso de un manipulador algebraico como PSPC.

En general R viene expresado en función de r y f , por lo que antes de abordar el cálculo de dicha función deben ser obtenidos los desarrollos de Hansen que nos dan la expresión de

$$\begin{aligned}
 &\left(\frac{r}{a}\right)^n \cos mf \\
 &\left(\frac{r}{a}\right)^n \sin mf
 \end{aligned}$$

en función de los elementos orbitales.

2. Desarrollos en función de la anomalía media

Para la obtención de estos desarrollos hemos utilizado relaciones que resultan mucho más sencillas de programar que las relaciones dadas por Giacaglia [8].

A partir de la Ecuación de Kepler

$$E - e \sin E = M$$

se obtendrá el desarrollo de la anomalía excéntrica, E , en función de la anomalía media, M . Para ello, se utilizarán las conocidas fórmulas de recurrencia

$$E_{n+1} = M + e \sin E_n = M + e \sin(M + e \sin E_{n-1})$$

comenzando en $E_0 = M$, y efectuando el desarrollo de Taylor en torno a M de la función $\sin(M + e \sin E_{n-1})$. De esta forma se obtiene hasta el orden 4 en e

$$\begin{aligned}
 E &= M \\
 &+ \left\{ e - \frac{1}{8}e^3 + \dots \right\} \sin(M) \\
 &+ \left\{ \frac{1}{2}e^2 - \frac{1}{6}e^4 + \dots \right\} \sin(2M) \\
 &+ \left\{ \frac{3}{8}e^3 + \dots \right\} \sin(3M) \\
 &+ \left\{ \frac{1}{3}e^4 + \dots \right\} \sin(4M) \\
 &+ \dots
 \end{aligned}$$

Para obtener el desarrollo del $\text{sen } E$, sustituiremos la expresión anterior en la fórmula

$$\text{sen } E = \frac{1}{e}(E - M)$$

obtenida de la ecuación de Kepler, resultando

$$\begin{aligned} \text{sen } E &= \left\{ 1 - \frac{1}{8}e^2 + \frac{1}{192}e^4 + \dots \right\} \sin(M) \\ &+ \left\{ \frac{1}{2}e^1 - \frac{1}{6}e^3 + \dots \right\} \sin(2M) \\ &+ \left\{ \frac{3}{8}e^2 - \frac{27}{128}e^4 + \dots \right\} \sin(3M) \\ &+ \left\{ \frac{1}{3}e^3 + \dots \right\} \sin(4M) \\ &+ \left\{ \frac{125}{384}e^4 + \dots \right\} \sin(5M) \\ &+ \dots \end{aligned}$$

El $\text{cos } E$ lo construiremos sustituyendo E por su desarrollo en serie y efectuando el desarrollo de Taylor de la función $\cos(M + \Delta)$ en torno a M , lo que nos da

$$\begin{aligned} \text{cos } E &= -\frac{1}{2}e + \dots \\ &+ \left\{ 1 - \frac{3}{8}e^2 + \frac{5}{192}e^4 + \dots \right\} \cos(M) \\ &+ \left\{ \frac{1}{2}e^1 - \frac{1}{3}e^3 + \dots \right\} \cos(2M) \\ &+ \left\{ \frac{3}{8}e^2 - \frac{45}{128}e^4 + \dots \right\} \cos(3M) \\ &+ \left\{ \frac{1}{3}e^3 + \dots \right\} \cos(4M) \\ &+ \left\{ +\frac{125}{384}e^4 + \dots \right\} \cos(5M) \\ &+ \dots \end{aligned}$$

Por medio del desarrollo del $\text{cos } E$ obtenemos el de $\frac{r}{a}$ sin más que considerar la relación

$$\frac{r}{a} = 1 - e \cos E$$

por tanto

$$\begin{aligned} \frac{r}{a} &= 1 + \frac{1}{2}e^2 + \dots \\ &+ \left\{ -e + \frac{3}{8}e^3 + \dots \right\} \cos(M) \\ &+ \left\{ -\frac{1}{2}e^2 + \frac{1}{3}e^4 + \dots \right\} \cos(2M) \\ &+ \left\{ -\frac{3}{8}e^3 + \dots \right\} \cos(3M) \\ &+ \left\{ -\frac{1}{3}e^4 + \dots \right\} \cos(4M) \\ &+ \dots \end{aligned}$$

$\left(\frac{r}{a}\right)^n$ se obtiene a partir del desarrollo de $\frac{r}{a}$ calculando la potencia de una serie de Poisson por medio de la correspondiente rutina creada en el PSPC.

Una función de gran utilidad definida en el PSPC nos permite efectuar la expansión binomial

$$(1 + I)^\alpha = 1 + \alpha I + \frac{\alpha(\alpha - 1)}{2 \cdot 1} I^2 + \dots + \binom{\alpha}{k} I^k \quad (2)$$

hasta el orden k , siendo I una serie de Poisson. Esta función permite calcular $\frac{a}{r}$ usando la relación

$$\frac{a}{r} = (1 - e \cos E)^{-1}$$

obteniendo de esta forma

$$\begin{aligned} \frac{a}{r} = & 1 \\ & + \left\{ e - \frac{1}{8}e^3 + \dots \right\} \cos(M) \\ & + \left\{ e^2 - \frac{1}{3}e^4 + \dots \right\} \cos(2M) \\ & + \left\{ \frac{9}{8}e^3 + \dots \right\} \cos(3M) \\ & + \left\{ \frac{4}{3}e^4 + \dots \right\} \cos(4M) \\ & + \dots \end{aligned}$$

La misma función (2) permite desarrollar $\sqrt{1 - e^2}$, y por tanto, obtener el $\text{sen } f$ a partir de la expresión

$$\text{sen } f = \frac{a}{r} \sqrt{1 - e^2} \text{sen } E$$

de esta forma

$$\begin{aligned} \text{sen } f = & \left\{ 1 - \frac{7}{8}e^2 + \frac{17}{192}e^4 + \dots \right\} \sin(M) \\ & + \left\{ e - \frac{7}{6}e^3 + \dots \right\} \sin(2M) \\ & + \left\{ \frac{9}{8}e^2 - \frac{207}{128}e^4 + \dots \right\} \sin(3M) \\ & + \left\{ \frac{4}{3}e^3 + \dots \right\} \sin(4M) \\ & + \left\{ \frac{625}{384}e^4 + \dots \right\} \sin(5M) \\ & + \dots \end{aligned}$$

Para el desarrollo del coseno de la anomalía verdadera, f , en función de la anomalía media bastará utilizar la expresión

$$\cos f = \frac{a}{r} (\cos E - e)$$

lo que nos da

$$\begin{aligned} \cos f = & -e \\ & + \left\{ 1 - \frac{9}{8}e^2 + \frac{25}{192}e^4 + \dots \right\} \cos(M) \end{aligned}$$

$$\begin{aligned}
& + \left\{ e - \frac{4}{3}e^3 + \dots \right\} \cos(2M) \\
& + \left\{ \frac{9}{8}e^2 - \frac{225}{128}e^4 + \dots \right\} \cos(3M) \\
& + \left\{ \frac{4}{3}e^3 + \dots \right\} \cos(4M) \\
& + \left\{ \frac{625}{384}e^4 + \dots \right\} \cos(5M) \\
& + \dots
\end{aligned}$$

El cálculo de $\sin f$ y $\cos f$ en función de M nos permite calcular $\sin jf$ y $\cos jf$. Para $j > 1$ para ello se utilizan las fórmulas de recurrencia

$$\sin jf = \sin[(j-1) + 1]f = \sin(j-1)f \cos f + \cos(j-1)f \sin f$$

$$\cos jf = \cos[(j-1) + 1]f = \cos(j-1)f \cos f - \sin(j-1)f \sin f$$

Para obtener los desarrollos de Hansen basta efectuar los productos de las series de Poisson correspondientes.

3. Obtención de R en el problema de tres cuerpos

Si llamamos $\mathbf{r}_1, \mathbf{r}_2$ a los vectores que definen el movimiento de P_1 respecto a P_0 y de P_2 respecto al centro de masas de P_0, P_1 , las ecuaciones del movimiento de este problema se expresarán como

$$\frac{d^2 \mathbf{r}_1}{dt^2} = \frac{\partial(U_1 + R_1)}{\partial \mathbf{r}_1}$$

$$\frac{d^2 \mathbf{r}_2}{dt^2} = \frac{\partial(U_2 + R_2)}{\partial \mathbf{r}_2}$$

siendo

$$U_1 = \frac{G(m_0 + m_1)}{r_1}$$

$$U_2 = \frac{G(m_0 + m_1 + m_2)}{r_2}$$

$$R_1 = \frac{G(m_1 + m_1)m_2}{m_0 m_1} \left(\frac{m_0}{r_{02}} - \frac{m_1}{r_{12}} \right)$$

$$R_2 = \frac{G(m_1 + m_1 + m_2)}{m_0 + m_1} \left[\frac{m_0 m_1 m_2}{r_1} + m_0 \left(\frac{1}{r_{02}} - \frac{1}{r_2} \right) + m_1 \left(\frac{1}{r_{12}} - \frac{1}{r_2} \right) \right]$$

Expresando r_{02} y r_{12} en función de r_1 y r_2 se llega a

$$R = \sum_{i=2}^{\infty} N_i \frac{a_1^i}{a_2^{i+1}} \left(\frac{r_1}{a_1} \right)^i \left(\frac{a_2}{r_2} \right)^{i+1} P_i(\sigma) \quad (3)$$

donde

$$N_i = \frac{m_0 m_1 m_2 [m_0^{i-1} - (-m_1)^{i-1}]}{(m_0 + m_1)^i}$$

siendo σ el coseno del ángulo formado por los vectores \mathbf{r}_1 y \mathbf{r}_2 , y $P_i(\cdot)$ el polinomio de Legendre de orden i .

Las funciones R_1, R_2 que corresponden a las perturbaciones de las órbitas de P_1 y P_2 respectivamente pueden ponerse como

$$R_1 = \frac{\mu_1}{m_0 m_1} \left[\frac{(m_0 + m_1) m_2}{r_2} + R \right]$$

$$R_2 = \frac{\mu_2}{(m_0 + m_1) m_2} \left[\frac{m_0 m_1 m_2^2}{r_1} + R \right]$$

El desarrollo de R puede ser usado en la integración de las 12 ecuaciones que se obtienen aplicando (1) para cada una de las órbitas.

Para integrar las ecuaciones (1) es preciso expresar R en función de los elementos orbitales, lo que puede resultar muy costoso según el orden de desarrollo al cual queremos llegar.

El primer paso es expresar σ , para lo cual debe tenerse en cuenta que

$$\sigma = \cos(\mathbf{r}_1, \mathbf{r}_2) = \frac{\mathbf{r}_1 \cdot \mathbf{r}_2}{r_1 r_2}$$

donde $\mathbf{r}_i = (x_i, y_i, z_i)$ siendo

$$x_i = r_i [\cos \Omega_i \cos(\omega_i + f_i) - \sin \Omega_i \sin(\omega_i + f_i) \cos i_i]$$

$$y_i = r_i [\sin \Omega_i \cos(\omega_i + f_i) + \cos \Omega_i \sin(\omega_i + f_i) \cos i_i]$$

$$z_i = r_i \sin(\omega_i + f_i) \sin i_i$$

que viene expresado en función de $\omega_1, \omega_2, \Omega_1, \Omega_2, i_1, i_2$ y f_1, f_2 .

Posteriormente debe calcularse $P_i(\sigma)$ lo que supone obtener las correspondientes potencias de σ .

Finalmente, debe separarse de $P_i(\sigma)$ los términos que dependen de f_i para expresarlos en función de los elementos orbitales junto con los términos $\left(\frac{r_1}{a_1}\right)^i \left(\frac{a_2}{r_2}\right)^{i+1}$, lo que conduce al desarrollo de las conocidas fórmulas de Hansen.

Para evitar un tiempo de computación excesivo, debido a la manipulación de un elevado número de términos trigonométricos, se ha optado en la práctica por la sustitución directa de los desarrollos de $\left(\frac{r_1}{a_1}\right)^i \left(\frac{a_2}{r_2}\right)^{i+1}$ por un lado, y los de $\sin^j f_i, \cos^j f_i$ por otro, tratando estos últimos como variables polinómicas en lugar de usar directamente las fórmulas de Hansen. Con este proceso, y desarrollando únicamente el primer sumando de R hasta el orden 3 en las excentricidades e_1, e_2 , se ha obtenido una expresión de 18281 términos, tras 3 horas y media de cálculos en un Macintosh SE/30, lo que da idea del volumen de trabajo de estos desarrollos.

Agradecimientos

Este trabajo ha sido financiado parcialmente por la C.I.C.Y.T. (Proyectos PB90-0921 y ESP91-0919).

Los autores están mencionados por orden alfabético.

Referencias

- [1] Abad, A., Sein-Echaluce, M.L.: 1989, Manipuladores Algebráicos en Mecánica Celeste. *Rev. Acad. Ciencias de Zaragoza*, **43**, 117-127.

- [2] Abad,A.: 1990, *Guia de Usuario del Procesador de Series de Poisson PSPC* . En preparación.
- [3] Abad, A., San Juan, J.F.:1991, Integración automatizada de las ecuaciones de Lagrange en el movimiento orbital. *Boletín del observatorio Astronómico de Madrid* Vol XII. Num 4. 37-45
- [4] Brouwer,D. and Clemence, G.M.:1961, *Celestial Mechanics*. Academic Press, New York.
- [5] Deprit, A. y Deprit, E.: 1990, Processing Poisson Series in Parallel. *J. Symbolic Computation*,10,179-201
- [6] Ling, J.F.:1989, Aplicación del método estroboscópico al problema estelar de tres cuerpos. Tesis doctoral. Universidad de Santiago de Compostela.
- [7] Fitzpatric, Ph. M.: 1970, *Principles of Celestial Mechanics*. Academic Press, New York.
- [8] Giacaglia, G. E. O. (1976): A note on Hansen's coefficients in Satellite Theory. *Celest. Mech.* 14, 515-523.

ABSTRACT

In this paper, we study the dynamics of a rigid body in a gravitational field. We first write the equations of motion in a suitable form, and then we use the method of averaging to obtain the secular motion. The results are compared with those obtained by other methods.

Now, let $O_1 = O_2$ through the origin of the body-fixed system $O_1x_1y_1z_1$ and the axis O_2z_2 belongs to a rigid body which is fixed in space. In the present case, the variables (q, \dot{q}) and their conjugate momenta (p, \dot{p}) are functions of the problem in function of the initial conditions.

1- MOTION OF A GYROSCOPE IN A GRAVITATIONAL FIELD

Basically, the motion of a gyroscope in a gravitational field is a problem of rigid body dynamics. The gyroscope is considered as a rigid body with respect to a reference frame $O_1x_1y_1z_1$ in such a way that the axis O_1z_1 of the outer gimbal is always parallel to the axis O_2z_2 , defining the angle $\psi = \psi(t)$. On the other hand, the inner gimbal (gimbal) determines the new orthogonal system $O_2x_2y_2z_2$ in such a way that the axis O_2z_2 is parallel to the axis O_1z_1 .

Finally, about of the O_2z_2 axis, a rotor that rotates with respect to the initial position a new angle ϕ .

MOTION OF A GYROSCOPE IN A SET OF REDUNDANT VARIABLES

R.CID PALACIOS and M.C.FERNANDEZ QUIÑONES

Departamentos de FÍSICA TEÓRICA (ASTRONOMÍA) y MATEMÁTICA APLICADA

Facultad de Ciencias

Universidad de Zaragoza

ABSTRACT

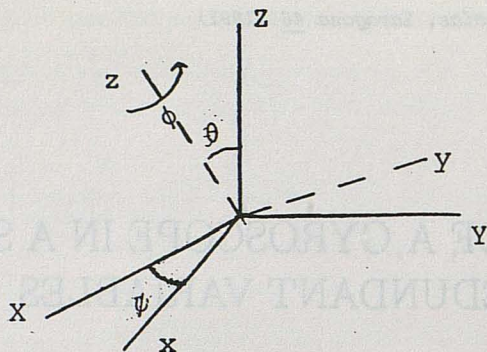
In this paper, we study the rotational motion of a gyroscope in a Cardan suspension, formulate in a set of redundant variables and their conjugate momenta.

Now, let $\theta_0 = \widehat{POZ}$ the angle that the vertical OP form with respect to the axis OZ , belonging to a right-hand rectangular trihedral system $OYYZ$ fixed in space. In the particular case $\theta_0 = 0$, we can introduce a set of two variables (q_0, q_1) and their conjugate momenta (p_0, p_1) , for the resolution of the problem in function of first integrals in involution.

1.- MOTION OF A GYROSCOPE IN A CARDAN SUSPENSION

Schematically, the motion of a gyroscope in a Cardan suspension is defined (E, Leimanis, 1965) with respect to a ortogonal fixed system $OXYZ$, in such a way that the axis Ox of the outer gimbal turn out around of the axis OZ , defining the angle $\psi = \widehat{XOx}$. On the axis Ox a second ring (inner gimbal) determine the new orthogonal system $Oxyz$, in such a way that $\theta = \widehat{ZOz}$.

Finally, about of the Oz axis turn a rotor that define with respect to an initial position a new angle ϕ



In these conditions, the kinetic energy of revolution T_3 , of the outer ring will be:

$$T_3 = \frac{1}{2} C_3 \dot{\psi}^2 \quad (1)$$

where C_3 is the moment of inertia about OZ .

Furthermore, given that the components of the angular velocity of the interior ring are:

$$\omega_1 = \dot{\theta}, \quad \omega_2 = \dot{\psi} \sin \theta, \quad \omega_3 = \dot{\psi} \cos \theta \quad (2)$$

The kinetic energy of revolution T_2 , adopts the form

$$T_2 = \frac{1}{2} (A_2 \dot{\theta}^2 + B_2 \dot{\psi}^2 \sin^2 \theta + C_2 \dot{\psi}^2 \cos^2 \theta) \quad (3)$$

where A_2, B_2, C_2 , are the corresponding moments of inertia. Analogously, given that the components of the angular velocity of the rotor are:

$$\omega_1 = \dot{\theta}, \quad \omega_2 = \dot{\psi} \sin \theta, \quad \omega_3 = \dot{\phi} + \dot{\psi} \cos \theta \quad (4)$$

the kinetic energy of revolution T_1 , will be:

$$T_1 = \frac{A_1}{2} (\dot{\theta}^2 + \dot{\psi}^2 \sin^2 \theta) + \frac{C_1}{2} (\dot{\phi} + \dot{\psi} \cos \theta)^2 \quad (5)$$

Consequently, the total kinetic energy of the system will be:

$$T = \frac{1}{2} [A \dot{\theta}^2 + (B - C \cos^2 \theta) \dot{\psi}^2 + C_1 (\dot{\phi} + \dot{\psi} \cos \theta)^2] \quad (6)$$

where

$$A = A_1 + A_2, \quad B = A_1 + B_2 + C_3, \quad C = A_1 + B_2 - C_2 \quad (7)$$

In order to calculate the potential energy V , we can suppose, without loss of generality, that the vertical OP lies in the $O\dot{X}Z$ - plane. Therefore, being $\theta_0 = \widehat{POZ}$, we obtain

$$V = mgs(\cos\theta_0 \cos\theta + \sin\theta_0 \sin\theta \sin\psi) \quad (8)$$

where, m is the mass of the rotor and the inner gimbal ring, s the distance from the center of mass to the origin O , and g the gravity acceleration.

It is known that the motion of a gyroscope can be expressed in Hamiltonian form by means of the Euler angles (ψ, θ, ϕ) and their conjugate momenta $(p_\psi, p_\theta, p_\phi)$. In this case, the hamiltonian H can be written

$$H = \frac{1}{2} \left[\frac{p_\theta^2}{A} + \frac{(p_\psi - p_\phi \cos\theta)^2}{B - C \cos^2\theta} + \frac{p_\phi^2}{C_1} \right] + V \quad (9)$$

and the problem has always two evident integrals

$$H = h = cte, \quad p_\phi = k_3 = cte \quad (10)$$

In the particular case $\theta_0 = 0$, the problem admit a third integral

$$p_\psi = k_1 = cte \quad (11)$$

since, now, we have

$$V = mgs \cos\theta = k \cos\theta \quad (12)$$

2.- MOTION OF A GYROSCOPE IN A SET OF REDUNDANT VARIABLES

R.Cid and M.E. San Saturio (1988) showed a proposition that we can resume in the following form:

"A canonical system

$$\dot{q}_i = \partial H / \partial p_i; \quad \dot{p}_i = -\partial H / \partial q_i \quad (i = 1, 2, \dots, n) \quad (13)$$

with a Hamiltonian $H(q, p, t)$ of class $C^{(2)}$, can be transformed in a new canonical system

$$\dot{Q}_\alpha = \partial H^* / \partial P_\alpha \quad \dot{P}_\alpha = -\partial H^* / \partial Q_\alpha \quad (\alpha = 1, 2, \dots, n+1) \quad (14)$$

of Hamiltonian $H^* = H[q(Q), p(Q, P), t]$, by means of the transformation

$$q_i = q_i(Q_\alpha), \quad P_\alpha = \sum_i p_i \left(\frac{\partial q_i}{\partial Q_\alpha} \right) \quad (15)$$

of class $C^{(2)}$, that verifies the condition

$$\det M^* = \begin{vmatrix} P_1 & \frac{\partial q_1}{\partial Q_1} & \dots & \frac{\partial q_n}{\partial Q_1} \\ P_2 & \frac{\partial q_1}{\partial Q_2} & \dots & \frac{\partial q_n}{\partial Q_2} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n+1} & \frac{\partial q_1}{\partial Q_{n+1}} & \dots & \frac{\partial q_n}{\partial Q_{n+1}} \end{vmatrix} = 0$$

In this case, every solution $[Q_\alpha(t), P_\alpha(t)]$ of the system (14), determines a solution $[q_i(t), p_i(t)]$ of the system (13)."

According to this proposition, the motion of a gyroscope in the general case ($\theta_0 \neq 0$) can be studied in a set of four variables (q_0, q_1, q_2, q_3) defined in the form (R.Cid and M.C. Fernandez Qnes, 1991)

$$q_0 = \cos \frac{\theta}{2} \cos \frac{\psi + \phi}{2} \quad q_1 = \sin \frac{\theta}{2} \cos \frac{\psi - \phi}{2} \quad (16)$$

$$q_2 = \sin \frac{\theta}{2} \sin \frac{\psi - \phi}{2} \quad q_3 = \cos \frac{\theta}{2} \sin \frac{\psi + \phi}{2} \quad (17)$$

while, in the particular case ($\theta_0 = 0$), it is sufficient to introduce two variables (q_0, q_1), since in this case exit three integrals

$$H = h, \quad p_\psi = k_1, \quad p_\phi = k_3 \quad (18)$$

3.- STATEMENT OF THE PROBLEM IN THE PARTICULAR CASE.

The reduction of the problem to quadratures, in the particular case ($\theta_0 = 0$), can now be accomplished as follows:

Putting

$$\sin \theta = q_0 \quad \cos \theta = q_1 \quad (19)$$

or else

$$\theta = \arctan(q_0/q_1) \quad q_0^2 + q_1^2 = 1 \quad (20)$$

the transformation $(\theta, p_\theta) \rightarrow (q_0, q_1, p_0, p_1)$ defined by the equalities (19) and the conditions

$$p_0 = p_\theta \frac{\partial \theta}{\partial q_0} = p_\theta \frac{q_1}{q_0^2 + q_1^2} = p_\theta q_1 \quad (21)$$

$$p_1 = p_\theta \frac{\partial \theta}{\partial q_1} = -p_\theta \frac{q_0}{q_0^2 + q_1^2} = -p_\theta q_0 \quad (22)$$

is a canonical transformation by virtue of the proposition established in paragraph 2.

According to this conditions, we shall write the relations

$$q_0^2 + q_1^2 = 1 \quad (23)$$

$$p_0^2 + p_1^2 = p_\theta^2 \quad (24)$$

$$q_0 p_0 + q_1 p_1 = 0 \quad (25)$$

Hence, in this case, the motion of the gyroscope is represented by the Hamiltonian

$$H = \frac{1}{2} \left[\frac{(p_0^2 + p_1^2)}{A} + \frac{(k_1 - k_3 q_1)^2}{B - C q_1^2} + \frac{k_3^2}{C_1} \right] + k q_1 \quad (26)$$

and the equations

$$\dot{q}_0 = \frac{\partial H}{\partial p_0} = \frac{p_0}{A} \quad \dot{p}_0 = -\frac{\partial H}{\partial q_0} \quad (27)$$

$$\dot{q}_1 = \frac{\partial H}{\partial p_1} = \frac{p_1}{A} \quad \dot{p}_1 = -\frac{\partial H}{\partial q_1} \quad (28)$$

$$(29)$$

On the other hand, the derivative of the equality (23) with respect to the time t assume the form

$$\frac{d}{dt}(q_0^2 + q_1^2) = 2(q_0 \dot{q}_0 + q_1 \dot{q}_1) = \frac{2a^2}{A}(q_0 p_0 + q_1 p_1) = 0 \quad (30)$$

that proves the integral character of the equation (23), so as the disappearance of the variable q_0 in the Hamiltonian (26).

In these conditions, we have

$$\dot{p}_0 = 0 \quad p_0 = k_0 = cte \quad (31)$$

$$\dot{p}_1 = \frac{(k_1 - k_3 q_1)}{(B - C q_1^2)^2} [k_3(B - C q_1^2) - C q_1(k_1 - k_3 q_1)] - k \quad (32)$$

It is obvious that the relation (25) is not a first integral of the system, since in general we have

$$\frac{d}{dt}(q_0 p_0 + q_1 p_1) \neq 0 \quad (33)$$

Finally, substituting the moment $p_1 = A \dot{q}_1$ in the Hamiltonian (26) we obtain the solution of the system by means of the hyperelliptic integral

$$dt = \sqrt{A} \sqrt{\frac{B - C q_1^2}{Q_0 + Q_1 q_1 + Q_2 q_1^2 + Q_3 q_1^3}} dq_1 \quad (34)$$

where the constants Q_i , are given by the equalities

$$Q_0 = B(2h - \frac{k_0^2}{A} - \frac{k_3^2}{C_1}) - k_1^2 \quad (35)$$

$$Q_1 = 2(k_1 k_3 - Bk) \quad (36)$$

$$Q_2 = -C(2h - \frac{k_0^2}{A} - \frac{k_3^2}{C_1}) - k_3^2 \quad (37)$$

$$Q_3 = 2Ck \quad (38)$$

Let us now see the involutory character of the integrals

$$H = h, \quad p_0 = k_0, \quad q_0^2 + q_1^2 = 1 \quad (39)$$

According to Whittaker(1988): If n distinct integrals $\phi_i(q, p, t) = k_i = cte$ of a canonical system $\dot{q}_i = \partial H / \partial p_i$, $\dot{p}_i = -\partial H / \partial q_i$, ($i = 1, 2, \dots, n$) are in involution, i.e. $\{\phi_\alpha, \phi_\beta\} = 0$, then solving these integrals in the form $p_i = f_i(q, k, t)$ and substituting the functions f_i in the expression $\sum p_i dq_i - H dt$ a perfect differential $dV(q, k, t)$ is obtained. Thus in this case the remaining integrals of the system are $\partial V / \partial k_i = h_i = cte$.

In other words, the knowledge of n integrals in involution of a dynamical system $\dot{q}_i = \partial H / \partial p_i$, $\dot{p}_i = -\partial H / \partial q_i$ ($i = 1, 2, \dots, n$) is equivalent to its total integration.

In our case, we have

$$\{H, p_0\} = \frac{dp_0}{dt} = 0 \quad (40)$$

$$\{H, q_0^2 + q_1^2\} = \frac{d}{dt}(q_0^2 + q_1^2) = 0 \quad (41)$$

$$\{p_0, q_0^2 + q_1^2\} = -2q_0 \neq 0 \quad (42)$$

$$(43)$$

So, we have two integrals in involution (H, p_0) , or $(H, q_0^2 + q_1^2)$ that proves the integrability of the system.

References

- [1] CID, R. and FERNANDEZ QUIÑONES, M.C. (1991): On the problem of movement of a gyroscope in Cardan suspension, studied in a set of redundant variables. *Spain-Urss Workshop on position Astronomy and Celestial Mechanics. Valencia*
- [2] CID, R. and SAN SATURIO, M.E. (1988): Motion of rigid bodies in a set of redundant variables. *Celes. Mech.*, 42, 263-277. U.S.A.
- [3] LEIMANIS, E. (1965): The General Problem of the motion of Coupled Rigid Bodies about a Fixed Point. *Springer-Verlag, Berlin, Heidelberg, New York.*
- [4] WHITTAKER, E.T. (1988): A treatise on the analytical dynamics of particles and rigid bodies. *Cambridge University Press. Cambridge. New York.*

GLACIACIONES Y ASTRONOMIA

A. Elipe

Grupo de Mecánica Espacial. Universidad de Zaragoza

M. A. Soriano

Dpto. de Geología. Universidad de Zaragoza

Abstract

In this paper, the astronomical theory of the Ice Ages is exposed. Presently, there seems to be a general consensus that the variations in the motion of the Earth on climate is real enough, an inter-disciplinary effort is needed for a better modeling of the climate in past epochs. In this note, we present an exposition of the discovery of the Ice Ages, the different theories for its explanation, the geological test, and the keys for further development in research for obtaining a deeper knowledge of the climate in the past and in the future.

1. Descubrimiento de las Glaciaciones

La teoría, hoy en día, universalmente aceptada de que gran parte de los continentes estuvieron cubiertos por capas de hielo de varios centenares de metros de espesor, es reciente y triunfó tras una serie de apasionadas batallas científicas entre los geólogos del pasado siglo. En efecto, de un modo escueto, podemos decir que dicha teoría fue iniciada por Louis Agassiz al lanzar esta idea en una Reunión de la Sociedad Suiza de Ciencias Naturales en Newchâtel el 26 de Julio de 1837. El origen de tan atrevida opinión surgía principalmente de la localización de enormes bloques de roca (boulders), distribuidos aleatoriamente por terrenos muy alejados de nieves perpetuas y glaciares, con profundas incisiones y compuestos por material que no se encontraba presente en bastantes kilómetros a la redonda. Si bien la existencia de estos bloques era conocida, nadie había osado formular en un contexto científico que su origen fuese glaciar, por el contrario, la idea más generalmente extendida era que procedían de alguna gran inundación en épocas remotas, sin desdeñar el *Diluvio Universal* (la teoría catastrofista era aceptada por un elevado número de naturalistas). Lyell [18], uno de los representantes del actualismo, junto con otros científicos sugirieron que los bloques habían llegado por el efecto de grades olas marinas, o habían sido desplazados por icebergs, aunque no fueron capaces de explicar el que aparecieran en lugares como Suiza, y en sitios especialmente elevados.

El impulso definitivo de esta teoría se produjo con la conversión a ella de geólogos de gran prestigio, como Buckland (Oxford) y el ya citado Lyell, y con la aparición de evidencias similares en Norte-América, como por ejemplo, una morrena terminal, llegando a alcanzar hasta 40 m, de altitud, extendiéndose desde Long-Island hasta el estado de Washington, aunque originada por varios focos de expansión. En general, podemos decir que esta idea fue defendida inicialmente

por científicos de Escocia, Suiza, Escandinavia, y que fue rechazada por habitantes de zonas llanas y costeras. A este respecto, conviene recordar que en aquellos momentos existía un gran desconocimiento sobre glaciares, que fue en 1852 cuando se comprobó que Groenlandia estaba cubierta por una gran capa de hielo, y que fue a finales del siglo pasado cuando se descubrió la verdadera dimensión del continente Antártico. Tras 25 años de continuo debate, por fin fue aceptada la existencia de una Glaciación (Ice Age).

Una cuestión que surgió desde el momento en que Agassiz expuso su teoría fue que, caso de ser cierta ésta, para que la capa de hielo cubriese parte de Suiza, debería tener un espesor de 1 Km., y como la única fuente suministradora de agua era el océano, el nivel de las aguas debería bajar unos 30 m, con la consiguiente modificación de la línea de costa. Sin embargo, aunque tal cantidad de agua efectivamente fue sustraída del océano, registros geológicos afirman que no se produjo una variación de tal magnitud en las costas, debido a que por el peso de la capa de hielo, la corteza terrestre se hundió. Por otra parte, cuando se produjo el deshielo, hubo dos tipos de respuesta: una inmediata, que consistió en un aumento del nivel del mar, y otra gradual, que produce una lenta elevación del terreno, fenómeno que todavía puede detectarse. Por ejemplo, en los Grandes Lagos de Norte-América, el suelo se eleva unos 40 cms. por siglo.

Hacia 1875, se puede decir que se había cartografiado completamente la glaciación [11], al mismo tiempo que se tenía evidencia de la existencia de diversas glaciaciones anteriores. Sin embargo, se abrió un amplio espectro de interrogantes, acerca del agente que las causó, porqué cesaron y sobre si volverían a producirse. Entre las diversas teorías que intentan dar solución a estas preguntas, es la Astronomía dinámica la que parece más plausible, aunque lógicamente, tenga sus detractores.

2. Historia de las distintas teorías astronómicas

El pionero en dar una interpretación astronómica de las glaciaciones fue Adhémar [1], quien reflexiona sobre el hecho de que la órbita de la Tierra es elíptica, y que el paso por el perihelio se produce sobre el 3 de Enero, próximo al solsticio de invierno; por este motivo, en el hemisferio Norte, la primavera y verano tienen 7 días más que el otoño y el invierno, sucediendo lo contrario en el hemisferio Sur, por lo que en el hemisferio austral el número de horas de oscuridad supera al de luminosidad, lo que hace que se vaya enfriando cada vez más. Esa sería la causa de una glaciación en el hemisferio sur. Además, conocido ya por D' Alembert, el movimiento de la precesión de los nodos junto con el movimiento del perihelio en la órbita forman un ciclo de 22.000 años, por lo que cada 11.000 años debería haber una glaciación, una correspondiente al hemisferio norte y otra al hemisferio sur.

Por desgracia, junto a su lógico razonamiento introdujo unas consecuencias que hicieron que su teoría perdiese credibilidad. En efecto, supuso que la capa de hielo del Antártico absorbería el agua del hemisferio Norte, originando un gran abultamiento en el antártico. Posteriormente, cuando el agua se fuese deshelando, el fenómeno resultante sería una especie de hongo, hasta que éste colapsase produciendo una gran ola hacia el hemisferio norte. Aparte de esta fantasía, el

método matemático fue criticado por algunos científicos, entre ellos von Humboldt, quien señaló que la idea principal (un hemisferio se calentaba mientras el otro se enfriaba) era incorrecta. En efecto, ya se conocía desde D' Alembert, que la cantidad total de calor recibido por un hemisferio durante un año es exactamente igual al recibido por el otro. La verdadera razón por la que el hemisferio Sur se encuentra más frío se descubrió varios años después y es debido por una parte, a la ausencia de las corrientes marinas que trasladan agua caliente y por otro, a que al reflejarse la luz en las grandes capas de nieve y hielo aumenta el frío. Aunque falsa, constituyó un gran e importante paso en el camino para comprender el misterio de las glaciaciones.

Motivado por la obra de Adhémar, Croll decidió seguir en esta línea. Croll conocía los recientes estudios de Leverrier (descubridor de Neptuno), donde se probaba que la excentricidad de la órbita terrestre poseía un periodo muy grande y pensó que esta lenta variación en la excentricidad podría ser una de las causas de las glaciaciones. Mediante las fórmulas de Leverrier, Croll calculó la excentricidad orbital para un periodo de 3.000.000 años, y las representó gráficamente, siendo el primer geólogo en estudiar la historia de la órbita de la Tierra. Un primer análisis indicó que la excentricidad sufría grandes cambios en magnitud, así, mientras que ahora es muy pequeña, hace 100.000 años era muy elevada, por lo que tendría que ser algún factor decisivo en la génesis de las glaciaciones.

Su hipótesis fue que una disminución en la cantidad de luz recibida durante el invierno favorece la acumulación de nieve, y además, cualquier aumento en la superficie cubierta por la nieve, resultará en una mayor pérdida de calor al reflejarse la luz, por lo cual, pequeños cambios en la radiación solar serían ampliados por los propios campos de nieve. Así pues, su principal tarea consistía en determinar qué factores astronómicos controlan la cantidad de luz solar recibida durante el invierno. El principal que encontró, fue el de la precesión de los equinoccios: si el invierno sucede cuando la tierra está cerca del Sol, los inviernos serán más templados que usualmente y recíprocamente. Por razones diferentes, llegó al mismo ciclo de 11.000 años de Adhémar, pero teniendo en cuenta la excentricidad orbital. Si la órbita fuese circular, la precesión no afectaría al clima, puesto que las estaciones sucederían a la misma distancia del Sol. Hoy en día la excentricidad es muy pequeña ($1/100$), por lo que apenas tiene efecto cualquiera que sea la posición en la órbita en que suceda el invierno. Sin embargo, en épocas de excentricidad mayor ($6/100$), inviernos extraordinariamente templados tienen lugar cuando el solsticio se produce cerca del perihelio, e inviernos extraordinariamente fríos en las proximidades del afelio. Así pues, de acuerdo con Croll, un hemisferio sufre una glaciación cuando dos condiciones se verifican simultáneamente: una órbita muy excéntrica y el solsticio de invierno sucede muy lejos del Sol. Consecuentemente, a las épocas de excentricidad muy elevada las denominó *Epocas Glaciares* y los intervalos entre ellas *Epocas Interglaciares*. La última época glacial comenzó aproximadamente hace 250.000 años y acabó hace unos 80.000 años.

Sin embargo, Croll estaba realmente preocupado con la pregunta de si tan pequeñas oscilaciones en la excentricidad, son capaces de explicar los registros geológicos, pues las masas de hielo deberían cubrir gran parte de Europa y Norte-América. Para explicar esto, sugirió que los

factores astronómicos serían el detonante que ayudaría a las corrientes templadas del Océano Atlántico, de tal modo, que cuando en un hemisferio se produce una glaciación, los vientos se moverían en una dirección tal que contribuirían a enfriarlo más todavía. Posteriormente, Croll añadió a su teoría la influencia de la variación de la oblicuidad de la eclíptica; sin embargo, al no disponer de una buena teoría de ésta, no pudo aportar conclusiones definitivas. En 1875 publicó *Climate and Time* [8], compendio de su teoría.

El gran impulsor de la teoría astronómica de las glaciaciones y cuyo nombre va necesariamente ligado a ésta es Milutin Milankovitch, Profesor de Matemática aplicada en la Universidad de Belgrado, donde explicaba Mecánica y Astronomía. Su primer propósito fué el desarrollar una teoría matemática capaz de describir los climas de la Tierra, Marte y Venus, tanto de hoy como del pasado. Para ello procedió a calcular la temperatura en los distintos paralelos de la Tierra. Así, de un modo teórico podría determinar la temperatura en el resto de los planetas, pues en todos los casos la fuente calorífica era común. Además, si fuera posible calcular el clima hoy en día, también lo sería el calcularlo en otras épocas más remotas, cuando las condiciones orbitales y geométricas de la Tierra fueran diferentes. Es decir, *obtendría información viajando en distancia y en el tiempo*. El desarrollar esta teoría le llevaría más de 30 años.

Al igual que Croll, encontró que la distribución de la radiación solar sobre las superficies de los planetas estaba determinada por tres parámetros: *la excentricidad de la órbita, la inclinación del eje de rotación y la posición de los equinoccios en su ciclo precesional*. Sin embargo, contaba con un nuevo factor a su favor: una sustancial mejora en los desarrollos matemáticos de Leverrier, llevados a cabo por el matemático alemán Pilgrim, lo que permitía extender los anteriores parámetros al último millón de años.

Su segundo objetivo era calcular la cantidad de radiación solar que llega a cada planeta durante una estación en un paralelo de latitud dada. Era conocido a partir de Newton que la radiación incidente del Sol dependía de dos factores geométricos, la distancia al Sol, y el ángulo con el que los rayos inciden en la superficie. Demostró que las variaciones en la órbita y excentricidad eran lo suficientemente elevadas como para originar contracciones y extensiones de las capas de hielo y que las oscilaciones en la oblicuidad de la eclíptica eran más importantes de lo que Croll había estipulado. Su teoría matemática del clima terrestre actual, de Marte y Venus [20], fue considerada como una de las principales contribuciones para la descripción del clima moderno.

Esta obra causó gran impacto en Wladimir Köppen, reconocido como uno de los mayores climatólogos del mundo, que junto con su yerno Alfred Wegener, geólogo, se pusieron en contacto con Milankovitch, resultando estas relaciones especialmente fructíferas para éste. Una vez disponible el mecanismo necesario para calcular la radiación solar en cualquier latitud y estación, el siguiente objetivo era proporcionar una descripción matemática del pasado climatológico de la Tierra. Para ello, procedió a representar gráficamente las curvas de radiación para ciertas latitudes, pero la cuestión era determinar cual de éstas era crucial para el desarrollo de las capas de hielo. La hipótesis principal tanto de Adhémar como de Croll era que el factor crítico había sido la radiación recibida en altas latitudes durante el invierno. Sin embargo, Mi-

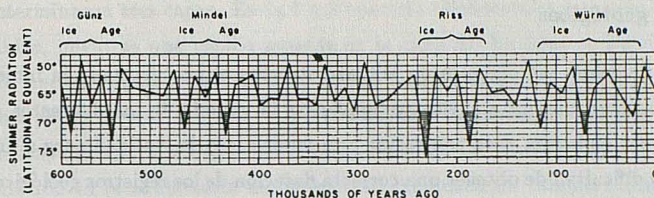


Figura 1.—Curva de radiación de Milankovitch para una latitud de 65° Norte. Adaptado de Köppen y Wegener (1924)

lankovitch dudaba de que ésta fuese la causa real y Köppen le sugirió que el factor decisivo sería **la disminución del calor durante el verano**. En efecto, según Köppen, los cambios en la radiación en invierno apenas tendrían efecto en la cantidad de nieve, pues las temperaturas son lo suficientemente bajas como para que se acumule. Sin embargo, durante el verano los glaciares se funden y por tanto, cualquier disminución en la intensidad de la luz del sol recibida impedirá la fusión, haciendo que la cantidad de nieve aumente y comience la expansión del glaciar.

Siguiendo este razonamiento, Milankovitch comenzó a calcular las curvas de la radiación recibida durante el verano a latitudes 55°, 60° y 65° durante los pasados 650.000 años (Figura 1). Estas gráficas, tal como Köppen comprobó, se ajustaban bastante bien a la historia de los glaciares Alpinos. Publicó sus resultados como un capítulo del libro de Köppen y Wegener [15] "Climas del pasado geológico", lo que aseguró una amplia divulgación de su teoría entre geólogos y meteorólogos. Posteriormente extendió sus tablas a ocho latitudes, situadas entre 5° y 75°N, publicando un volumen titulado "Mathematical Climatology and the Astronomical Theory of Climatic Changes". Con estas curvas de radiación, quedaba clara por primera vez la contribución de los ciclos astronómicos en la radiación solar. La influencia de la oscilación de la oblicuidad (ciclo de 41.000 años) es mayor en los polos y menor en el ecuador, pero por el contrario, la del ciclo de precesión (ciclo de 22.000 años) es menor en los polos y mayor en el ecuador.

Con esto, comenzó a preparar su cuarto y último objetivo, calcular de qué modo las capas de hielo podían responder a un cierto cambio en la radiación solar. Para ello, la principal dificultad estribaba en determinar la importancia de la disminución de la temperatura por reflexión de la luz, tal como ya había sugerido Croll. Milankovitch resolvió el problema estudiando la altitud a la cual existe nieve perpetua. Esta difiere con la latitud, así en la zona ecuatorial está en montañas muy elevadas, mientras que en las zonas polares está a nivel del mar. Milankovitch encontró una formulación matemática de la relación entre la radiación recibida durante el verano, y la altitud de la línea de nieve, publicando sus resultados en 1938. Con todo lo cual, quedaba completo el trabajo [21] y se podía calcular la latitud de las capas de hielo en cualquier época dentro de los últimos 650.000 años.

3. Registros geológicos

La aparición de la teoría astronómica de Milankovitch fue objeto de atención inmediata, pues se ofrecía por primera vez un modelo que explicaba la génesis de las diferentes glaciaciones que se intuían. Sin embargo, para contrastar este modelo con la realidad surgía un obstáculo, consistente en la dificultad de obtener una correcta datación de los registros geológicos, siendo el método más frecuente el de asignar edades por el grosor de las capas sedimentarias. No obstante, a finales del siglo pasado, se llegó a afirmar la existencia de cuatro glaciaciones, separadas por un periodo interglaciario similar al que ahora disfrutamos. En las curvas de Milankovitch, se observaban mínimos en 25.000, 72.000 y 115.000 años, con lo que algunos geólogos veían confirmada la teoría astronómica, pero como las edades de los registros se conocían vagamente, otros se mostraron en contra, sin llegar a ningún acuerdo.

Un método diferente fue empleado por Albrecht Penck [22] estudiando los valles alpinos. En efecto, encontró en todos ellos tres terrazas separadas por profundos escarpes, lo que le sugirió que cada uno de éstos se formó en épocas frías, al comienzo de las épocas interglaciares. Como encontró cuatro capas de gravas, dedujo que se habían producido cuatro glaciaciones en el Pleistoceno, que denominó por orden de antigüedad: *Günz*, *Mindel*, *Riss* y *Würm*. Analizando el espesor de los sedimentos postglaciares en los lagos alpinos y la velocidad con que se depositan, estimó que la última se había producido hace unos 20.000 años, la anterior hace 60.000, la previa a ésta, que llamó *Gran interglaciario*, había durado 240.000 años y estimó una duración del Pleistoceno de 650.000 años.

Estos periodos coincidían con la teoría que Milankovitch obtuvo de modo independiente años después, lo que hizo que las curvas de radiación solar de éste se empleasen durante varios años como método de datación. Sin embargo, hacia 1955 esta técnica fue abandonada al aparecer otra más precisa, empleada en la datación de fósiles del Pleistoceno: el carbono radiactivo. Desarrollado por Libby [17], solamente era válido para un periodo de tiempo no superior a 40.000 años. Con el radiocarbono, se llegó a la conclusión de que la última glaciación alcanzó su máximo hace 18.000 años, desapareciendo rápidamente hace 10.000 años. Además, se encontraron depósitos correspondientes a épocas templadas de 25.000 años de edad, justamente en periodos en que Milankovitch predecía una glaciación. Por otra parte, extendiendo el periodo de validez de los 40.000 años del carbono radiactivo, aparecieron más épocas glaciares de las predichas por la teoría astronómica.

Sin embargo, todavía persistía el problema de que solamente habían sido considerados registros tomados en la superficie terrestre, por lo que se imponía el análisis de muestras del fondo marino. En 1875, una corbeta inglesa tomó muestras del fondo marino en una expedición alrededor del mundo, con el resultado de que en aguas muy profundas (> 4.000 m) no existen registros orgánicos pues sus esqueletos son disueltos, mientras que a profundidades menores de 4.000 m había abundancia de foraminíferos en aguas calientes, y en aguas frías se detectaban restos de radiolarios y diatomeas. Sin embargo, para obtener mayor información era preciso realizar perfiles de los sedimentos marinos mediante testigos y los medios técnicos necesarios para conseguirlos no fueron suficientemente buenos hasta los años 20 de este siglo. De este modo, con cortes de 40

cm., se determinaron tres capas. En la 1 y 3 aparecía *Globorotalia menardii*, correspondiente a aguas cálidas, mientras que estaba ausente en la capa 2. En 1947 se logró un perfil de 15 m. En estos nuevos registros se descubrió que la concentración de carbonato cálcico variaba drásticamente en las distintas capas. Una circulación de agua distinta en el Pacífico, según se estuviese en etapa glacial o interglacial, explicaría este hecho. Esto, junto con la convicción de que una variación en la abundancia de *Globorotalia* indicaba un cambio climático, le permitió a Ericson [10] determinar la existencia de 9 glaciaciones durante el Pleistoceno.

Posteriormente se utiliza un nuevo método de datación basado en la composición isotópica de los átomos de oxígeno en foraminíferos fósiles. Urey, había probado que la concentración de O^{18} depende de la temperatura del agua. En aguas frías, la concentración de O^{18} es mayor que la de O^{16} , con lo que sería posible calcular la temperatura del agua mediante la relación en las concentraciones O^{18}/O^{16} en los fósiles. Sin embargo, la temperatura también depende de la composición isotópica del agua. Emiliani [9], tras el análisis de varias muestras marinas concluyó que durante los últimos 300.000 años había habido 7 glaciaciones, y que durante éstas, las aguas del Caribe descendían unos $6^{\circ} C$. Las curvas que obtuvo, venían a apoyar la teoría astronómica de Milankovitch.

Ante la discrepancia entre los resultados obtenidos por Ericson y Emiliani, Imbrie sugiere la posibilidad de que otros factores diferentes a la temperatura podían influir en el contenido de foraminíferos, por ejemplo, la concentración de sal o la cantidad de alimento disponible para los foraminíferos. Por una parte, Imbrie y Kipp [13], y por otra, Broecker y Jan van Donk [5] analizan independientemente una misma muestra, probando que Ericson estaba equivocado, y que Emiliani decía solamente una verdad a medias, pues donde Emiliani deducía temperaturas frías, ambos grupos encontraron aguas templadas. La temperatura disminuía solamente $2^{\circ} C$ y no los 6° de Emiliani, por lo que la mayor parte de la variación isotópica era debida a cambios en el volumen de las capas de hielo, no a cambios en la temperatura.

Un fenómeno que contribuyó en el proceso de datación, fue el de la inversión magnética. Descubierta por Brunhes (1906), Matuyama también detectó que durante el Pleistoceno, al menos una vez, se había invertido el campo magnético terrestre. Posteriormente, Cox et al.[7] lo confirman, y denominan al último periodo de polaridad normal *Epoca de Brunhes*, y a la previa (de polaridad invertida) *Epoca de Matuyama*. Mediante la técnica de K-Ar se estimó con gran precisión su duración en 700.000 años

En 1970 surge el proyecto internacional CLIMAP, en el que participaron más de 100 científicos de Dinamarca, Francia, Alemania, Inglaterra, Noruega, Suiza y Holanda con dos propósitos fundamentales:

- Elaborar un mapa de la superficie de la Tierra durante la última glaciación
- Medir las oscilaciones del clima durante el Pleistoceno.

En 1976 el grupo [6] publicó un mapa global mostrando las temperaturas del Océano y la distribución de glaciares a la altura de la última glaciación, hace 18.000 años (Figura 2). Para ello, hubo que dividir la Epoca de Brunhes (700.000 años) de la última inversión magnética



Figura 2.—Temperaturas en el nivel del mar, extensión de las capas de hielo, su elevación y albedo hace 18.000 años. Adaptado de CLIMAP (1976)

en zonas estratigráficas, es decir, en estratos capaz de reconocerse y correlacionarse entre los diferentes testigos de sondeos. Con ello pudieron identificarse distorsiones locales, tales como ausencia de erosión, corrientes de turbidez, etc. Esta tarea, ya comenzada por Ericson, tenía que ser válida para todo el Océano.

De notable importancia fue el estudio de una muestra (V28-238) obtenida en aguas poco profundas del Océano Pacífico. Correspondía a un periodo muy largo de tiempo, con lo que también contenía evidencias de la última inversión magnética (Matuyama) y con métodos especialmente diseñados para este caso se consiguieron dos curvas isotópicas [25], que han sido consideradas como la *Piedra Rosetta* de las glaciaciones [14].

La primera (Figura 3) reflejaba la variación en la composición isotópica de conchas de plancton cerca de la superficie del agua. Los primeros 17 estados correspondían exactamente a los dados por Emiliani para el Caribe y todavía aparecían dos más hasta la base de la Epoca de Brunhes. Faltaba por ver que estas fluctuaciones eran globalmente sincronas. Esto se comprobó con su segunda curva, correspondiente a los foraminíferos del fondo del mar, pues coincidía con los del plancton.

Como la temperatura en el fondo del Océano siempre está muy fría, próxima a helarse, no podría haberse enfriado demasiado en la glaciación, por lo que ambas curvas señalaban cambios en la composición de isótopos ligeros en el Océano y no variaciones en la temperatura del agua; como ésta, mediante las corrientes, se mezcla rápidamente, cualquier cambio químico en una parte se reflejaría en cualquier lugar durante unos 1000 años. Cuando los glaciares se extendieron, isótopos ligeros del Oxígeno pasaron del mar a las capas de hielo alterando la proporción isotópica del oxígeno en el agua del mar. Cuando se fundió el hielo, los isótopos volvieron al mar, recuperando éste su composición inicial. El efecto en la temperatura fue demasiado pequeño como para que pudiese ser detectado.

Los límites de los 19 estados isotópicos estaban muy bien definidos, al principio por medio de la datación por radiocarbono y al final mediante la inversión magnética. Interpolando se pudo fijar las épocas de estos estados, lo que permitió una datación muy precisa de los fenómenos del Pleistoceno. Así pudo observarse el ciclo de 100.000 años y también oscilaciones de periodos más cortos, de 44.000 y de 22.000 años, pero estos últimos muy amortiguados. La explicación a este amortiguamiento fue sugerido por Hays, quien pensó que cuando la velocidad de acumulación es menor de 2 mm./siglo, las actividades de los animales que viven en el fondo marino pueden enmascarar la frecuencias altas. Mediante el análisis de una nueva muestra de precipitación rápida (3 mm./siglo) se observó un ciclo largo (100.000 años) y sobrepuestos tres más cortos, uno de 41.000 años, correspondiente a la oblicuidad de la eclíptica, y otros dos de 23.000 y 19.000 años respectivamente, debidos a la precesión de los equinoccios. Resultados análogos se obtuvieron con dos métodos distintos mediante el análisis espectral isotópico y el de la curva radiolarios-temperatura, con lo cual, Hays, Imbrie y Shackleton concluyeron [12, p. 1131]: *It is concluded that changes in the earth's orbital geometry are the fundamental cause of the sucesion of Quaternary ice ages .*

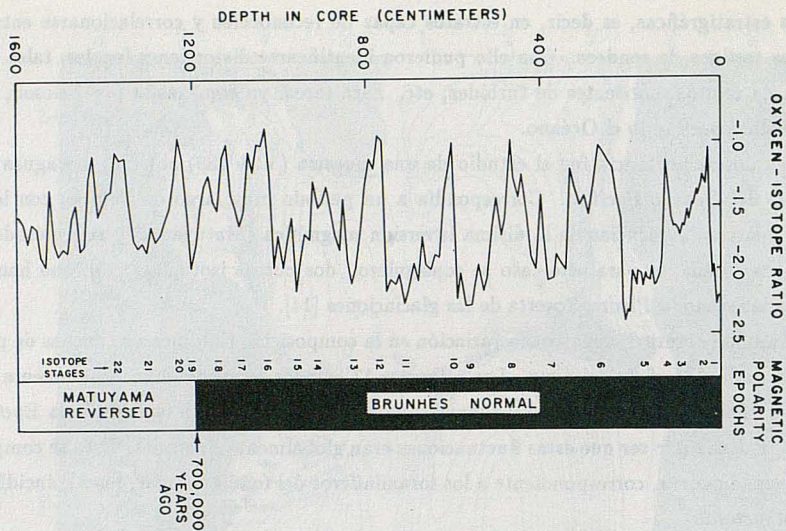


Figura 3.—Curvas isotópicas y magnéticas realizadas por Shackleton y Opodyke (1973) en la muestra V28-238. Adaptado de CLIMAP (1976)

4. Elementos orbitales terrestres

El problema de calcular la órbita y la rotación de la Tierra durante un periodo muy grande de tiempo está relacionado directamente con la estabilidad del sistema Solar. Según una definición de Weierstrass, el sistema solar es estable si los elementos orbitales de los planetas se pueden expresar como sumas de términos periódicos. Los autores clásicos toman en las teorías unos pequeños parámetros que suele ser el cociente entre la masa de un planeta y la del Sol ($\mu_i = m_i/m_0$). Lagrange ya demostró que hasta el primer orden en los parámetros, las ecuaciones diferenciales de los semiejes mayores contienen solamente términos de corto periodo. Tisserand [26] demuestra el mismo resultado hasta el segundo orden en μ . Sin embargo, los resultados principales se deben a Poincaré [23], quien afirma que las series de las teorías de perturbación son no convergentes en cualquier abierto del espacio fásico, y además, añade que las soluciones de las ecuaciones de movimiento para el problema de $n + 1$ cuerpos se pueden construir formalmente hasta cualquier orden de los pequeños parámetros como series de Fourier, siempre que nos olvidemos de la cuestión de la convergencia de dichas series. El problema radicaba en encontrar estas series de un modo explícito. La aparición en los años 1960 de métodos de perturbaciones basados en transformaciones de Lie ha permitido el obtener estas series de un modo más simple.

Otro factor que ha posibilitado el avance en la deducción de teorías analíticas, ha sido la introducción de los ordenadores en Mecánica Celeste. En efecto, hasta hace solamente unos años, las teorías solamente podían contrastarse con las observaciones de los planetas, lo que solamente daba garantías de la validez para un corto periodo de tiempo, pero con la aparición

de los ordenadores, se han podido calcular numéricamente las órbitas de los planetas y comparar los resultados obtenidos con las efemérides deducidas analíticamente, así como obtener distintas integraciones de la órbita de la Tierra, válidas para varios millones de años. ([19,24]).

Por otra parte, con el desarrollo de manipuladores algebraicos ha sido posible el obtener teorías planetarias de orden mucho más elevado, con garantías de no introducir errores y en un tiempo asombrosamente corto. Entre estos intentos, podemos citar las teorías planetarias del Bureau de Longitudes y del Observatorio de París [4,16].

Muy someramente, vamos a describir las ecuaciones diferenciales que aparecen en este problema. Suponiendo que cada planeta es un punto de masa m_i , y que su vector posición con respecto a un sistema heliocéntrico es r_i , su movimiento viene regido por la ecuación diferencial

$$\frac{d^2 r_i}{dt^2} = -\frac{G(m_0 + m_i)}{\|r_i\|^3} r_i + \sum_{j=1}^9 \frac{Gm_j}{\|r_i - r_j\|^3} (r_i - r_j) - \sum_{j=1}^9 \frac{Gm_j}{\|r_j\|^3} r_j.$$

Se trata por lo tanto de un sistema diferencial de orden 54. Además, a estas ecuaciones hay que añadir las correspondientes a la rotación de la Tierra:

$$\frac{dI\omega}{dt} = M,$$

donde I es el momento de inercia terrestre, ω es el vector velocidad angular, y M es el momento resultante de las fuerzas externas (atracción de la Luna, Sol y demás planetas).

La precisión que se obtenga en la integración dependerá en gran medida de la exactitud con que se tomen los valores iniciales r_i , \dot{r}_i , ω , I , M , así como de la estimación de las masas m_i y de la constante de gravitación G , por lo que resultan imprescibles las observaciones astrométricas de gran precisión, así como los datos suministrados por las sondas espaciales *Pioneer* y *Voyager*.

Con respecto al movimiento de rotación de la Tierra, en la actualidad, no existe una teoría analítica que permita calcular los elementos de la rotación para varios millones de años con suficiente exactitud (≈ 0.015 radianes en los movimientos medios por millón de años), debido principalmente a la presencia de funciones elípticas, muy complicadas de manejar. Con el desarrollo de software específico, se abren nuevas perspectivas para la solución de este problema.

Una nueva dificultad se viene a añadir a las ya mencionadas y es el producir las curvas de insolación en las distintas latitudes a partir de los resultados de la integración, bien analítica, bien numérica, o semi-analítica. Dignos de destacar son los trabajos de Berger [2,3], quien obtiene las curvas de insolación de Milankovitch para los últimos 5 millones de años.

5. Conclusiones

La teoría astronómica de las Glaciaciones ha supuesto un esfuerzo multidisciplinar (Geólogos, Geofísicos, Astrónomos, Matemáticos, Meteorólogos, Químicos, Biólogos, etc.) con el fin de aportar luz sobre el origen de las distintas glaciaciones. Puede concluirse que hay una correlación significativa entre las curvas de insolación y las geológicas, y que las variaciones en los parámetros de la órbita terrestre son un factor determinante en el modelado del clima. Proporciona un reloj absoluto con el que datar sedimentos del Cuaternario con una precisión varias veces mayor que

la que ahora es posible. Predice la variación climática natural en una escala de tiempo geológico para los próximos 100.000 años, un periodo de tiempo compatible con la evolución de elementos radiactivos procedentes de residuos nucleares.

Referencias

- [1] Adhémar, J.A.:1842 *Revolutions of the Sea*, publicado privadamente. Paris
- [2] Berger, A.:1988, Milankovitch theory and Climate. *Review of Geophysics*26, 624-647
- [3] Berger, A. et al.:1991, Testing the astronomical theory with a coupled climate-ice-sheet model. *Palaeogeography, Paleoclimatology, Paleoecology(Global and Planetary Change Section)* 89, 125-141
- [4] Bretagnon, P.: 1974. Termes a longues periodes dans le systeme solaire. *Astronomy and Astrophysics*, 30,141-154
- [5] Broecker, W.S. and van Dock, J.:1970, Insolation changes, ice volumes and the O¹⁸ record in deep-sea cores, *Reviews of Geophysics an Space Physics*, 8, 169-197
- [6] CLIMAP Projects Members:1976, The surface of the Ice-Age Earth. *Science*,191, 1131-1144
- [7] Cox, A., Doell, R.R., Dalrymple, G.B.:1963. Geomagnetic polarity epochs and Pleistocene geochronometry. *Nature*, 198, 1049-1051
- [8] Croll, J.: 1875 *Climate and Time* , Appleton and Co. New York.
- [9] Emiliani, C.: 1955, Pleistocene temperature. *Journal of Geology*,63 538-578
- [10] Ericson, D.B. et al.: 1963, Pliocene-Pleistocene climates and deep-sea sediments, *Science*, 139, 727-737
- [11] Geikie, J.:1874-1894, *The greatice age*, Daldy, Isbister & Co., London
- [12] Hays, J.D., Imbrie, J. y Shackleton, N.J.:1976. Variations in the Earth's Orbit: Pacemaker of Ice Ages. *Science*,194, 1121-1132
- [13] Imbrie, J. and Kipp, K.P.:1971 A new micropaleontological method for quantitative paleoclimatology: application to a late Pleistocene Caribbean core, in *Late Czenozoic glacial ages*, (K.K. Trekian, ed.) Yale Univ. Press, 71-181.
- [14] Imbrie, J. and Imbrie, K.P.:1979 *Ice Ages: Solving the mistery*,Enslow, Hillside, N.J.
- [15] Köppen, W. and Wegener, A.:1924, *Die Klimate der Geologischen Vorzeit*, Gebründer Borntraeger, Berlin
- [16] Laskar, J.:1985, *Astronomy and Astrophysics*, 144,133
- [17] Libby, W.F.:1952, *Radiocarbon dating* Univ. Chicago Press. Chicago
- [18] Lyell, C.: 1830, *Principles of Geology*, John Murray, London.
- [19] Milani, A., Nobili, A. and Carpino, M.: 1987, Secular variations of the semimajor axes: Theory and experiments *Astronomy and Astrophysics*,172,265

- [20] Milankovitch. M.: 1920, *Théorie Mathématique des Phénomènes Thermiques Produits par la Radiation Solaire*, Academie Yougoslave des Sciences et des Arts de Zagreb, Gauthier-Villars, Paris.
- [21] Milankovitch. M.: 1941, *Kanon der Erdbestrahlung R. Serbian Acad. Spec. Publ. 132, Sect. Math. Nat. Sci., 33* (Canon of insolation and the ice-age problem, Traducción inglesa por Israel Program for Scientific Translations, Jerusalem, 1969)
- [22] Penck, A. and Brückner, E.: 1909, *Die Alpen im Eiszeitalter*, Tauchnitz, Leipzig
- [23] Poincaré, H.: 1893. *Les nouvelles methodes de la Mecanique Celeste* Gauthiers-Villars. Paris
- [24] Quinn, T.R., Tremaine, S. and Duncan, M.: 1991, A three million year integraton of the Earth's orbit *Astronomical Journal*, 101, 2287-2305
- [25] Shackleton, N.J. and Opdyke, N.D.: 1973, Oxigen isotope and paleomagnetic stratigraphy of ecuatorial Pacific core V28-238: Oxigen isotope temperatures and ice volumes on a 10^5 and 10^6 year scale, *Quaternary Research*, 3, 39-55
- [26] Tisserand, F.: 1876, *Traité de Mecanique Celeste* Gauthiers-Villars. Paris

CONSTRAINTS IN THE FILTERING AND INVERSION OF LONG- AND SHORT-PERIOD SEISMIC DATA AND UNCERTAINTY IN INTERPRETATION

J. BADAL

Departamento de Física Teórica-Geofísica, Universidad de Zaragoza,
Plaza de San Francisco, 50009 Zaragoza, Spain

ABSTRACT

Modelling of layered earth structures based on the analysis of dispersed seismic signals and 2D theoretical earth models of shear wave velocity as a function of depth, require phase and group velocity calculations of surface waves and also inversion of the dispersion data obtained. Firstly, errors affecting surface wave velocity measurements are assessed. We discuss present-day constraints in the filtering and inversion processes that we have used for both long- and short-period seismic data and we comment on the uncertainty in interpretation. Concerning the determination of velocity data, different parameters involved in the techniques used for the filtering of dispersed seismic signals in order to minimize effects of modal contamination, are considered. The choice of velocity data for inversion, based on the wavelength-distance "proportionality relationship", is also discussed in relation to deep structures. Other factors taken into account are the characteristic parameters of the inversion method used and the initial earth model chosen for the inversion process. Finally, some aspects which constrain the interpretation of the inversion results are examined.

Key words: dispersion, inversion, surface waves.

INTRODUCTION

During the past two decades the structure of the crust and upper mantle of the Iberian Peninsula has been studied by means of surface wave dispersion analysis. Several earth models have been obtained for different domains, not only for those of continental structure but also for those of oceanic structure near the peninsula.

These models have changed with time as the data and processing techniques have improved (Corchete, Payo & Badal, 1991). Analysis of surface waves across the Iberian Peninsula and adjacent zones (Payo, 1965 and 1970; Payo & Ruiz de la Parte, 1974), even the most recent (Corchete *et al.*, 1990; Badal *et al.*, 1992a), have been based on analog data recorded at the long-period Iberian stations. Three years ago, however, as part of the ILIHA Project (Paulssen, 1990), fourteen broadband stations of the portable Network of Autonomously Recording Stations (NARS) were installed in Spain and Portugal. Most of these stations operated for a period of one year (February 1988-March 1989) and recorded body and surface waves generated by distant and local earthquakes. The installation of the NARS array on Iberian territory has provided a large set of digital data and substantially increased the path coverage for single-station and two-station Rayleigh wave velocity measurements, thus making possible a more detailed study of the crust and upper mantle structure than was previously possible. With the ILIHA data set now available, it has been possible for instance to investigate the northern Iberian region (Payo *et al.*, 1992), an area which was till now unexplored by long-period data, and also the lithospheric elastic layer structure of the southern half of the peninsula (Badal *et al.*, 1992b), where such an important zone like the Betic Cordillera is located which coincides with the region of convergence of the African and Eurasian plates.

In previous work (Badal *et al.*, 1992a and b; Payo *et al.*, 1992), we have studied the deep elastic structure of the Iberian Peninsula using digital filtering techniques which provide a significant improvement in signal-to-noise ratio (Corchete *et al.*, 1989), and an inversion method of seismic wave velocities based on the stochastic inverse operator (Badal, 1990). Digital filtering techniques are used to obtain phase and group velocity dispersion curves of surface waves. The stochastic inverse is used to solve the inverse problem according to the generalized inversion theory (Aki & Richards, 1980; Menke, 1984; Tarantola, 1987), and so obtain theoretical earth models.

The techniques used to determine deep structures can also be used to determine shallow structures (Russell, 1987). Several studies support this assertion (Macbeth & Burton, 1985 and 1986; Mokhtar, Herrmann & Russell, 1988). We have used the multiple filter technique and the stochastic inverse operator to filter and invert short-period seismic data (0.2-2.5 s) in order to determine the shallow structure of a small region in northwestern Iberia (Sarrate *et al.*, 1992). In this study higher Rayleigh modes are analyzed for the first time to obtain the structure of the

sedimentary layers. This research line based on observations of high frequency Rayleigh waves (0.4-5.0 Hz) generated in the course of seismic refraction experiments, provides not only the fine S-velocity structure of the medium at shallow depths, but also energy attenuation models showing the anelastic properties of the shallower earth materials.

The aims of this paper are to discuss present-day constraints in the filtering and inversion processes that we have used for both long- and short-period seismic data and to comment on the uncertainty in interpretation.

ERRORS AFFECTING SURFACE WAVE VELOCITY MEASUREMENTS

Single-station group velocity measurements are affected by *random errors*, like uncertainties in surface wave travel times, *systematic errors*, such as instrumental errors, and also *errors due to noise* associated with the seismic signal, such as those due to propagation effects. Although we have almost always performed two-station velocity measurements, an assessment of these types of error is given below.

Change of initial phase with frequency

This effect may be neglected if epicentral distances are sufficiently large (Kanamori & Abe, 1968).

Uncertainties in surface wave travel times

They are almost entirely due to uncertainties in origin time and source location. The rms error arising from these effects is 5 s. For the error in origin time, an upper bound of 4 or 5 s is quite acceptable (Nakanishi & Anderson, 1984; Hadiouche & Jobert, 1988; Montagner & Tanimoto, 1990). For an average path length of 650 km, this would correspond to a mislocation of about 15 km (Mindevalli & Mitchell, 1989). For surface wave paths including the effects of source finiteness and of digitizing error, the rms error is also 5 s (Forsyth, 1975). Those errors can be assumed to be very small when using small events (m_b magnitude less than 4.8) and digital data.

Instrumental errors

These errors are negligible within the period range for which the instrumental signal-to-noise ratio is optimum.

Error due to digitization

Non-existent with digital data.

Errors due to propagation effects

They are caused by interference with other phases, lateral refraction, or incomplete separation of modes. The rms errors in group times are of the order of 3.0-3.5 s for Love waves and 3.5-4.0 s for Rayleigh waves (Mindevalli & Mitchell, 1989).

Estimates of the total error in single-station velocity measurements

The total error in group times is of the order of 8.0-8.5 s for Love waves, and 8.5-9.0 s for Rayleigh waves. Thus, for an average path length of 650 km and Love wave group velocities of 2.9 to 3.9 km s⁻¹, and Rayleigh wave group velocities of 2.7 to 3.3 km s⁻¹, we have an error in Love wave group velocity of 0.11-0.21 km s⁻¹ and an error in Rayleigh wave group velocity of 0.10-0.14 km s⁻¹. Nevertheless, these estimates of the error in single-station velocity measurements can be substantially reduced through filtering of seismograms.

Two-station phase and group velocity measurements are also affected by errors of this type. In order to estimate the total error in the phase and group velocities, the factors listed below were taken into account.

Error due to the origin time

This error arises from the source finiteness and the source time function. It is small for seismic events that have m_b magnitude greater than 5.5. Most of the earthquakes we have considered are of $m_b > 5.5$. Two-station measurements give better results since they are not affected by source uncertainties, hence in our case the uncertainty in the origin time is not an error factor.

Error due to the mislocation of the earthquake

Mislocations of nuclear explosions provide an estimate of the uncertainty in epicentral location of the order of 20 km. For a typical group velocity of 3.80 km s⁻¹, this would correspond to an error of 0.015 km s⁻¹.

Error due to the seismograph system

Instrumental errors are negligible within the period range for which the instrumental signal-to-noise ratio is optimum. WWSSN stations have a good response over the range 10-100 s. Other stations with long-period instruments, also used by us, have a similar response over these periods. Within this period range, we verified using a real seismic signal that the error due to the amplitudes not being reduced to ground motion is negligible.

Error due to digitization

It is present only when dealing with analog data. Elimination of the very high frequencies and base-line corrections (James & Linde, 1971) can be applied in an attempt to avoid scatter in the data. In order to estimate this type of error, an analytical function was digitized and the points so obtained were then fitted step by step by a polynomial function. The differences between these two functions were determined and thus the error in digitization was obtained. This error is not significant and thus does not affect the phase and group velocity measurements.

Error due to multipathing

Propagation effects, such as lateral refractions, can be minimized by means of a rigorous control of approach directions of wavefronts. We verified the paths followed by the seismic energy from an epicentre to the peninsula. We chose epicentres very close to the great circle determined by each pair of stations. In all cases the angle between the two great circles drawn from an epicentre to its corresponding pair of stations was less than 2 degrees. As an additional control of the line up of the two stations, we considered the angle formed by the line joining those two stations and the longer of the two great circles and restricted this to be less than 4 degrees.

Error due to higher mode interference

Effects of modal contamination can be minimized using digital filtering techniques (Corchete *et al.*, 1989), such as the multiple-filter technique (MFT) (Dziewonski, Bloch & Landisman, 1969; Dziewonski, Mills & Bloch, 1972; Dziewonski & Hales, 1972; Herrmann, 1973), a time-variable filter (TVF) (Cara, 1973), or a phase-matched filter (PMF) (Herrin & Goforth, 1977; Goforth & Herrin, 1979). We use the MFT and a TVF (Badal *et al.*, 1990) to remove this kind of interference and to isolate the fundamental mode from Rayleigh wavetrains.

PARAMETERS INVOLVED IN THE FILTERING OF DISPERSED SEISMIC SIGNALS

We obtain group velocity dispersion using the MFT. A TVF allows us to efficiently remove higher mode interference. Once the signal is filtered, the MFT is used again to directly compute group arrival times at each station. From these times, the interstation group velocity can easily be calculated. We determine phase velocity dispersion by applying a spectral method to the filtered seismograms that we have previously obtained by means of multiple filtering analysis and time variable filtering, which is the standard frequency-domain Wiener deconvolution (Hwang & Mitchell, 1986). The interstation Green's function required by this technique, which is the ratio of the smoothed cross-spectrum to the smoothed autospectrum of the first station, is calculated by windowing the corresponding correlation functions with trapezoidal windows of different lags. In this manner, the Wiener filtering is optimized to achieve the greatest noise reduction possible. The interstation phase velocity is calculated from the phase spectrum of the transfer function using the classical formula of the two-station method. Some unknown parameters involved in the filtering process of long-period seismic data are worth mentioning and will now be discussed.

The relative bandwidth of the Gaussian filter used in the MFT

As is well known, this technique makes use of a symmetric band-pass filter whose relative bandwidth is not fixed *a priori*. The value assigned to this filter parameter ($\alpha^{-1/2}$) depends on the dispersion of the signal, the noise level and possible modal contamination, such that a good frequency-time resolution is obtained; for instance, $\alpha=20$ (Corchete *et al.*, 1989), or $\alpha=40$ (Badal *et al.*, 1990).

The time lag of the window used when applying a TVF

To check the smoothness and bias of the isolated spectrum, a Tukey window can be used as it is best for signals with a moderately peaked spectrum (Hwang & Mitchell, 1986). In accordance with Landisman, Dziewonski & Sato (1969) and Hwang & Mitchell (1986), a Tukey window with a time lag of five times the period considered is large enough for a correct spectral estimate (Badal *et al.*, 1990).

The lags of the windows applied in order to optimize Wiener filtering

Trapezoidal windows of different lags are applied to correlation functions calculated from isolated seismograms. Their lags can be optimized visually in order to achieve the largest possible noise reduction. A modified Tukey window which

takes into account the respective lags of a rectangular window and a Parzen window, can also be used (Badal *et al.*, 1990).

A CONSTRAINT IN THE CHOICE OF VELOCITY DATA

Surface wave velocity dispersion data over an arbitrarily large period range cannot be used for inversion. In the case of single-station measurements from observations of surface waves due to earthquakes near the Iberian Peninsula, with epicentres either in the Atlantic Ocean or in North Africa, or in the Pyrenees, it is not easy to have well-dispersed wave packets with clear amplitudes. This is certainly a serious difficulty when trying to obtain group velocities corresponding to periods greater than 50 s. As far as studies of the crustal structure based on surface wave velocity calculations is concerned, a period of 35 s is a real upper limit for operations.

In the case of two-station measurements from observations of surface waves due to teleseismic events, an upper limit for the period or wavelength of the waves must be fixed. The reason for initially adopting such an upper limit is none other than the "proportionality relationship" that exists between the wavelength and the interstation distance travelled by the waves, i. e. the greater the wavelength the greater the interstation distance must be in order to obtain a reasonable accuracy (McEvelly, 1964). The different trajectories between pairs of long-period seismological stations in the Iberian Peninsula which the waves traverse following nearly great-circle paths, cover a minimum distance of 351 km (Toledo-Alicante, Toledo-Málaga) and a maximum distance of 760 km (Porto-Alicante, Porto-Ebro). As a consequence, only periods up to 100 s should be considered (Badal *et al.*, 1992a). With ILIHA data obtained using the broadband NARS stations, the distances between stations vary considerably, although a representative interstation distance is of the order of 500 km and therefore the periods considered should be lower than 90-100 s (Payo *et al.*, 1992).

FACTORS AFFECTING THE INVERSION PROCESS OF SURFACE WAVE VELOCITY DISPERSION DATA

Modern inversion theory has been widely used in geophysics since its development by Backus & Gilbert (1970). Several variations of the basic procedures

have since been developed. For instance, stochastic inversion is equivalent to the damped least-squares method and minimizes the magnitude of the solution and error vectors (Franklin, 1970; Jackson, 1972; Wiggins, 1972). Differential inversion minimizes the magnitude of the error vector plus the differences between adjacent solution elements (Twomey, 1977). For the same set of dispersion data the stochastic inversion method and the differential inversion method yield nearly the same models if a proper initial model can be supplied. We have adopted an inversion method based on the stochastic inverse operator (Badal, 1990) for the determination of theoretical shear velocity models. Numerical experiments show however, that stochastic inversion fails to give a stable solution when starting from a half-space model, whereas differential inversion works well (Hwang & Mitchell, 1987). We consider initially that the distribution of the elastic constants in the earth is approximated by a finite number of plane-horizontal, homogeneous layers overlying a half-space. Factors affecting the inversion process are examined below.

The damping factor

In the stochastic inversion scheme, the Lavenberg-Marquardt parameter or damping factor ϵ^2 (Lawson & Hanson, 1974), defined as the ratio of the variance associated with noise to the variance associated with the model parameters, permits control of the inherent instability of the inversion process (Badal, 1990). ϵ^2 degrades the resolution, but reduces the error in the solution and stabilizes it, thus reducing the covariance. A small ϵ^2 value would cause the calculation to diverge while a large value would cause the convergence to be too slow. The parameter can be chosen such that models have the best possible resolution within the constraint that they be realistic, i. e. they do not contain large fluctuations in velocity. A value near or equal to 30 for the damping constant can be considered (Badal *et al.*, 1992a), although experience has demonstrated the necessity of reducing this value to 10 or even 5, for problems involving joint inversion of phase and group velocities (Córchete *et al.*, 1990).

Variations of the data vector elements

If the data vector is assumed to be composed of independent random variables, the variances associated with the solution vector are a function of the variances associated with the data vector. Appropriate filtering increases the signal-to-noise ratio and reduces the scatter in the velocity data (Badal *et al.*, 1990 and 1992a).

The amount of information

In previous studies of surface wave propagation across the Iberian Peninsula, simultaneous inversion of phase velocities and group velocities has never been carried out. Either phase velocities or group velocities have been inverted, but in no case has the task of performing joint inversion of velocity data been undertaken. Joint inversion of Rayleigh wave phase and group velocities in Iberia has been performed only recently (Corchete *et al.*, 1990; Badal *et al.*, 1992a and b; Payo *et al.*, 1992). In these cases, it is obvious that a greater quantity of information is handled, which consequently affects the computation process. Nevertheless, in spite of the added difficulties that the joint inversion of phase and group velocities involves, this type of inversion results in a more realistic earth model due to the greater number of constraints which must be adhered to.

The number of parameters specifying the initial earth model

The *a priori* relationship $\alpha = [2(1-\sigma)/(1-2\sigma)]^{1/2}\beta$ between compressional wave velocity α and shear wave velocity β , where σ is Poisson's ratio, can be used in order to reduce the total number of model parameters via Poisson's ratio. Also, a connection between density and shear velocity can be established via an empirical relation (Macbeth & Burton, 1985 and 1986). This reduction of the model's independent variables is justified as the Rayleigh waves are in general more sensitive to changes in the shear constants than to other structural parameters (Bloch, Hales & Landisman, 1969). At shallow depths Poisson's ratio can be assumed to be known reasonably well, thus constraining compressional velocities while empirical relations can constrain density values. Mindevalli & Mitchell (1989) have assumed, for example, that Poisson's ratio is 0.25 in the crust and 0.27 in the upper mantle, and have calculated densities using known relations between density and compressional wave velocity in the crust (Talwani, Sutton & Worzel, 1959) and mantle (Birch, 1964).

Two common rock types have comparable P velocity at the temperature and pressure at the Moho: the ultrabasic rock *peridotite*, and *eclogite* which is the high pressure form of the basic rock gabbro. Observed seismic velocities suggest that peridotite is dominant in the topmost mantle. The values of Poisson's ratio calculated from P_n and S_n phases range from 0.24 to 0.27 which is consistent with peridotite (but falls short of the few experimentally determined values for eclogite, which exceed 0.30). So, we have constrained the compressional velocities to a value for Poisson's ratio of 0.25 (Badal *et al.*, 1992a), and we have made a connection

between density and shear velocity via an empirical relation by considering an average mantle density of 3370 kg m^{-3} down to a depth of 200 km in accordance with PREM model (Dziewonski & Anderson, 1981).

If the initial discrete earth model is defined by its structural parameters, like the layer thicknesses and the corresponding shear velocity values in these layers, the layer thicknesses are fixed and the shear velocity in each layer is estimated from the dispersion data.

The dimension or number of degrees of freedom of the inverse problem

The most prudent form of action is to fix a priori the number of layers in the initial earth model bearing in mind that a trade-off exists between the number of degrees of freedom of the problem to be solved and the structure to be determined. An initial earth model having 20 layers has been considered at shallow depths (Sarrate *et al.*, 1992) in order to appreciate better the variations in shear velocity with depth at crustal depths. At subcrustal depths, we have chosen from the first an initial discrete earth model having 25 layers (Badal *et al.*, 1990).

UNCERTAINTY IN INTERPRETATION

Interpretation of our inversion results is obviously constrained by several factors, each of a very different nature. They all contribute to uncertainties, which we comment below.

Uncertainty due to mixed-path velocity measurements

We perform surface wave dispersion analysis on mixed-path velocity measurements and not on pure-path measurements, thus strictly speaking the data that are inverted are non-regionalized and so our theoretical earth models might not be accurate. However, the observed velocity variations over the relatively short seismic paths travelled by the surface waves through the peninsula, are small; therefore, we believe that regionalization based on the main geotectonic units of Iberia, is not absolutely necessary.

Uncertainty due to the model initially considered

The inversion procedure requires a starting model. One should be cautious not to influence results with *a priori* assumptions, i. e. not to bias final models with any

assumptions concerning the location of velocity discontinuities in the initial model. We have previously referred to the elastic parameters which define the starting earth model as well as the number of layers in this initial model. Discrete distributions of compressional wave velocity and density as a function of depth are restricted by means of Poisson's ratio and an empirical relationship, respectively. The thicknesses of the main layers and the position of some of the velocity discontinuities in the initial model, and also the smooth or sharp character of these discontinuities, could affect the earth models obtained and be a cause of uncertainty in the interpretation of the final inversion results. Unfortunately, in relation to the shear velocity structure of the Iberian Peninsula at crustal and subcrustal depths, the question of whether all these geometric features are resolved or not, arises.

Uncertainty due to the resolution of the problem

A smoothed solution is obtained by stochastic inversion, in the sense that each parameter of the model obtained is a 'weighted' mean of the parameters that define the real earth model (Badal, 1990). The resolving kernels, which are nothing more than the weights involved, give the resolution of the problem as a measure of the degree of agreement between the true and computed solutions. In fact, these resolving kernels, which are usually represented graphically as continuous functions of depth, inform on the exactness achieved in calculating the solution with respect to depth. Thus different resolving kernels, for different reference depths, due to the approximations made in the calculation of models via generalized inversion, must be examined. In relation to the recent inversion results obtained till now of the deep structure of the Iberian Peninsula (Badal *et al.*, 1992a and b; Payo *et al.*, 1992), we can conservatively say that the theoretical models that we have obtained appear to be valid for depths ranging between 20 and 200 km.

Uncertainty due to a violation of the linearity assumption in the inversion

Smooth starting models in the linearized inversion may introduce non-linear effects. Our resolution calculations should therefore be interpreted as a rough order-of-magnitude calculation. Experiments show that the thickness of the high-velocity lid determines the resolution in the density; a smaller lid thickness results in an improvement in resolution. Lid thickness can be substantially reduced by allowing a transverse isotropic parametrization of velocity in the inversion (Lerner-Lam, 1985). In our present inversion this is not accounted for.

Uncertainty due to the possible anisotropy of the medium

A further issue is that of possible anisotropy of the medium. Possible anisotropy based on the so-called Love-Rayleigh wave discrepancy could take place in the Iberian region. This phenomenon is considered to be a typical effect of anisotropy, less likely to be influenced by the presence of lateral heterogeneities than the polarizations of the surface waves or azimuthal variations of their velocities. Preliminary results inferred from ILIHA data, indicate that in a NW-SE azimuth, Love and Rayleigh wave phase velocities appear compatible with an isotropic model, whereas in a NE-SW direction, Love waves are slightly faster than would be predicted from an isotropic model derived from Rayleigh waves (Paulssen, 1990). Another results also inferred from ILIHA data, are the anomalously low Rayleigh wave amplitudes observed in South Spain which, in addition, appear to exhibit extremely large time- and frequency-dependent polarization anomalies. The highly attenuated and distorted surface waves in the South of Iberia suggest that, apart from anelastic attenuation, multiple scattering and reverberation mechanisms are operative within a SW-NE trending zone at the Betic Cordillera (Paulssen *et al.*, 1990). Since there is at present no adequate Love wave data set at hand for the whole Iberian region, we have interpreted our data assuming an isotropic model. Nevertheless, a possible alternative explanation for a Love/Rayleigh wave incompatibility could be a laterally varying structure or a failure to include a sufficient number of layers in the inversion process (Mitchell, 1984). Models obtained from the inversion of Rayleigh wave velocities that are significantly different (by more than the model uncertainties) from those obtained from the inversion of Love wave velocities, is a result sometimes taken to imply polarization anisotropy of the upper mantle material (McEvelly, 1964). Polarization anisotropy might be present throughout portions of the structure. However, if the observations were explained equally well by combined inversion of both datasets after solving the corresponding direct problems, polarization anisotropy would not be required to explain the observed long-period seismic data (Mindevalli & Mitchell, 1989).

Joint inversion of the data, erroneously assuming an isotropic model, may produce unrealistically high or low shear velocity values through some portions of the models in an attempt to fit both sets of data, and so give systematic errors rather than random errors. The reason for this is that the maximum sensitivities of Rayleigh and Love waves to elastic properties, for a given period, generally occur through different depth ranges. These errors could be important at depths

where resolution is very poor. This is a source of uncertainty that should be considered in the future.

In relation with the small lateral variations of velocity observed at present in the Iberian area, the following question remains: are the observed lateral variations of Rayleigh wave velocity due to azimuthal anisotropy or can they rather be explained by regional variations of velocity? At the moment the reply seems to be the latter. Clear evidence of azimuthal anisotropy based on differences in Rayleigh wave velocity, has not been reported yet (G. Payo, private communication). Available geological and geophysical data (Banda, 1988) are more consistent with the idea of regional variations in velocity rather than with an interpretation based on anisotropy. For instance, heat flow values inferred from the map compiled by Banda, Albert-Beltrán and Fernández (1989) show a rough correlation with surface wave velocity values, and since high heat flow values usually correspond to low velocity values and viceversa, an interpretation based on regional variations seems to be more appropriate.

ACKNOWLEDGMENTS

The author would like to thank the Director of the Geophysical Observatory of Toledo of the National Geographical Institute, Dr. Gonzalo Payo, for his useful comments and review. E. Gurría revised the language. This research was supported by the Dirección General de Investigación Científica y Técnica (DGICYT), Projects PB89-0349 and PB90-0599-C03-02.

REFERENCES

- Aki, K. and P. G. Richards (1980). *Quantitative Seismology. Theory and Methods*, vol. II, Freeman, San Francisco, 675-717 pp.
- Backus, G. and F. Gilbert (1970). Uniqueness in the inversion of inaccurate gross earth data, *Phil. Trans. R. Soc. A.*, **266**, 123-192.
- Badal, J. (1990). Inversion of seismic wave velocities by means of the stochastic inverse operator, *Rev. Acad. Ciencias Zaragoza*, **45**, 127-150.

Badal, J., V. Corchete, G. Payo, J. A. Canas, L. Pujades and F. J. Serón (1990). Processing and inversion of long-period surface-wave data collected in the Iberian Peninsula, *Geophys. J. Int.*, **100**, 193-202.

Badal, J., V. Corchete, G. Payo, F. J. Serón, J. A. Canas and L. Pujades (1992a). Deep structure of the Iberian Peninsula determined by Rayleigh wave velocity inversion, *Geophys. J. Int.*, **108**, 71-88.

Badal, J., V. Corchete, G. Payo, J. A. Canas and L. Pujades (1992b). Lithospheric elastic layer structure of the southern Iberian region, (submitted to *Tectonophysics*).

Banda, E. (1988). Crustal parameters in the Iberian Peninsula, *Phys. Earth planet Inter.*, **51**, 222-225.

Banda, E., J. Albert-Beltrán and M. Fernández (1989). Regional geothermal gradients and lithospheric structure in Spain, Springer Verlag (in press).

Birch, F. (1964). Density and composition of mantle and core, *J. geophys. Res.*, **69**, 4377-4388.

Bloch, S., A. L. Hales and M. Landisman (1969). Velocities in the crust and upper mantle of southern Africa from multi-mode surface-wave dispersion, *Bull. seism. Soc. Am.*, **59**, 1599-1629.

Cara, M. (1973). Filtering of dispersed wavetrains, *Geophys. J. R. astr. Soc.*, **33**, 65-80.

Corchete, V., J. Badal, G. Payo and F. J. Serón (1989). Filtrado de ondas sísmicas dispersadas (Filtering of dispersed seismic waves), *Rev. Geofísica*, **45**, 39-58.

Corchete, V., J. Badal, G. Payo, J. A. Canas, L. Pujades and F. J. Serón (1990). An attempt at joint inversion of Rayleigh wave phase and group velocities in Iberia, *Rev. Geofísica*, **46**, 83-96.

Corchete, V., G. Payo and J. Badal (1991). Evolution of the crust-mantle models in the Iberian region by surface waves, *Proc. of the XXII ESC Meeting*, Barcelona, September 1990.

Dziewonski, A. M. and D. L. Anderson (1981). Preliminary reference Earth model, *Phys. Earth planet. Inter.*, **25**, 297-356.

Dziewonski, A. M. and A. L. Hales (1972). Numerical analysis of dispersed seismic waves, *Meth. compt. Phys.*, **11**, 39-85.

Dziewonski, A. M., S. Bloch and M. Landisman (1969). A technique for the analysis of transient seismic signals, *Bull. seism. Soc. Am.*, **59**, 427-444.

Dziewonski, A. M., J. Mills and S. Bloch (1972). Residual dispersion measurement. A new method of surface analysis, *Bull. seism. Soc. Am.*, **62**, 129-139.

Forsyth, D. A. (1975). The early structural evolution and anisotropy of the oceanic upper mantle, *Geophys. J. R. astr. Soc.*, **43**, 103-162.

Franklin, J. N. (1970). Well-posed stochastic extensions of ill-posed linear problems, *J. Math. Anal. Appl.*, **31**, 681-716.

- Goforth, T. and E. Herrin (1979). Phase-matched filters: applications to the study of Love waves, *Bull. seism. Soc. Am.*, **69**, 27-44.
- Hadiouche, O. and N. Jobert (1988). Geographical distribution of surface-wave velocities and 3-D upper-mantle structure in Africa, *Geophys. J.*, **95**, 87-109.
- Herrin, E. and T. Goforth (1977). Phase-matched filters: applications to the study of Rayleigh waves, *Bull. seism. Soc. Am.*, **67**, 1259-1275.
- Herrmann, R. B. (1973). Some aspects of band-pass filtering of surface waves, *Bull. seism. Soc. Am.*, **63**, 663-671.
- Hwang, H. J. and B. J. Mitchell (1986). Interstation surface wave analysis by frequency-domain Wiener deconvolution and modal isolation, *Bull. seism. Soc. Am.*, **76**, 847-864.
- Hwang, H. J. and B. J. Mitchell (1987). Shear velocities, Q_β , and the frequency dependence of Q_β in stable and tectonically active regions from surface wave observations, *Geophys. J. R. astr. Soc.*, **90**, 575-613.
- Jackson, D. D. (1972). Interpretation of inaccurate, insufficient and inconsistent data, *Geophys. J. R. astr. Soc.*, **28**, 97-109.
- James, D. E. and A. T. Linde (1971). A source of major error in the digital analysis of World Wide Standard Station seismograms, *Bull. seism. Soc. Am.*, **61**, 723-728.
- Kanamori, H. and K. Abe (1968). Deep structure of island arcs as revealed by surface waves, *Bull. Earthq. Res. Inst. Tokyo Univ.*, **46**, 1001-1025.
- Landisman, M., A. Dziewonski and Y. Sato (1969). Recent improvements in the analysis of surface wave observations, *Geophys. J. R. astr. Soc.*, **17**, 369-403.
- Lawson, Ch. L. and R. J. Hanson (1974). *Solving Least Squares Problems*, Prentice-Hall, New York.
- Lerner-Lam, A. L. (1985). Tradeoff between lid thickness and anisotropy using linearized waveform inversion (abstract), *EOS, Trans. Am. geophys. Un.*, **66**, 966-967.
- Macbeth, C. D. and P. W. Burton (1985). Upper crustal shear velocity models from higher mode Rayleigh wave dispersion in Scotland, *Geophys. J. R. astr. Soc.*, **83**, 519-539.
- Macbeth, C. D. and P. W. Burton (1986). Propagation of 0.7-2.5 Hz Rayleigh waves in Scotland, *Geophys. J. R. astr. Soc.*, **84**, 101-120.
- McEvelly, T. V. (1964). Central U. S. crust-upper mantle structure from Love and Rayleigh wave phase velocity inversion, *Bull. seism. Soc. Am.*, **54**, 1997-2015.
- Menke, W. (1984). *Geophysical Data Analysis: Discrete: Inverse Theory*, Academic Press, Orlando, FL.
- Mindevalli, O. Y. and B. J. Mitchell (1989). Crustal structure and possible anisotropy in Turkey from seismic surface wave dispersion, *Geophys. J. Int.*, **98**, 93-106.

- Mitchell, B. J. (1984). On the inversion of Love- and Rayleigh-wave dispersion and implications for Earth structure and anisotropy, *Geophys. J. R. astr. Soc.*, **76**, 233-241.
- Mokhtar, T. A., R. B. Herrmann and D. R. Russell (1988). Seismic velocity and Q models for the shallow structure of the Arabian shield from short-period Rayleigh waves, *Geophysics*, **53**, 1379-1387.
- Montagner, J. P. and T. Tanimoto (1990). Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. geophys. Res.*, **95**, 4797-4819.
- Nakanishi, I. and D. L. Anderson (1984). Measurements of mantle wave velocities and inversion for lateral heterogeneities and anisotropy. Part II, Analysis by the single station method, *Geophys. J. R. astr. Soc.*, **78**, 537-617.
- Paulssen, H. (1990). The Iberian Peninsula and the ILIHA project, *Terra Nova*, **2**, 429-435.
- Paulssen, H., A. L. Levshin, A. V. Lander and R. Snieder (1990). Time- and frequency-dependent polarization analysis: anomalous surface wave observations in Iberia, *Geophys. J. Int.*, **103**, 483-496.
- Payo, G. (1965). Iberian Peninsula crustal structure from surface wave dispersion, *Bull. seism. Soc. Am.*, **55**, 727-743.
- Payo, G. (1970). Structure of the crust and upper mantle in the Iberian Shield by means of a long period triangular array, *Geophys. J. R. astr. Soc.*, **20**, 493-508.
- Payo, G. and E. Ruiz de la Parte (1974). Dispersion of surface wave in the Iberian Peninsula and the adjacent Atlantic and Mediterranean areas, *Geofisica Internacional*, **14**, 89-102.
- Payo, G., V. Corchete, J. Badal, F. Serón, J. A. Canas and L. Pujades (1992). First two-station Rayleigh wave velocity measurements for the northern Iberian region, *Bull. seism. Soc. Am.* (in press).
- Russell, D. R. (1987). *Multi Channel Processing of Dispersed Surface Waves*, Ph. D. Dissertation Thesis, Saint Louis University, Saint Louis.
- Sarrate, J., J. A. Canas, L. Pujades, J. Badal, V. Corchete and G. Payo (1992). Shallow structure of a region in northwestern Iberia from short-period Rayleigh wave observations, *Tectonophysics* (accepted).
- Talwani, M., G. Sutton and J. Worzel (1959). A crustal section across the Puerto Rico trench, *J. geophys. Res.*, **64**, 1545-1555.
- Tarantola, A. (1987). *Inverse Problem Theory. Methods for Data Fitting and Model Parameter Estimation*, Elsevier, Amsterdam.
- Twomey, S. (1977). *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*, Elsevier, Amsterdam.
- Wiggins, R. A. (1972). The general linear inverse problem: Implications of surface waves and free oscillations for earth structure, *Rev. Geophys. Space Phys.*, **10**, 251-285.

DISEÑO METODOLOGICO DE UNA CAMPAÑA DE PROSPECCION GEOQUIMICA MULTIELEMENTAL EN LA RED DE DRENAJE DEL PALEOZOICO DE LA CORDILLERA IBERICA.

J. Mandado, M. Arce, L.F. Auqué, M^aJ. Gimeno y J.M. Tena.

Departamento de Geología. Facultad de Ciencias. Universidad de Zaragoza. 500009 Zaragoza.

Fecha de entrega: 5 de diciembre de 1991.

ABSTRACT

Statement of any multielemental geochemical exploration program needs previous delimitation of the more suitable methodological conditions. Material sampled, geological environment and program objectives must be evaluated from a cost/efficacy point of view.

Study of a reduced zone in Paleozoic materials of Eastern Iberian Chain allows to evaluate the importance of several factors (granulometric character of aluvial materials, chemical extracting techniques and dimensions of sampling net) in order to define the dispersion aureoles.

The more suitable methodological conditions can be summarised as follows: sample, <0,16 mm fraction of aluvial sediments; extracting techniques, warm acid attack with nitric; sampling density, one sample every 300 m.

Study of selective extracting techniques allows to establish the surficial dispersion pattern and the better elements to characterize them in the zone.

1. INTRODUCCION

Las técnicas empleadas en prospección geoquímica, en sedimentos aluvionares de redes de drenaje, varían ampliamente en función de la naturaleza del objetivo, del tipo de terreno, características del material de muestreo disponible y el costo relativo de métodos alternativos (Govett, 1983). La elección de las técnicas y métodos de trabajo óptimos para un programa determinado es, tras la elección del área, la decisión más crítica con vistas a una deseada efectividad global de la campaña (Rose *et al.*, 1979).

En este marco, se consideró preceptivo el desarrollo de una campaña previa orientativa como guía para valorar las distintas opciones metodológicas disponibles y diseñar un plan de trabajo aplicable al desarrollo de una campaña de prospección geoquímica aluvionar en la red de drenaje de la Cordillera Ibérica, que minimice la relación coste/efectividad.

Nuestro objetivo en esta campaña se concreta especialmente en tres aspectos: 1) Elección de la fracción de muestra más favorable, 2) valoración de las técnicas de extracción más apropiadas, y 3) planificación de una densidad de muestreo razonable en términos de la relación coste/resultados.

El área muestreada, de unos 4.5 km², comprende el Barranco de la Tejera, Barranco de Ontanat y el tramo final del Arroyo de Alpartir, situados en el sector noroccidental de la Sierra de Algairén (provincia de Zaragoza). Se trata de arroyos con actividad más o menos discontinua y, en algun caso, de curso activo sólo en épocas de tormenta, como el caso del

Barranco de Ontanat. El interés de esta zona concreta para el desarrollo de la campaña previa reside en la existencia en ella de una antigua explotación, próxima a la desembocadura del Barranco de la Tejera (Casa de las Minas), que garantiza la existencia de una anomalía y su dispersión, y una cata profunda con una importante escombrera en la zona alta del Barranco del Ontanat.

2. METODOLOGIA DE MUESTREO

Las muestras se recogieron en el lecho activo de la red de drenaje (figura 1), y están constituidas por material pobremente clasificado de procedencia local. Suelen presentar un grado de edafización variable, con mezcla de material orgánico más o menos degradado y material inorgánico, todo ello colonizado por raíces. No son infrecuentes los aportes coluvionares laterales de importancia variable.

El intervalo de muestreo inicial fue de 200 m a lo largo del cauce, aumentando la densidad de toma a 100 m en las proximidades de la Casa de las Minas, por considerarse una zona de especial interés. En la escombrera del cargadero de esta explotación, cerrada desde hace décadas, se recogió una muestra para control analítico.

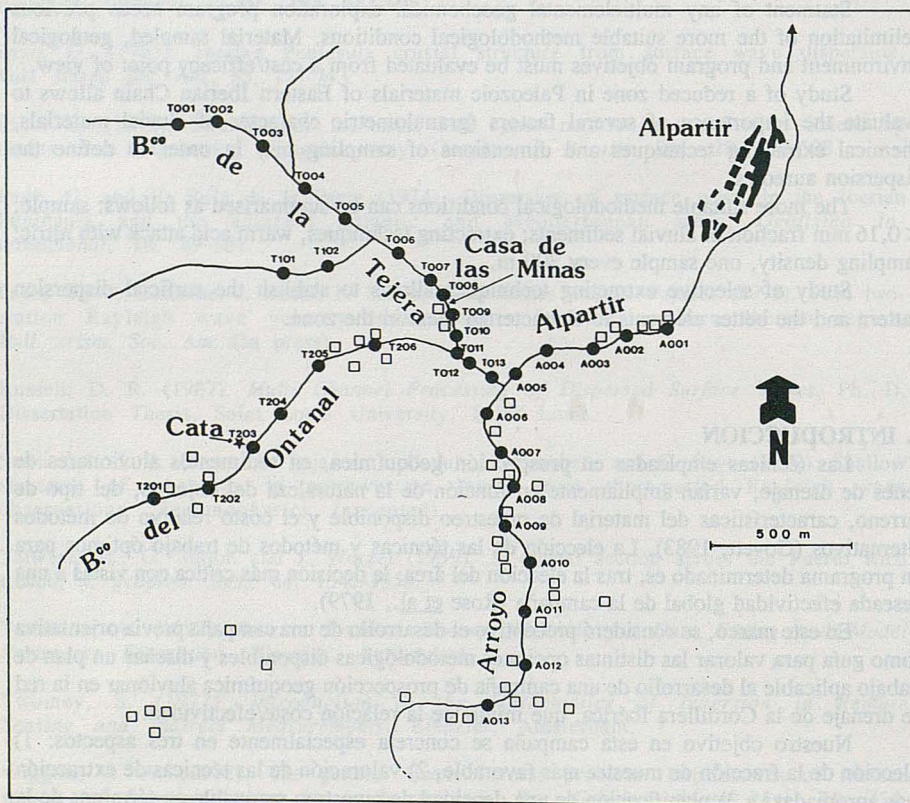


Figura 1.- Esquema de situación de las muestras.

Con el propósito de disminuir la variabilidad muestral, en cada punto se tomaron diferentes fracciones del lecho vivo en un radio de 1 a 2 m, anotando en cada caso las condiciones de muestreo y el posible aporte coluvionar (como agente de dilución de los valores de concentración). En cada punto se recogieron aproximadamente 2 kg de muestra y, en los casos en que la falta de humedad lo permitiese, se procedió al tamizado para eliminar la fracción superior a 2 mm. El número total de muestras recogidas se cifra en 34, lo que corresponde a una densidad de muestreo de 7.5 muestras por km².

Para la preparación de las muestras se procedió en primer lugar al secado del material recogido en el campo en una estufa a 60° C. Posteriormente se tamizaron, separándose las fracciones correspondientes al intervalo comprendido entre 1mm y 0.16 mm, y el inferior 0.16 mm, desechándose el resto. La fracción 1-0.16 mm fue molida para conseguir un tamaño inferior a 0.16 mm. Por último se prepararon por cuarteo dos submuestras de 50 gr para ambas fracciones.

3. METODOLOGIA DE EXTRACCION

El objetivo principal de las técnicas de extracción consiste en pasar a disolución los elementos traza contenidos en las muestras. No obstante, existen muchas variaciones en los métodos de disolución que aportan además una valiosa información en cuanto a la forma de presentación de los elementos en la muestra, en cuanto a la selección del medio de muestreo (materia orgánica, oxihidróxidos secundarios, etc.) más apropiado para obtener un mayor contraste de la anomalía, e, incluso, como medio de clasificación de anomalías según el modelo de extracción obtenido (Dijkstra *et al.*, 1987). En la bibliografía se registra una gran cantidad de técnicas de extracción aplicadas a prospección geoquímica, sintetizadas en los trabajos de Chao (1984) y Fletcher (1981).

Las técnicas de disgregación pueden reunirse en dos grandes grupos: totales y parciales. Las extracciones parciales pueden ser selectivas o no (si únicamente atacan determinadas fases minerales).

La elección de la técnica concreta a utilizar pasa, lógicamente, por un estudio preliminar de las fases presentes en las muestras de sedimento aluvial de que se dispone. Hemos empleado para ello los estudios de Rose (1975), Rose *et al.* (1979) y Chao y Theobald (1976) sobre la estabilidad y comportamiento geoquímico de las fases presentes en las muestras.

Para este trabajo se optó por un método consistente en la aplicación de series de extracción selectiva por separado, junto con una extracción parcial no selectiva mediante ataque ácido en caliente. La razón de ello es la de poder valorar las concentraciones de elementos traza asociados con oxihidróxidos de hierro-manganeso, minerales de la arcilla y materia orgánica, obtenidas mediante las extracciones parciales selectivas, y comparar los resultados con los procedentes de una disgregación parcial, aunque energética, no selectiva.

Las extracciones selectivas practicadas fueron:

- Reactivo de Tamm, oxalato amónico 0.275 M en medio clorhídrico a un pH de 3.3, con una relación muestra/reactivo de 5 gr en 50 mls y catalizado por luz natural. En estas condiciones se asegura la disolución de oxihidróxidos de hierro (amorfo y cristalino) y de manganeso. Con este tipo de tratamiento no se analizaron Pb ni Ag, que forman compuestos insolubles con el oxalato.

- $MgCl_2$, según el método de Cole y Rose (1984): 5 gr de muestra en 25 gr de $MgCl_2$ 1M y HCl hasta un pH de 5, ataque durante una hora, con agitación periódica, centrifugación y enrase hasta 50 ml. Este método sirve para extraer los cationes intercapa de los minerales de la arcilla. En estas condiciones de pH no se disgrega la red de estos minerales.

- Método de Gibbs (1973), hipoclorito sódico en medio clorhídrico a un pH de 7 ± 1 , con una relación muestra/reactivo de 5 gr en un volumen final de 200 mls. Mediante este tratamiento se destruye la materia orgánica oxidable presente en la muestra y la mayoría de los sulfuros.

La extracción no selectiva se realizó con ácido nítrico concentrado, en las siguientes condiciones: 5 gr de muestra, con 15 ml de HNO_3 y 5 ml de H_2O_2 , sobre placa calefactora y durante 90 minutos, al cabo de este tiempo se añadieron 5 ml más de HNO_3 y 1 ml de H_2O_2 , completando el volumen hasta 50 ml con agua destilada.

La aplicación de esta metodología comparada de extracción aportará una interesante información a distintos niveles:

1- Servirá como criterio para escoger la extracción que ofrezca mejor contraste entre anomalía y fondo geoquímico.

2- Podremos interpretar los resultados para valorar de forma cualitativa la distribución de los elementos traza en las distintas fases mayoritarias que componen el sedimento.

3- Proporciona información sobre el mecanismo mayoritario de dispersión para los distintos elementos relacionados con la anomalía.

En la valoración de los puntos 2 y 3 será preciso tener siempre en cuenta la influencia de los distintos factores geológicos que conforman las distintas muestras: material fuente de la anomalía, condiciones ambientales de meteorización, actividad de la red de drenaje, etc.

4. METODOLOGIA ANALITICA

Los métodos de análisis en prospección geoquímica han de ser elegidos dentro del espectro de métodos rápidos y de bajo coste, dado el importante número de muestras que usualmente se manipula. Por tanto, la precisión con que se trabaja es más pobre que la que podría conseguirse con métodos más costosos. La razón implícita en la elección del método es que, en prospección geoquímica, el objetivo primordial es descubrir tendencias regionales dentro de materiales de muestreo con una composición generalmente muy variable incluso en áreas locales. Cuando la variabilidad que genera el muestreo es grande, no es razonable derrochar recursos para reducir notablemente el error de laboratorio mediante análisis más lentos y costosos. En cualquier caso, es más útil emplear estos recursos en reducir (o cuando menos, controlar) el error de muestreo. Como apunta Miesch (1976), si el error de muestreo constituye una parte dominante del error experimental total, incluso una técnica analítica perfecta no sería útil para alcanzar la precisión experimental requerida; esto es, es más efectivo reducir el error experimental total mediante un método rápido de laboratorio con un gran número de muestras que mediante un tratamiento costoso de sólo unas pocas.

En este sentido, los métodos analíticos espectrométricos gozan de gran difusión en exploración geoquímica, en especial aquellos que permiten determinaciones multielementales, como la espectroscopía en plasma de acoplamiento inducido (ICP) que, a pesar de sus limitaciones analíticas con respecto a otros métodos (mayor límite de detección y menor precisión), permite una rápida y menos costosa elaboración.

En nuestro caso, los análisis fueron llevados a cabo mediante un espectrómetro de emisión atómica (ICP) Perkin Elmer, modelo 40 SP, del Servicio Central de Análisis de la Universidad de Zaragoza, analizándose simultáneamente As, Sb, Zn, Cd, Co, Fe, Mn, Cu y Ag. Las condiciones instrumentales de análisis y calibrado del equipo se recogen en Villareal (1991).

5. DISCUSION DE LOS RESULTADOS ANALITICOS

5.1 Valoración de la fracción de muestra más apropiada.

La granulometría de sedimento más apropiada para análisis se determinó a partir de los datos obtenidos en cuatro muestras tomadas al azar. En cada una de ellas se analizaron las fracciones 1-0.16 mm e inferior a 0.16 mm. Igualmente se analizó como muestra de control la recogida en la escombrera del cargadero de la Casa de las Minas, ya que sus valores dan una estimación de la concentración de los elementos estudiados en la fuente de la anomalía.

A partir de los datos de la extracción nítrica, que es la técnica que proporciona valores más elevados, se apreció la ausencia de diferencias importantes entre las dos fracciones para los elementos traza, a excepción de Zn y Cd. Estos elementos tienden a concentrarse en la fracción gruesa, en tanto que el Sb presenta una distribución por fracciones muy irregular. El resto de los elementos traza (Pb, Cu, As y Co) no muestran diferencias significativas, con un valor ligeramente mayor para la fracción más fina en la muestra de control CM. Si suponemos a esta muestra como indicadora del comportamiento de estos elementos traza, cabe esperar ciertas ventajas analíticas utilizando la fracción menor, por cuanto presenta en ella un cierto enriquecimiento en trazas. A esta ventaja analítica se le añade otra de tipo metodológico, como es el ahorro de una operación de molienda que supone.

5.2 Valoración de las extracciones.

El análisis de la totalidad de las muestras, para la fracción menor de 0,16 mm tras la extracción con ácido nítrico, se recoge en la tabla 1.

El resto de técnicas de extracción muestran contenidos mucho más bajos para la mayoría de los elementos, situándose muchos de ellos por debajo del límite de detección. Se comprueba así que estas extracciones tienen un efecto diferencial según el elemento traza en cuestión: Zn, Cu y As aportan mayor cantidad de valores significativos con la extracción nítrica; Sb, Cd y Co lo hacen para la extracción con oxalato, aunque en mucha menor proporción. Para Pb y Ag sólo se empleó el ataque nítrico, puesto que ambos elementos forman compuestos insolubles en la extracción con oxalato amónico.

La razón de estudiar los valores de Fe y Mn asociados a los elementos traza radica en el importante control que los oxihidróxidos de ambos elementos pueden ejercer sobre la concentración de muchos de los elementos traza. Los valores de Fe y Mn liberados mediante los ataques nítrico y de oxalato presentan tendencias de evolución muy semejantes entre sí y radicalmente distintas a las observadas en la extracción de $MgCl_2$. En estas condiciones, y asumiendo el comportamiento de las extracciones aplicadas, podemos suponer que la mayor parte del Fe y Mn analizados procede de la disolución parcial de sus oxihidróxidos.

TABLA 1.- Datos analíticos de la totalidad de las muestras, para la extracción con ácido nítrico concentrado (valores expresados en ppm).

MUESTRA	As	Sb	Zn	Pb	Fe	Mn	Cu
T001	58.00	48.30	19.80	----	13460.0	281.80	7.00
T002	8.50	----	----	3.10	12480.0	206.00	----
T003	----	----	13.60	----	11863.0	223.80	----
T004	----	----	22.60	----	13700.0	153.40	----
T005	----	----	21.20	----	14798.0	267.70	----
T006	----	----	20.50	----	13563.0	214.80	----
T007	----	----	21.00	----	14789.0	237.30	----
T008	----	----	12.20	----	9996.0	137.70	----
T009	----	0.40	17.50	----	14462.0	343.30	----
T010	----	----	15.70	----	12652.0	197.50	----
T011	534.30	----	30.10	----	18548.0	434.60	4.50
T012	1.50	----	34.70	----	17985.0	262.70	13.60
T013	2.90	0.10	20.70	----	14438.0	77.10	9.20
T101	----	0.20	24.90	----	14098.0	70.40	----
T102	1.20	----	----	35.40	16541.0	229.10	----
T201	----	----	----	15.50	17910.0	233.40	----
T202	----	----	----	14.60	15149.0	224.80	----
T203	6.60	----	----	11.70	20368.0	340.30	----
T204	2.50	0.60	----	7.30	18668.0	200.10	----
T205	22.30	0.90	1.50	33.80	20957.0	443.50	25.70
T206	----	----	----	20.80	18602.0	307.20	----
A001	38.75	----	----	23.65	21577.0	411.90	30.70
A002	22.00	----	----	19.75	16346.5	405.92	18.82
A003	23.50	----	----	13.05	14293.0	290.80	9.75
A004	28.00	----	----	28.10	19342.5	445.60	27.30
A005	26.75	----	----	25.18	20191.0	546.40	16.00
A006	24.00	----	0.30	25.90	19440.0	915.40	15.00
A007	19.50	----	----	29.95	19156.8	454.00	17.30
A008	30.00	----	9.20	53.00	19775.0	502.00	14.40
A009	41.50	----	8.00	66.55	24526.5	766.30	65.20
A010	----	0.40	30.30	----	11534.0	159.80	----
A011	----	----	57.80	19.10	17239.0	307.10	9.60
A012	----	0.20	71.90	31.80	19268.0	465.10	8.00
A013	----	----	63.30	18.30	17966.0	460.00	6.10

Los valores significativamente por debajo del límite de detección se indican con ----.

Los límites de detección especificados son:

Fe - 100 ppm

Mn - 10 ppm

Ag - 5 ppm

As, Sb, Zn, Pb y Cu - 2 ppm

La extracción con $MgCl_2$ tan sólo consigue valores significativos para Fe, Mn y Cu, en tanto que el ataque con $NaClO$ lo hace para As y Cu.

Son abundantes los casos bibliográficos que registran este hecho. Rose y Suhr (1971) estudiaron las fuentes de variación del fondo geoquímico sobre los sedimentos de 40 muestras aluvionares en Pensilvania, comparando el contenido de elementos traza en óxidos de Fe, materia orgánica y cationes de cambio de arcillas. Estos autores concluyeron que los mayores contenidos de aquellos elementos se correspondían con altos contenidos en óxidos de Fe, fase que explicaba la mayor parte de la variabilidad de los elementos traza. Brundin y Nairis (1972) analizaron sedimentos aluvionares, material orgánico y aguas en una campaña

geoquímica en el NW de Suecia; realizaron un análisis factorial sobre los resultados obtenidos en su campaña con el fin de interpretar las intercorrelaciones entre los elementos relacionados en los tres ambientes y extrajeron dos factores dominantes: el primero (Mn-Ca- Zn-Fe-Sr-Ni) lo interpretaron como reflejo de la precipitación de oxihidróxidos de Fe y Mn (limonita); el segundo factor (Ba-Ca-Mg- Cu-Zn-Pb-Mn-Fe-Ni) lo interpretaron como reflejo de la materia orgánica presente.

Estas consideraciones sugieren una importante variación del fondo geoquímico en relación con el ambiente de muestreo que conviene estudiar. Discerniremos para ello entre el Arroyo de Alpartir por un lado, y el Barranco de la Tejera y del Ontanat por otro. El Arroyo de Alpartir presenta un estado de evolución mayor, con un área de drenaje mucho más amplia y un lecho más rico en finos, permaneciendo activo prácticamente durante todo el año. El Barranco de la Tejera, de menor desarrollo y actividad, presenta un material de relleno más heterométrico; su interés reside en el estudio de la posible anomalía y su dispersión en torno a la Casa de las Minas.

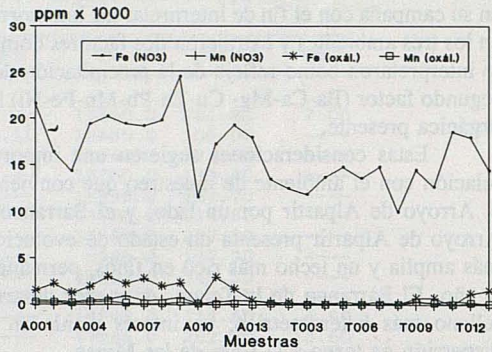
El análisis de los resultados de esta primera campaña lo vamos a centrar en las extracciones con nítrico concentrado y oxalato amónico por ser, como ya se vio anteriormente, las únicas que ofrecen patrones comparativos para todos los elementos en que estamos interesados.

En el Arroyo de Alpartir (figura 2), la extracción con oxalato no permite visualizar anomalías para ningún elemento, y únicamente el Sb presenta valores mayores que los obtenidos por ataque nítrico. Para este último ataque sí que es posible observar diferentes comportamientos según el elemento en cuestión. Así, el Zn presenta una anomalía para la extracción nítrica a partir de la muestra A009, no detectada por el ataque de oxalato, lo que indujo a pensar que estaba relacionada con la presencia de minerales detríticos. Por su parte, As, Cu y Pb muestran una tendencia similar a las pautas marcadas por Fe y Mn, con un máximo marcado para la muestra A009 y un mínimo desarrollado a partir de la muestra A010. En especial, el Pb presenta una curva muy similar a la del Fe.

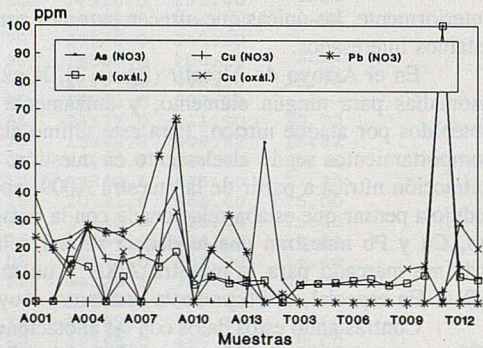
Contrastando estos datos con las anotaciones de campo, hemos podido comprobar que 100 mts aguas arriba de la muestra A009 se localiza una pequeña represa sobre el arroyo que altera la carga de sedimentos. Parece claro, por tanto, que este tipo de acciones, que conducen a alteraciones de la dinámica del arroyo, tienen una influencia notable sobre el proceso de dispersión de los elementos. Las formas poco móviles, como el Zn asociado a minerales detríticos, se concentran aguas arriba de la represa (hay que señalar que el Zn es uno de los pocos elementos que presentan valores sensiblemente superiores en la fracción gruesa que en la fina), en tanto que As y Cu, más móviles, pueden ser lavados y dispersarse aguas abajo de la represa.

En el Barranco de la Tejera se obtienen, en general, menores valores de concentración para los elementos que en el Arroyo de Alpartir (figura 2). El Sb tan sólo ofrece valores significativos con la extracción de oxalato, al igual que ocurría en el arroyo anterior, en tanto que para el Pb no se registran concentraciones medibles. En este caso, puede detectarse una anomalía mediante ambas extracciones, marcándose a partir de la muestra T011, situada aguas abajo de la Casa de las Minas. Esta anomalía presenta mejor contraste para Sb (con extracción de oxalato) y para Cu (con ambas extracciones). Este resultado parece acorde con el tipo de paragénesis registrado para este yacimiento, incluido en el tipo Sb-Cu(Ag) (García Gil *et al.*, 1988; Osacar y Besteiro, 1988).

Fe-Mn



As-Cu-Pb



Sb-Zn

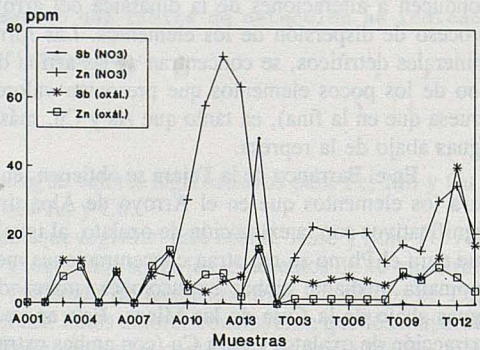


Figura 2.- Representación de los valores analíticos obtenidos mediante las extracciones nítrica y oxálica, para las muestras del A° de Alpartir y del B° de la Tejera.

Por otra parte, se comprobaron los resultados obtenidos puntualmente con las muestras de campo para las que se suponía abundancia de materia orgánica, sin llegar a resultados positivos. Si tenemos en cuenta los resultados obtenidos con la extracción con hipoclorito sódico (para materia orgánica) en las muestras de control, parece indicarse que la extracción sobre materia orgánica hubiera tenido más éxito con la fracción 1-0.16 mm, aunque los pocos datos disponibles no permiten mayores aseveraciones.

La conclusión que parece desprenderse de estos datos es que la movilidad de estos elementos está muy controlada por la presencia de óxihidróxidos de Fe y Mn, especialmente el Pb (que es el elemento teóricamente menos móvil de los investigados). En cuanto al Zn, con una movilidad relativamente alta, presenta ciertas anomalías que interpretamos correspondientes a la presencia de minerales detríticos. As y Cu se presentan igualmente asociados de forma patente a estos mismos óxidos.

5.3 Planificación de la red de muestreo.

Existen métodos estadísticos de tratamiento de datos que permiten una cuantificación objetiva para valorar la densidad de muestreo apropiada. Dijkstra (1976) propone un método, basado en la construcción de covariogramas, que informa sobre el intervalo de correlación significativa para muestras sucesivas. Al aplicar esta técnica de autocorrelación para las muestras de esta campaña obtuvimos resultados poco significativos, con reducida intercorrelación entre las muestras, que en este caso es atribuible a la escasez de puntos de muestreo disponibles en los distintos barrancos, lo que probablemente convierte a la variabilidad analítica y de muestreo en un importante agente introductor de aleatoriedad.

Enfocando el problema desde otro punto de vista, planteamos la utilización de las ideas de Hawkes (1976). Este autor propone un modelo teórico que relaciona el contenido de una muestra aluvial anómala (C_{ca}) y el área de drenaje aguas arriba (A_a), con el grado de mineralización (C_{cm}) y superficie expuesta a erosión (A_m) del área mineralizada, suponiendo un fondo geoquímico (C_{cf}) conocido y constante. Propone la siguiente fórmula:

$$C_{cm} A_m = A_a (C_{ca} - C_{cf}) + A_m C_{cf}$$

Este modelo da cuenta de la dilución en la concentración de un elemento aguas abajo por aporte de material estéril, lo que se traduce en un progresivo menor contraste de la anomalía ($C_{ca} - C_{cf}$), hasta que llega a confundirse con el fondo geoquímico.

Hawkes indica una interesante aplicación a su fórmula de dilución, como es la planificación de la red de muestreo. Puede resolverse la ecuación anterior para A_a , el máximo tamaño de la cuenca de drenaje en la cual todavía puede reconocerse la anomalía:

$$A_a = A_m (C_{cm} - C_{cf}) / (C_{ca} - C_{cf})$$

Teniendo en cuenta la mineralización de Cu de Alpartir como fuente de la anomalía, estudiaremos la aplicación de la fórmula para este elemento con la extracción nítrica, que en el anterior apartado se reveló como la más significativa. Las condiciones locales de este antiguo yacimiento sugieren que la escombrera de la mina es actualmente la principal fuente de la anomalía, por lo cual utilizaremos el valor de la muestra de control CM como parámetro C_{cm} de la fórmula.

El área mineralizada A_m (escombrera) tiene unos 100 m^2 , con un contenido en Cu para $C_{em} = 3413.2 \text{ ppm}$. Tomando como fondo geoquímico del Cu en el Barranco de la Tejera 2 ppm , puede definirse, en una primera aproximación, como valor anómalo crítico dos veces el fondo geoquímico local: $C_{ca} = 4 \text{ ppm}$. En estas condiciones:

$$A_a = 10^4 (3413.2 - 2)/(4 - 2) = 0.17 \text{ km}^2$$

Un área de 0.17 km^2 , en términos de distancia linear a lo largo del arroyo supone unos 340 m . Valor que debe interpretarse como límite superior aconsejable en la densidad de muestreo para detectar anomalías con las indicaciones de contraste anteriormente supuestas.

Elijiendo unas condiciones de contraste mayores, por ejemplo, de 2 ppm como fondo y 6 ppm como valor anómalo crítico, se obtiene un resultado $A_a = 0.08 \text{ km}^2$ o, en términos de longitud linear a lo largo del arroyo, 170 mts .

La elección del contraste investigado y, por tanto, la densidad de muestreo apropiada, supone un compromiso entre la efectividad del método en términos de magnitud de los contrastes discernibles y el coste en cuanto a cantidad de muestras necesarias. En este sentido, la elección de un contraste crítico de dos veces el valor del fondo geoquímico ($C_{cf} = 2 \text{ ppm}$ y $C_{ca} = 4 \text{ ppm}$) supone un espaciado máximo de muestreo de 340 mts . Ambas condiciones se estiman razonables dentro de los objetivos perseguidos en la campaña.

El resultado de 340 mts como máximo espaciado de muestreo es fiable en la medida en que el caso que nos ocupa se acerque al modelo teórico de Hawkes. De entre las premisas teóricas del modelo existe una que puede afectar particularmente a esta campaña, la variabilidad analítica y de muestreo, y que ya se indicó al comienzo del apartado como posible responsable de la escasa intercorrelación entre las muestras. Tales circunstancias aconsejan valorar el anterior resultado como optimista y optar por una densidad algo menor en la medida en que el coste presupuestario lo aconseje. De cualquier forma, y habida cuenta que la elección del fondo geoquímico y contraste son en cierto modo subjetivos y locales, el resultado es válido sólo como indicación aproximada.

6. CONCLUSIONES

La aplicación de las extracciones selectivas sobre los sedimentos aluvionares ha aportado resultados que pueden interpretarse para valorar la forma de presentación de los elementos dispersos, y la información precisa para la elección del método de extracción más apropiado para la campaña global.

En cuanto a la granulometría de la fracción a investigar, se optó por la inferior a 0.16 mm por las ventajas analíticas (valores algo mayores para gran parte de los elementos) y metodológicas (menor tiempo de preparación puesto que evita una molienda).

Los elementos investigados, presentes en la mineralización, en principio dan suficientes valores significativos para As, Sb, Zn, Pb, Cu, Fe y Mn. Se desestimó el análisis de Ag, Cd y Co, que en esta campaña previa no han aportado valores operativos, ya que el fondo general de los mismos se encuentra situado por debajo del límite instrumental de detección.

En lo que refiere a su forma de presentación, parecen encontrarse asociados principalmente a oxihidróxidos de Fe y Mn, y en menor proporción al resto de las posibles

formas apuntadas anteriormente. La fracción orgánica retiene sólo pequeñas cantidades de Cu y As, y de las superficies de cambio de las arcillas sólo es posible extraer, con la metodología utilizada, reducidas cantidades de Fe, Mn y Cu. Se consideró por tanto apropiado, ya que ninguna de las extracciones selectivas daba por sí sola valores suficientemente significativos, emplear para la campaña global la extracción con ácido nítrico concentrado, menos selectiva.

En cuanto a la densidad de muestreo, teniendo en cuenta las consideraciones del anterior apartado, se optó por un espaciado entre muestras entre 250-300 mts, apropiado en cuanto a la relación coste/efectividad se refiere. En principio parece aconsejable mantener este espaciado de muestreo lo más constante posible, siempre que las condiciones de accesibilidad al lecho así lo permitan.

Además, se sugieren ciertas pautas de comportamiento para los elementos que, de todas formas, son aún provisionales, habiéndose de comparar con los resultados obtenidos en la campaña global. Los bajos valores obtenidos para el Sb se explican por su alta movilidad, que propicia su lavado de los sedimentos del lecho. Por su parte, el Zn presenta un modelo de dispersión relacionado en parte con la presencia de minerales detríticos. Este hecho aconseja aplicar algún tipo de control analítico del submuestreo en la campaña global, con el objeto de contrastar posibles inhomogeneidades en las muestras. En cuanto a As, Cu y Pb, en especial este último, muestran valores de concentración muy relacionados con el contenido en Fe y Mn. Además, el Fe parece controlar de forma más patente el contenido en elementos traza que el Mn, hecho que probablemente se debe a que se presenta en mucha mayor proporción.

AGRADECIMIENTOS

Este trabajo se ha podido realizar gracias a la ayuda concedida por el Vicerrectorado de Investigación de la Universidad de Zaragoza, dentro del Programa de Apoyo a la Investigación (Proyecto 221-14).

BIBLIOGRAFIA

- Brundin, N.H. and Nairis, B. (1972): *J. Geochem. Explor.*, 1, 7-46.
- Cole D.P. and Rose, A.W. (1984): *J. Geochem. Explor.*, 20, 137-160.
- Chao, T.T. (1984): *J. Geochem. Explor.*, 20, 101-135.
- Dijkstra, S. (1976): *Geol. Mijnbouw*, 55, 105-109.
- Dijkstra, S.; Van den Hull, H.J.; Jubeli, Y.M.; Ramanamurthy, M.V. and Sen, T.K. (1987): *J. Geochem. Explor.*, 29, 247-258.
- Fletcher, W.K. (1981): *Handbook of Exploration Geochemistry*, V.2, Ed. Elsevier, 462 pp.

- García Gil, E.; Vindel E. y Lunar, R. (1988): Bol. Geol. Min., 99, 88-99.
- Gibbs, R.J. (1973): Science, 180, 71-73.
- Govett, G.J.S. (1983): Handbook of Exploration Geochemistry, V.3, Ed. Elsevier, 461 pp.
- Hawkes, H.E. (1976): J. Geochem. Explor., 6, 345-358.
- Miesch, A.T. (1976): U.S. Geol. Surv. Prof. Paper, 954-A, 39 pp.
- Osacar, M.C. y Besteiro, J. (1988): Bol. Soc. Esp. Min., 11, 129-132.
- Rose, A.W. (1975): Vancouver Symposium Volume, 691-705.
- Rose, A.W.; Hawkes, H.E. and Webb, J.S. (1979): Academic Press, 657 pp.
- Rose, A.W. and Shur, N.H. (1971): Toronto Symposium Volume, 587-593.
- Villareal, A.R. (1991): Tesis Licenciatura, Univ. Zaragoza, 211 pp., inédita.

LAS SURGENCIAS TERMALES DE LOS BAÑOS (TERUEL). I. CARACTERIZACION FISICOQUIMICA Y APROXIMACION A LOS PROCESOS DE INTERACCION AGUA-ROCA.

Auqué, L.F.; Tena, J.; Gimeno, M.J.; Mandado, J. y Fernández, J.

Departamento de Geología. Facultad de Ciencias. Universidad de Zaragoza. 50009 ZARAGOZA.

ABSTRACT

Springs of *Los Baños* (Teruel) are manifestations of a low enthalpy thermal system (21 °C of surface temperature) with sulfate-calcium type waters, associated with carbonatic-evaporitic materials and residence time 30 years up.

Relationships between dissolved concentrations of Ca^{2+} , Mg^{2+} , and HCO_3^- enable a distinction to be made between the principal carbonate mineral reactions. Dissolution of gypsum and/or anhydrite, and incongruent dissolution of dolomite are the main water-rock interaction processes deduced for the system.

Speciation-solubility calculations (with PHREEQE code) show that these waters are near saturation (equilibrium) with calcite and undersaturated with dolomite and gypsum/anhydrite, at surface conditions. These results are consistent with the dissolution processes proposed earlier.

1. INTRODUCCION

El estudio de los recursos termales y mineromedicinales de la provincia de Teruel ha sido objeto de distintos trabajos (IGME, 1982; Fernández, 1987; Fernández et al., 1988) de ámbito general y normalmente enfocados hacia un análisis directo de prospección y evaluación del potencial geotérmico.

Los resultados obtenidos en estos trabajos han puesto de manifiesto la existencia de recursos y sistemas de baja/muy baja entalpía, sin delimitación específica de posibilidades dentro de esta tipología. Además se ha comprobado que las técnicas geotermométricas clásicas utilizadas en prospección geotérmica muestran numerosas deficiencias al ser aplicadas a este tipo de aguas, de baja entalpía y asociadas a reservorios sedimentarios.

Cualquier intento de precisar el potencial geotérmico de estas surgencias pasa, por tanto, por un análisis más detallado de sus rasgos geoquímicos y de los procesos de interacción agua-roca que controlen el quimismo de estas soluciones. Ello se traduce en la necesidad de definir las condiciones fisicoquímicas de las mismas, inicialmente en condiciones de surgencia, para lo que resulta indispensable la utilización de técnicas de modelización geoquímica.

El objetivo de este trabajo se centra en la realización de un análisis detallado de las características geoquímicas y fisicoquímicas de los manantiales de Los Baños de Teruel, uno de los puntos termales-mineromedicinales de la provincia a priori más interesantes. Se hará una especial consideración en la definición de los posibles procesos de interacción agua-roca responsables del quimismo de estos manantiales, así como en el cálculo de las posibles relaciones de equilibrio entre las aguas y distintos minerales.

2. SITUACION GEOGRAFICA Y GEOLOGICA

A 6 km de Teruel, tomando la carretera de Alfambra, se encuentra la antigua estación de ferrocarril de Los Baños (Hoja topográfica nº 567) y junto a ésta lo que en tiempos fue el balneario de Los Baños de Teruel, actualmente en ruinas. Se trata de un conjunto de manantiales con un caudal de 8 l/s y una temperatura de 20.7 a 21.4°C. Las coordenadas topográficas de estos manantiales son: latitud 40° 22' 48.6'', longitud 01° 05' 49'' W y altitud 920 m.

Geológicamente, los Baños de Teruel se sitúan en la convergencia de las fosas del Jiloca y de Alfambra, ligados al accidente principal de la zona representado por la falla de Conclud, de dirección NO-SE (fig. 1). Este accidente pone en contacto los materiales del Muschelkalk con los del Rethiense-Hettangiense, y al Buntsandstein con sedimentos miopliocenos. Los puntos de surgencia de estos manantiales, a escasos metros unos de otros, se sitúan en el aluvial del río Alfambra, en las proximidades de materiales carbonatados jurásicos afectados por la citada falla.

Los materiales terciarios y cuaternarios son los que alcanzan una mayor extensión en esta zona, aunque es en las litologías del Triásico Superior y Jurásico aflorantes en el bloque levantado de la Fosa del Jiloca (al N de las poblaciones de Conclud y Caudé), donde posiblemente se sitúa el acuífero de estos manantiales (Fernández, 1987). Estos materiales forman un sinclinal de dirección NO-SE limitado en su flanco sur por la mencionada falla de Conclud, a favor de la cual parecen situarse las surgencias. Los materiales impermeables, arcillas, margas y yesos del Keuper, no llegan a aflorar pero previsiblemente pueden constituir la base de este acuífero.

No pueden realizarse mayores precisiones acerca de los materiales jurásicos que constituyen en concreto el acuífero de los manantiales de Los Baños. Estas litologías jurásicas, pese a presentar materiales con diferente comportamiento hidrológico (fig. 1), pueden constituir un único acuífero indiferenciado debido a la previsible interconexión de las distintas capas acuíferas individualizadas, a favor de la fracturación y carstificación existente.

Estudios previos realizados por IGME (1982) han puesto de manifiesto, mediante la realización de análisis de tritio en la surgencia del balneario (con contenidos de 0.0 ± 2.1 U.T.¹), tiempos de residencia superiores a 35 años para estas soluciones en el acuífero, además de señalar la inexistencia de procesos de mezcla de estas aguas con soluciones más superficiales y frías.

¹U.T., unidades de tritio.

BAÑOS DE TERUEL

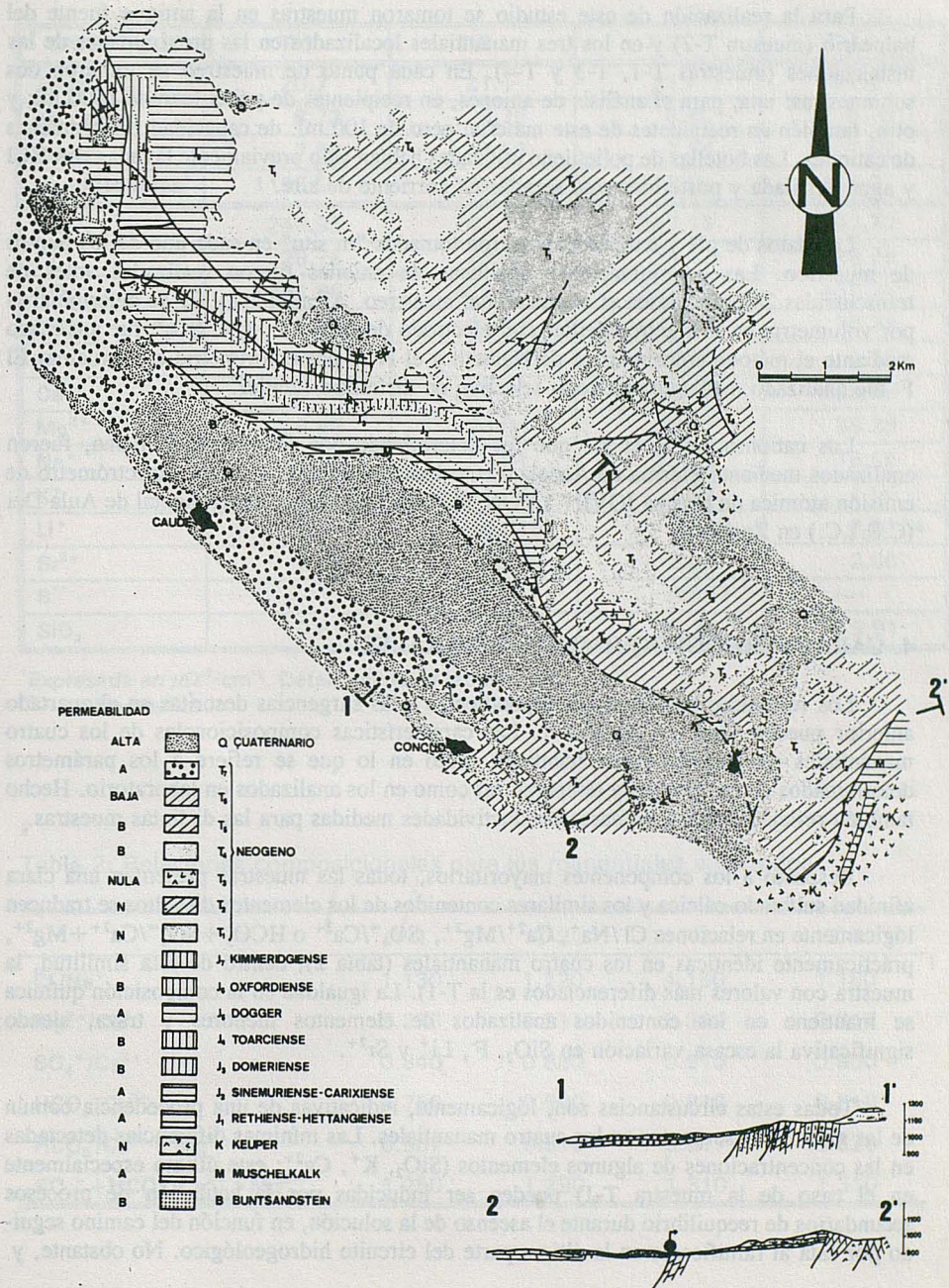


Figura 1. Situación de Los Baños de Teruel. Mapa y cortes geológicos de la zona (modificado de IGME, 1983).

3. METODOLOGÍA ANALÍTICA Y DE MUESTREO

Para la realización de este estudio se tomaron muestras en la antigua fuente del balneario (muestra T-2) y en los tres manantiales localizados en las proximidades de las instalaciones (muestras T-1, T-3 y T-4). En cada punto de muestreo se tomaron dos submuestras: una, para el análisis de aniones, en recipientes de polietileno de 1000 ml; y otra, también en recipientes de este material pero de 100 ml. de capacidad, para análisis de cationes. Las botellas de polietileno utilizadas habían sido previamente lavadas con HCl y agua destilada y posteriormente secadas en corriente de aire.

Los datos de pH y temperatura fueron tomados "in situ" en cada uno de los puntos de muestreo. Las determinaciones analíticas de aniones fueron realizadas antes de transcurridas 24 horas desde el momento del muestreo. Alcalinidad y Cl⁻ se determinaron por volumetría en titrmetro Mettler con electrodo de punto final; el SO₄⁼ fue analizado mediante el método colorimétrico de Nemeth y el NO₃⁻ por colorimetría con brucina. El F⁻ fue analizado mediante electrodo selectivo en titrmetro Mettler.

Los cationes, salvo el K⁺ que fue determinado por fotometría de llama, fueron analizados mediante Plasma de Acoplamiento Inducido (I.C.P.) con un espectrómetro de emisión atómica de plasma JOVIN-YVON 38, en el laboratorio experimental de Aula Dei (C.S.I.C.) en Zaragoza.

4. CARACTERES GEOQUÍMICOS GENERALES

Los resultados analíticos correspondientes a las surgencias descritas en el apartado anterior pueden verse en la tabla 1. Las características composicionales de los cuatro manantiales son prácticamente idénticas, tanto en lo que se refiere a los parámetros determinados en campo (pH y temperatura) como en los analizados en laboratorio. Hecho también reflejado en las similares conductividades medidas para las distintas muestras.

Respecto a los componentes mayoritarios, todas las muestras presentan una clara afinidad sulfatado-cálcica y los similares contenidos de los elementos disueltos se traducen lógicamente en relaciones Cl/Na⁺, Ca²⁺/Mg²⁺, SO₄⁼/Ca²⁺ o HCO₃⁻+SO₄⁼/Ca²⁺+Mg²⁺, prácticamente idénticas en los cuatro manantiales (tabla 2); dentro de esta similitud, la muestra con valores más diferenciados es la T-1). La igualdad en la composición química se mantiene en los contenidos analizados de elementos menores y traza, siendo significativa la escasa variación en SiO₂, F, Li⁺ y Sr²⁺.

Todas estas circunstancias son, lógicamente, indicativas de una procedencia común de las soluciones surgentes en los cuatro manantiales. Las mínimas diferencias detectadas en las concentraciones de algunos elementos (SiO₂, K⁺, Ca²⁺; este último especialmente en el caso de la muestra T-1) pueden ser inducidas por la actuación de procesos secundarios de reequilibrio durante el ascenso de la solución, en función del camino seguido por ésta al ramificarse en la última parte del circuito hidrogeológico. No obstante, y

Tabla 1: Análisis químicos de los manantiales de Los Baños de Teruel. Las concentraciones están expresadas en mgr/l.

MUESTRAS	T-1	T-2	T-3	T-4
Temp. (°C)	20.7	21.1	21.4	21.4
pH campo	6.72	6.79	6.80	6.78
pH lab.	7.55	7.56	7.54	7.62
Conductividad*	1128.0	1135.0	1120.0	1113.0
HCO ₃ ⁻	248.95	247.73	247.12	247.73
Cl ⁻	29.78	29.42	29.42	29.78
SO ₄ ⁼	441.86	427.45	437.06	432.26
NO ₃ ⁻	16.31	17.17	17.17	16.74
F ⁻	0.551	0.551	0.551	0.551
Ca ²⁺	218.03	201.60	198.80	198.80
Mg ²⁺	36.58	38.04	38.77	38.28
Na ⁺	17.01	17.01	16.55	16.55
K ⁺	1.37	1.84	1.37	1.56
Li ⁺	0.019	0.0187	0.019	0.0184
Sr ²⁺	2.06	2.06	2.06	2.06
B ^{**}	----	----	0.857	----
SiO ₂	9.61	9.91	10.21	9.91

*Expresada en $\mu\Omega^{-1}\cdot\text{cm}^{-1}$. Determinada en laboratorio.

** ----- No analizado.

Tabla 2: Relaciones composicionales para los manantiales analizados.

	T-1	T-2	T-3	T-4
Cl/Na ⁺	1.130	1.120	1.150	1.160
Mg ²⁺ /Ca ²⁺	0.276	0.311	0.321	0.317
SO ₄ ⁼ /Ca ²⁺	0.840	0.880	0.910	0.900
HCO ₃ ⁻ /Ca ²⁺	0.750	0.800	0.816	0.818
HCO ₃ ⁻ /Ca ²⁺ + Mg ²⁺	0.587	0.610	0.617	0.620
SO ₄ ⁼ + HCO ₃ ⁻ /Ca ²⁺ + Mg ²⁺	1.250	1.290	1.310	1.310

en la mayoría de los casos, las variaciones discernibles en las concentraciones analizadas se encuentran dentro del propio rango de error analítico.

La homogeneidad química de estos manantiales, su parecida temperatura de surgencia y sus similares valores de pH, sugieren la inexistencia de procesos secundarios de otro tipo (mezcla con soluciones más frías, pérdida de CO₂, etc.) o al menos de efectos selectivos (con distinta intensidad) durante el tramo final del circuito hidrogeológico de estas aguas.

Y, en general, estas características geoquímicas son indicativas de una circulación de las soluciones a través de litologías predominantemente sulfatado-carbonatadas, hecho que permite relacionarlas con los materiales jurásicos próximos a las surgencias.

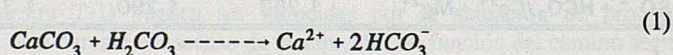
5. PROCESOS DE INTERACCION: DELIMITACION Y EVALUACION

Las características geoquímicas generales que acabamos de señalar delimitan, implícitamente y a grandes rasgos, los procesos de interacción agua/roca responsables de ese quimismo. Una discusión más detallada sobre esos procesos debe partir, necesariamente, de una idealización de las fases minerales involucradas. Teniendo presente las características litológicas de los materiales factiblemente relacionados con la circulación de esas aguas, pueden incluirse en el análisis calcita, dolomita, yeso y/o anhidrita.

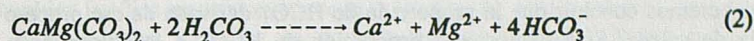
Los procesos de reacción de una solución con minerales carbonatados como los que acabamos de mencionar están implícitamente reflejados en el pH y concentración de Ca²⁺, Mg²⁺ y HCO₃⁻ de esa solución. En el caso de los manantiales que nos ocupan, el evidente proceso de disolución de yeso y/o anhidrita, responsable del carácter sulfatado de las aguas, también contribuirá a la cantidad total de Ca²⁺ en disolución.

Admitiendo que la contribución en moles de Ca²⁺ procedente de la disolución de yeso y/o anhidrita es la misma que la de SO₄⁼, puede calcularse la proporción de aquél elemento derivada exclusivamente de la interacción con carbonatos restando la concentración de SO₄⁼ de la de Ca²⁺ total en solución. En la tabla 3 se indican los resultados obtenidos, en los que claramente se pone de manifiesto la mayor contribución de los sulfatos (10 veces superior como término medio) que la de los carbonatos al Ca²⁺ total disuelto.

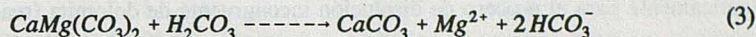
Planteada esta corrección para el Ca²⁺, y suponiendo que las concentraciones de Mg²⁺ y HCO₃⁻ derivan directamente de la interacción de las aguas con materiales carbonatados constituídos por calcita y dolomita, podemos utilizar las relaciones Ca²⁺:Mg²⁺:HCO₃⁻ de las soluciones para delimitar las características de esos procesos de interacción. Las reacciones de disolución congruente de calcita y dolomita pueden escribirse:



y



mientras que la reacción de disolución incongruente de dolomita viene definida por:



Según estas ecuaciones, la disolución congruente de calcita daría lugar a relaciones $m\text{Ca}^{2+}/m\text{HCO}_3^- = 1 : 2$ en la solución, mientras que para la disolución congruente e incongruente de dolomita (reacciones 2 y 3) serían: $m\text{Ca}^{2+} : m\text{Mg}^{2+} : m\text{HCO}_3^- = 1:1:4$ y $0:1:2$, respectivamente.

Las concentraciones molales de estos elementos obtenidas en los manantiales de Los Baños pueden verse en la tabla 3, en la que también se indican en cada caso los contenidos normalizados de Ca^{2+} , Mg^{2+} y HCO_3^- respecto a la proporción de Mg^{2+} . Como puede comprobarse estas últimas relaciones no se ajustan a ninguna de las proporciones idealizadas derivadas de las reacciones (1), (2) o (3).

La desproporción de Mg^{2+} frente a Ca^{2+} implica inicialmente un predominio de la interacción de estas aguas respecto a materiales dolomíticos; y atendiendo a las estequiometrías definidas, nos situaríamos próximos a un proceso de disolución incongruente de dolomita, aunque con una proporción de HCO_3^- y Ca^{2+} ligeramente superior a la deducida para el correspondiente proceso idealizado.

Tabla 3: Concentraciones molales de Ca^{2+} , Mg^{2+} y HCO_3^- para las muestras consideradas y valores normalizados para la proporción de Mg^{2+} correspondiente (ver texto).

MUESTRA	$m\text{Ca}^{2+}$ ⁽¹⁾	$m\text{Mg}^{2+}$	$m\text{HCO}_3^-$	Ca^{2+}	Mg^{2+}	HCO_3^-
T-1	$0.84 \cdot 10^{-3}$	$1.50 \cdot 10^{-3}$	$4.08 \cdot 10^{-3}$	0.50	1.00	2.72
T-2	$0.58 \cdot 10^{-3}$	$1.56 \cdot 10^{-3}$	$4.06 \cdot 10^{-3}$	0.37	1.00	2.60
T-3	$0.41 \cdot 10^{-3}$	$1.59 \cdot 10^{-3}$	$4.05 \cdot 10^{-3}$	0.26	1.00	2.54
T-4	$0.46 \cdot 10^{-3}$	$1.57 \cdot 10^{-3}$	$4.06 \cdot 10^{-3}$	0.30	1.00	2.60

⁽¹⁾ Concentración corregida sustrayendo los moles de SO_4^{2-} analizados en cada manantial.

Sin embargo, si suponemos que la proporción excedente de Ca^{2+} indicada procede de la disolución congruente de calcita, según la estequiometría definida por la reacción (1) deberemos concluir que la proporción de HCO_3^- derivada de ese proceso es el doble que la de calcio. Si restamos esta proporción de HCO_3^- a la determinada analíticamente, obtendremos la teórica contribución de un proceso de disolución de dolomita a la concentración de este anión. Las relaciones $\text{mMg}^{2+}:\text{mHCO}_3^-$ así obtenidas son de 1:2.2, 1:1.9, 1:2.0 y 1:2.0 para cada muestra, prácticamente coincidentes con las deducidas teóricamente para el proceso de disolución incongruente de dolomita (reacción 3).

Las estequiometrías minerales utilizadas como referencia teórica en la identificación de este proceso corresponden, tal como se advirtió al principio de este apartado, a fórmulas idealizadas simples. Lógicamente, la intervención en el proceso de calcita magnesiana o de dolomita de distinta estequiometría a la considerada supondría una modificación de las relaciones obtenidas.

En el sistema estudiado no ha podido identificarse la estequiometría real de las fases carbonatadas involucradas; no obstante, los resultados obtenidos parecen indicar que, en cualquier caso, los posibles efectos debidos a la existencia de estequiometrías no ideales se compensan, suministrando valores comparables a los de las reacciones teóricas utilizadas para la discusión.

El proceso de disolución incongruente de dolomita identificado, debería producirse teóricamente (Stumm & Morgan, 1981; Drever, 1986) en un sistema acuoso que hubiese alcanzado la saturación respecto a la calcita y que, por tanto, se encontrara en un estado de equilibrio termodinámico (o en un estado asimilable al de equilibrio termodinámico como por ejemplo una situación de estado estacionario) respecto a este mineral. La verificación de este hecho requiere la definición de las condiciones fisicoquímicas de la solución mediante cálculos de especiación-solubilidad.

6. CARACTERIZACION TERMODINAMICA DE LAS SOLUCIONES: INDICES DE SATURACION.

Los modelos de especiación-solubilidad permiten el cálculo de la distribución de especies en disolución a través de su análisis químico y de los parámetros de temperatura y pH, así como la determinación de su estado de saturación respecto a determinados minerales. Este estado de saturación se expresa normalmente en función del parámetro denominado *Indice de Saturación* (I.S.; por ej. Stumm & Morgan, 1981; Truesdell & Jones, 1974), que es el utilizado en este trabajo.

En la tabla 4 se expresan los I.S. obtenidos respecto a distintos minerales para los cuatro manantiales analizados, calculados mediante el código PHREEQE (Parkhurst et al., 1980). Como puede apreciarse, y pese al carácter predominantemente sulfatado de estas aguas, los I.S. respecto a yeso y anhidrita indican la existencia de marcadas condiciones de subsaturación. Ello, lógicamente, implica que la capacidad de estas aguas para disolver las fases minerales señaladas no ha sido totalmente satisfecha.

El cuarzo es el único mineral de los indicados en la tabla 4 que presenta unas ligeras condiciones de sobresaturación a la temperatura de surgencia, mientras que la calcedonia (y obviamente también la sílice amorfa) se encuentra subsaturada. Los resultados obtenidos respecto a las sales solubles incluidas en la base de datos del PHREEQE indican condiciones de marcada subsaturación.

Tabla 4. Índices de saturación respecto a distintos minerales para las distintas muestras consideradas (cálculos realizados con el código PHREEQE).

	Calcita	Arag.	Dolomita	Yeso	Anhid.	Cuarzo	Calced.
T-1	-0.271	-0.418	-1.011	-0.700	-0.966	0.279	-0.221
T-2	-0.197	-0.344	-0.806	-0.738	-0.999	0.286	-0.212
T-3	-0.202	-0.344	-0.798	-0.736	-0.994	0.295	-0.203
T-4	-0.200	-0.346	-0.799	-0.740	-0.997	0.281	-0.216

Respecto a los minerales carbonatados considerados en la modelización, se observan claras diferencias en los I.S. obtenidos. Los valores para la calcita indican una situación de ligera subsaturación, mientras que los de dolomita representan una situación más marcada de subsaturación. Aceptando los rangos de error asimilables al cálculo de I.S. respecto a estos minerales (± 0.25 y ± 0.5 unidades de I.S. para calcita y dolomita respectivamente) señalados en la bibliografía sobre el tema, obtenemos para la calcita un estado prácticamente indistinguible de una situación de equilibrio (I.S. = 0.0 ± 0.25) mientras que para la dolomita sigue manteniéndose una situación de subsaturación.

Los ligeros valores de subsaturación detectados respecto a la calcita pueden, no obstante, representar una situación real de las condiciones de surgencia y no únicamente los rangos de error numéricos respecto a los que referir una situación de equilibrio ideal (I.S. = 0.0). Si, como es previsible, la temperatura en el reservorio del sistema es mayor que la de surgencia y la solución se encuentra en equilibrio respecto a la calcita en esas condiciones, un enfriamiento por conducción durante el ascenso de las aguas provocará su subsaturación respecto a la calcita (recuérdese que la solubilidad de este mineral aumenta al disminuir la temperatura).

Puede concluirse, por tanto, que las aguas se encuentran en un estado de equilibrio respecto a la calcita y, por tanto se reafirma la existencia de un proceso de disolución incongruente de dolomita como uno de los responsables del quimismo de estos manantiales.

Todos los procesos y situaciones descritos hasta el momento llevan a pensar en la presencia de un fenómeno global de dedolomitización asociado al quimismo y circulación de estas aguas. La disolución de sulfato en estas soluciones provocaría la precipitación de calcita (y por tanto una situación de equilibrio o estado estacionario de la solución respecto a este mineral) y disolución de dolomita, o lo que es lo mismo, el mencionado proceso de disolución incongruente de dolomita.

Obviamente, la existencia de este proceso únicamente puede considerarse como resultado de un balance general, deducido a partir de la composición final de los manantiales. Una mejor delimitación del mismo requeriría obtener datos analíticos de la solución en distintos puntos a lo largo de su recorrido por el acuífero para, de esta manera, establecer balances más precisos.

En cualquier caso, la existencia de este proceso de dedolomitización justificaría las buenas correlaciones obtenidas entre SO_4^- y Mg^{2+} para distintos manantiales de la provincia (p. ej. Fernández et al., 1988), sin necesidad de recurrir a la presencia de hipotéticas fases sulfatado-magnésicas y cuya existencia es siempre de difícil justificación.

7. CONCLUSIONES

Las características geoquímicas de los manantiales de Teruel, de afinidad sulfatado-cálcica, son indicativas de su circulación preponderante a través de materiales sulfatado-carbonatados. Un análisis más detallado atendiendo a las estequiometrías idealizadas de los principales minerales involucrados, pone de manifiesto la existencia de distintos procesos de disolución como responsables del quimismo de estas aguas: por un lado, y de manera dominante, el de disolución de sulfato cálcico y, por otro, un proceso de disolución incongruente de dolomita.

Los índices de saturación calculados respecto a estos minerales concuerdan con los procesos propuestos. La necesaria saturación en calcita de las aguas (para hacer efectivo el proceso de disolución incongruente de dolomita) parece alcanzarse en los cuatro manantiales analizados. Las ligeras subsaturaciones respecto a este mineral pueden incluirse dentro del rango de error asimilable a una situación de equilibrio en este tipo de cálculos, o bien puede corresponder realmente a un estado de equilibrio respecto a la calcita pero a mayor temperatura, en profundidad.

El estado de subsaturación respecto a la dolomita en las surgencias analizadas puede corresponder también a una situación de equilibrio en profundidad y a mayor temperatura (la solubilidad de este mineral respecto a la temperatura sigue idéntica pauta a la señalada para la calcita). La relación I.S. - temperatura respecto a estos minerales puede utilizarse como medio para calcular la teórica temperatura en profundidad de estas aguas (ver Tena et al., 1991, en este mismo volumen).

La proporción de Ca^{2+} y HCO_3^- , que en principio se ha asociado a un proceso de disolución de calcita en el análisis estequiométrico, no tiene por qué representar un efectivo proceso de disolución. Probablemente corresponda al producto de actividad iónica necesario para delimitar un estado de equilibrio que haya requerido, a su vez, la presencia activa de un proceso de precipitación de calcita. El efecto de ión común producido por la disolución de yeso y/o anhidrita en estas aguas debería provocar, en algún momento, su sobresaturación respecto a la calcita. Lógicamente, la situación de equilibrio detectada respecto a este mineral en las surgencias implicaría la existencia de un proceso de precipitación.

Estas características definirían la existencia de un posible proceso global de dedolomitización asociado a la circulación de estas aguas; proceso que justificaría la estrecha relación $\text{SO}_4^- - \text{Mg}^{2+}$ detectada para distintos manantiales de Teruel, sin recurrir a hipotéticas fases sulfatado-magnésicas para justificarla.

BIBLIOGRAFIA

- DREVER, J.I. (1982). *The Geochemistry of Natural Waters*. Prentice Hall, 437 pp.
- FERNÁNDEZ, J. (1987). *Ambiente geológico y caracterización geoquímica de las aguas termales en la provincia de Teruel*. Tesis de Licenciatura, Fac. Ciencias, Universidad de Zaragoza, 163 pp. (No publicado).
- FERNÁNDEZ, J.; AUQUÉ, L.F. Y MANDADO, J. (1988). Quimismo de las aguas termales y mineromediciales de la provincia de Teruel (España). *Rev. Geociencias Aveiro*, 3, 1-2, 249-264.
- IGME (1982). *Estudio de las manifestaciones termales de Extremadura-Salamanca-Aragón-Rioja, orientadas a su posible explotación como recursos geotérmicos*. Memoria Tomo II, Aragón-Rioja. (No publicado).
- IGME (1983). *Teruel (Hoja 567). Mapa Geológico de España*. E. 1:50.000. Servicio de Publicaciones del Ministerio de Industria y Energía. 1 Mem., 1 mapa.
- PARKHURST, D.L.; THORSTENSON, D.C. & PLUMMER, L.N. (1980). *PHREEQE. A computer program for geochemical calculations*. U.S. Geological Survey Water Resources Investigations, 80-96, 157 pp.
- STUMM, W. & MORGAN, J.J. (1981). *Aquatic Chemistry* (2nd Ed.). John Wiley & Sons, New York, 780 pp.
- TENA, J.; AUQUE, L.F.; MANDADO, J. y GIMENO, M.J. (1991). Las surgencias termales de Los Baños (Teruel). II. Cálculos geotermométricos. *Revista de la Academia de Ciencias de Zaragoza*, en este mismo volumen.

TRUESDELL, A.H. & JONES, B.F. (1974). WATEQ. A computer program for calculating chemical equilibria of natural waters. *Jour. Res. U.S. Geol. Surv.*, 2-2, 233-248.

AGRADECIMIENTOS

Este trabajo fue financiado en parte por una Beca de la Diputación General de Aragón (L.F. Auqué) y una ayuda del Centro de Estudios Turolenses (J. Fernández). Queremos expresar así mismo nuestro agradecimiento al Dr. Francisco Alberto de la Estación Experimental de Aula Dei (C.S.I.C.) en Zaragoza por permitir la realización de los análisis químicos.

FACTORES CONDICIONANTES DE LA REFLECTIVIDAD SISMICA DE LA CORTEZA INFERIOR

GOMEZ, Javier

Area de Petrología y Geoquímica, Dpto. de Geología, Univ. de Zaragoza,
50009 Zaragoza

ABSTRACT

The existence of a seismically layered lower crust has been, without any doubt, the most interesting new aspect imaged by the deep seismic reflection profiles. This layered character is visible in many profiles acquired all around the world (Europe, United States, Australia, etc.), and contrasts abruptly with the transparent appearance of the middle and upper crust and also of the upper mantle. The question soon arose concerning the causes of the reflectivity of the lower crust and the factors that influence its character.

This paper deals with the present state of knowledge about the physical causes of the reflectivity and the modulating factors of the final character of the reflections in the seismic profiles. The first part review the geometric-descriptive technique of interpretation, which is made in three levels: (i) reflection attributes, (ii) seismic fabrics, and (iii) reflectivity patterns. The second part is a compilation of many of the factors that affected the final aspect of the reflection field in the seismic profiles. Finally a summary of the possible causes of the reflectivity in the lower crust and a short comment on the importance of each one is made.

INTRODUCCION

Desde hace aproximadamente quince años se viene trabajando con bastante intensidad en el estudio de la corteza inferior y el manto superior mediante sísmica de reflexión profunda (Barazangi & Brown, 1986a y b, Matthews & Smith, 1987 y Leven *et al.*, 1990). Esta técnica sísmica es una extensión de la que se utiliza en la industria del petróleo para localizar y prospectar campos de petróleo y gas. En primera instancia, la única diferencia con ella es el tiempo de escucha de los receptores (geófonos o hidrófonos), que en los perfiles de prospección petrolífera es de unos 6 segundos y en los perfiles de sísmica profunda suele ser de 20 s, aunque puede llegar hasta los 30 s. Esto permite alcanzar profundidades del orden de los 80-100 km, suficientes para investigar gran parte de la litosfera continental.

Hemos dicho que, *en primera instancia*, ésta es la única diferencia. Sin embargo, más que una extensión de la técnica clásica de reflexión, se puede hablar de una técnica nueva, tanto en lo que respecta a la adquisición y el procesado de los datos como, sobre todo, en el método de interpretación de los perfiles. La mayor parte de los perfiles para prospección petrolífera se realizan sobre rocas sedimentarias (ya que son éstas las que contienen el petróleo), y en este sentido se ha elaborado un complejo cuerpo de doctrina sobre la manera de procesar e interpretar las reflexiones de las rocas sedimentarias. Sin embargo, por debajo de los primeros 15 km de corteza (en algunos lugares mucho menos), las rocas sedimentarias desaparecen y son sustituidas por rocas ígneas y metamórficas (que reciben la denominación conjunta de rocas cristalinas), para las cuales no se tiene un conocimiento

tan preciso de su comportamiento sísmico. Esto es especialmente válido para la corteza inferior, donde se carece de un control geológico preciso de su estructura, composición y estado físico. Todo ello dificulta, e incluso imposibilita, la interpretación de las reflexiones generadas en esta parte de la corteza.

Uno de los objetivos básicos de la sísmica de reflexión profunda en la actualidad es precisamente el de identificar las causas de la reflectividad de la corteza inferior. En este artículo se pretende resumir el estado actual de los conocimientos acerca de los *factores* que influyen en el carácter de las reflexiones de la corteza inferior y las *causas* físicas invocadas en la literatura para explicar su reflectividad.

REFLECTIVIDAD DE LA CORTEZA CRISTALINA

Ya desde los trabajos pioneros de Junger (1951), Shor (1955) y Dohr (1957) se pudieron observar ciertas regularidades en la distribución de las reflexiones en un perfil profundo: reflexiones coherentes y fuertes de las formaciones sedimentarias superficiales, ausencia de reflexiones (salvo algunas con buzamientos apreciables) en la corteza superior y media y reflexiones fuertes, de poca continuidad lateral, en la corteza inferior. Las reflexiones de las capas sedimentarias no eran difíciles de explicar, era algo que los geólogos y geofísicos del petróleo llevaban investigando durante más de 70 años, pero se planteó el interrogante de cuál era la causa de las reflexiones en la corteza cristalina. ¿Por qué era transparente la corteza superior?, ¿Qué eran las reflexiones inclinadas que de vez en cuando aparecían en la parte superior de la corteza?, ¿Por qué era reflectora la corteza inferior?

Las causas que influyen en la reflectividad de las formaciones sedimentarias son bien conocidas (e.g. O'Doherty & Austery, 1971), ya que han sido el objetivo fundamental de la prospección sísmica llevada a cabo por la industria del petróleo durante los últimos 70 años. Por el contrario, se sabe muy poco acerca de los parámetros que influyen en la reflectividad de las rocas cristalinas (ígneas y metamórficas) en la corteza terrestre. La información estructural de los terrenos cristalinos es muy inferior a la de las formaciones sedimentarias, lo que influye decisivamente a la hora de interpretar las reflexiones. El problema se ve agravado por el hecho de que las rocas sedimentarias suelen situarse cerca de la superficie, en los primeros 5-10 km de los perfiles sísmicos, mientras que las rocas ígneas y metamórficas del basamento cristalino se sitúan por debajo de ellas. Esto tiene una consecuencia inmediata, que es la degradación de las reflexiones registradas al aumentar la profundidad del reflector.

En la calidad de una reflexión influyen una gran cantidad de factores de índole muy diversa, algunos geológicos (características físicas del propio reflector), no geológicos la mayoría (dependientes de los sistemas de adquisición y procesado de los datos). La interacción de todos estos factores determina el aspecto final de una reflexión en el perfil sísmico, lo que se conoce como el *carácter* de una reflexión. Correlacionar el carácter de una reflexión con las causas que lo han originado es una de las tareas pendientes más prioritarias que tiene la sísmica de reflexión profunda en terrenos cristalinos. Reconocer y aislar cada una de las variables que influyen en el carácter de una reflexión permite luego operar a la inversa: dada una reflexión en el perfil sísmico y conocidas las variables del sistema, reconocer la causa que la ha provocado.

Estudiar el carácter de las reflexiones aisladas es sólo el primer paso en la comprensión de la reflectividad de la corteza cristalina. Una vez expuestos los *atributos* de una reflexión, es necesario tener en cuenta las relaciones entre reflexiones, entendidas éstas como disposiciones coherentes

de reflectores a una escala regional. Es lo que Allmendinger *et al.* (1987a) denominan *fábricas sísmicas*. Las fábricas sísmicas nos servirán, a su vez, para definir una serie de *modelos de reflectividad* por superposición vertical o lateral de distintas fábricas sísmicas en un perfil de toda la corteza (incluyendo tal vez la parte más somera del manto superior).

DESCRIPCIÓN DE LOS PERFILES DE SÍSMICA DE REFLEXIÓN PROFUNDA

La mayor parte de los resultados y conclusiones que se han obtenido del estudio de la reflectividad de la corteza cristalina han tenido como unidad básica de trabajo los perfiles de reflexión, que no son otra cosa que la representación gráfica final, según una coordenada espacial horizontal y otra temporal vertical, del comportamiento sísmico de una región de la corteza terrestre. Parece apropiado, por tanto, conocer las características de los perfiles que más se utilizan para caracterizar la reflectividad de la corteza cristalina (con el objetivo puesto en su interpretación final en términos de estructuras y procesos geológicos). Dicha descripción se hace en varios niveles, que podemos resumir en tres principales:

1. Características definibles sobre una sola reflexión: *elementos de la reflexión*.
2. Características definibles sobre un conjunto de reflexiones de igual orientación y carácter penetrativo: *fábricas sísmicas*.
3. Características definibles en una sección vertical completa del perfil: *modelos de reflectividad*.

Elementos de una reflexión.

Podemos caracterizar una reflexión sísmica por medio de los siguientes atributos: (i) amplitud, (ii) polaridad, (iii) duración, (iv) frecuencia dominante, (v) continuidad y (vi) orientación. Los cuatro primeros atributos describen la geometría de las deflexiones en una traza aislada del perfil sísmico y los dos restantes la manera en que estas deflexiones se correlacionan de una traza a otra.

El conocimiento actual de la reflectividad de la corteza cristalina (y en especial de la corteza inferior), no permite identificar por separado los parámetros que influyen en cada uno de los atributos de las reflexiones, por lo que es conveniente, desde un punto de vista descriptivo, reducir a tres los atributos característicos: (i) orientación, (ii) amplitud y (iii) continuidad lateral. Estos son los únicos atributos que permanecen en un *line drawing* realizado a partir de la sección sísmica. Un *line drawing* es el dibujo esquemático que resulta de sustituir manual o automáticamente las reflexiones de una sección sísmica por trazos lineales que reproducen únicamente la longitud y la inclinación de las reflexiones más prominentes; normalmente se añade la información referente a la amplitud de las reflexiones en el perfil original utilizando trazos de diferente grosor.

La *orientación* de una reflexión suele darse en términos cualitativos. La distinción más importante se hace entre reflexiones (sub)horizontales, características sobre todo de la capa sedimentaria más superficial y de la corteza inferior, y reflexiones inclinadas, más propias de la corteza superior cristalina y del manto superior. A su vez las reflexiones inclinadas se subdividen de una forma grosera en función de la dirección del buzamiento, que en una sección bidimensional como los perfiles sísmicos se reduce a dos posibles inclinaciones, una hacia cada lado del perfil. Así, si el perfil lleva una dirección E-W, se distingue entre reflexiones que buzaban hacia el este y reflexiones que buzaban hacia el oeste.

La *amplitud* es una medida de la fuerza con que se recibe y registra una reflexión en el perfil sísmico. Es un atributo que hay que tratar con más cuidado, ya que no se preserva, en general, durante muchos de los pasos utilizados con el software comercial de procesado. En particular, el stack CDP no conservará la información sobre la amplitud absoluta de las reflexiones si existen correcciones estáticas imperfectas, si la corrección NMO utiliza curvas no hiperbólicas o si hay variaciones laterales en la amplitud, la frecuencia o la forma de la reflexión. Otro paso habitual en el procesado de los perfiles de reflexión es la equalización de amplitudes, que reduce el peso de una traza en proporción inversa a la amplitud del ruido que la degrada. Por todo ello, cuando se estudian cuantitativamente los coeficientes de reflexión de la corteza cristalina se suele trabajar con perfiles de cobertura simple, sin stack y no procesados, con la desventaja de que entonces sólo se pueden investigar las reflexiones de mayor intensidad.

La *longitud* de una reflexión es una medida de la continuidad lateral de una reflexión correlacionable por fase entre trazas adyacentes. Es un atributo muy susceptible a las variaciones en la relación señal/ruido a lo largo de un perfil, así como a la calidad de las correcciones estáticas. La longitud de las reflexiones suele usarse como índice semicuantitativo de la reflectividad de la corteza inferior, por lo que debe tener en cuenta las variaciones de la relación señal/ruido a lo largo de un perfil.

Fábricas sísmicas.

Allmendinger *et al.* (1987a) definen las fábricas sísmicas como disposiciones coherentes de reflexiones sobre un área determinada de un perfil sísmico. Pueden considerarse como la conjunción de dos conceptos: las características de las reflexiones aisladas (longitud, inclinación, etc.) y el modo en que esas reflexiones se asocian. Según estos autores, las fábricas sísmicas son similares a sus homónimas estructurales y metamórficas en el sentido de ser penetrativas (a la escala apropiada) con orientaciones similares. Pero también advierten que no deben confundirse con ellas, es decir, no hay que suponer que una fábrica sísmica determinada se debe siempre a la presencia de una determinada fábrica estructural o metamórfica. Además, aunque se han hecho muchos esfuerzos para eliminar la influencia de la adquisición y el procesado de los datos en el carácter de las reflexiones (ver más abajo), parece probable que algunos de los constituyentes de las fábricas sísmicas no se corresponden con estructuras reales en la Tierra o no se localizan directamente debajo del perfil.

Pese a estas limitaciones, las fábricas sísmicas pueden ser un eslabón importante para interpretar los perfiles de sísmica de reflexión profunda. Se han propuesto varias clasificaciones con diferentes nomenclaturas (e.g. Allmendinger *et al.*, 1987a y b, DEKORP, 1990), pero los tipos de fábricas identificados pueden resumirse en los siguientes (figura 1):

1. *Fábrica no reflectiva.*
2. *Fábrica cruzada:* reflexiones y/o difracciones con buzamientos hacia ambos lados y ausencia de reflexiones horizontales de gran continuidad lateral (los 'cocodrilos' en la nomenclatura de DEKORP, *op. cit.*).
3. *Fábrica inclinada:* reflexiones con buzamientos predominantes en una dirección.
4. *Fábrica estratificada:* reflexiones subhorizontales de longitud variable (las 'lamelas' de DEKORP, *op. cit.*)
5. *Cúmulos de difracción.*

La inspección de gran cantidad de perfiles sísmicos ha proporcionado además una idea de las zonas de la corteza donde es más habitual cada una de las fábricas sísmicas descritas. Así, la fábrica no reflectiva, aunque puede encontrarse en cualquier punto de un perfil sísmico, aparece casi siempre en la corteza superior (si no hay rocas sedimentarias) y el manto superior. La fábrica de reflexiones cruzadas aparece sobre todo en la corteza media y en la parte superior de la corteza inferior. La fábrica inclinada unidireccional es típica de la corteza superior (cuando se asocia a fallas de bajo ángulo), pero también puede afectar a toda la corteza. La fábrica estratificada tiene dos zonas de aparición: la parte más superficial de la corteza superior (cuencas sedimentarias), en cuyo caso la continuidad lateral de las reflexiones suele ser grande, y la corteza inferior, ya sea abarcando toda su extensión vertical o concentrándose en bandas, en cuyo caso la continuidad lateral es pequeña.

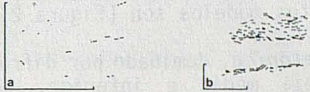
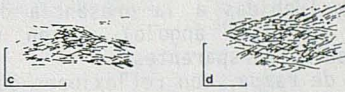

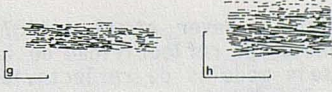
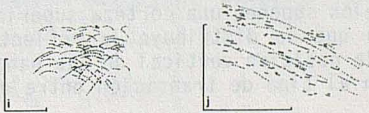
FABRICAS SISMICAS	
TIPOS	EJEMPLOS
NO REFLECTIVA	
CRUZADA	
INCLINADA	
ESTRATIFICADA	
CUMULOS DE DIFRACCION	

Figura 1. Tipos de fábricas sísmicas y ejemplos de cada una de ellas. a: perfil 2T, según Tomek et al. (1987). b: perfil DRUM, según McGeary & Warner (1985). c: perfil SWAT5, según Matthews (1987). d: perfil GLIMPCE, según Behrendt et al. (1990). e: perfil WINCH 2A, según Matthews (1986). f: perfil ECORS-Alpes, según Bois & ECORS (1990). g: perfil ECORS-Francia Norte, según Bois & ECORS (1990). h: perfil SWAT 8, según Bois et al. (1987). i: perfil DEKORP 2S, según Meissner et al. (1987). j: perfil 2T, según Tomek et al. (1987).

Modelos de reflectividad.

La identificación de modelos de reflectividad es otro intento más por clasificar la respuesta sísmica de los materiales de la corteza con el objetivo de llegar a identificar sus características. Los modelos de reflectividad se construyen a partir de las fábricas sísmicas, a las que se añade la información referente a su distribución en los perfiles. Es decir, un modelo de reflectividad incluye simultáneamente la información sobre las características de la reflectividad y sobre su distribución vertical y horizontal en los perfiles sísmicos.

Los modelos de reflectividad se pueden construir sobre una base cualitativa o cuantitativa. En el primer caso se utiliza la información sobre la continuidad lateral y la amplitud de las reflexiones para realizar line drawings esquemáticos que muestran tanto las características básicas de la reflectividad como su distribución en profundidad. En el segundo caso se construyen curvas de la frecuencia de las reflexiones en función de la profundidad (tiempos dobles); la forma de la curva y otros índices deducidos a partir de ellas sirven para clasificar los tipos de reflectividad.

Allmendinger et al. (1987b) proponen cuatro modelos cualitativos de reflectividad basándose en el estudio de los perfiles de reflexión profunda adquiridos por COCORP en una serie de líneas sísmicas que atraviesan las Rocosas, la provincia Basin and Range y el Colorado Plateau, USA, a una latitud de 40N. Estos modelos son (figura 2):

1. *Modelo cratónico*, dominado por difracciones y reflexiones inclinadas en la corteza media e inferior, pero sin un Moho de reflexión prominente.
2. *Modelo de deformación de 'piel fina'*, con reflexiones someras abundantes (debidas a la presencia de rocas sedimentarias y fallas listricas de bajo ángulo) y con una corteza media e inferior prácticamente transparentes.
3. *Modelo de rampa*, con reflexiones que buzcan en una sola dirección y que se pueden seguir desde la superficie hasta la corteza inferior, donde se pueden mezclar con reflexiones subhorizontales. El Moho de reflexión puede existir o no.
4. *Modelo estratificado*, con una corteza inferior altamente reflectora, con predominio de las reflexiones subhorizontales de continuidad lateral variable y un Moho de reflexión muy visible. La corteza superior y media son transparentes, aunque puede haber reflexiones asociadas a cuencas sedimentarias o fallas.

McGeary (1987) y Wever et al. (1987) se centran únicamente en la descripción de modelos de reflectividad de la corteza inferior. McGeary (op. cit.) distingue seis modelos de reflectividad a partir del estudio de los perfiles BIRPS en la plataforma continental que rodea las Islas Británicas. Todos los modelos suponen una corteza superior transparente y se diferencian en el modo en que se distribuye la reflectividad en la corteza inferior, basándose en la posición vertical de las bandas de reflexiones, en el número de bandas y en el tipo de transición entre ellas (brusco o gradual) (figura 3):

Tipo 1. Corteza inferior reflectiva en toda su extensión.

Tipo 2. Corteza inferior también reflectiva en toda su extensión, pero con un aumento de la densidad de las reflexiones en el Moho de reflexión.

- Tipo 3. Corteza inferior debilmente reflectiva, pero con un Moho de reflexión muy reflectivo.
- Tipo 4. Corteza inferior con dos bandas reflectivas, una a techo de la corteza inferior reflectiva y otra en el Moho de reflexión.
- Tipo 5. Transición lateral de una corteza inferior reflectiva a otra transparente, según un línea oblicua.
- Tipo 6. Corteza inferior reflectiva con variaciones en la profundidad del techo de las reflexiones.

MODELOS de REFLECTIVIDAD

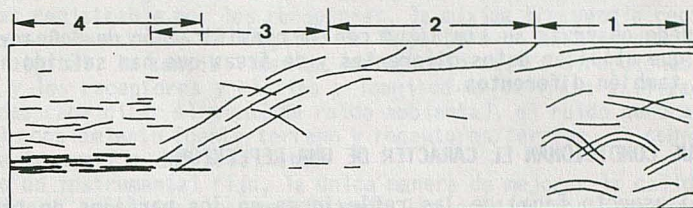


Figura 2. Modelos de reflectividad de la corteza en zonas orogénicas (según Allmendinger et al., 1987a). 1: modelo cratónico. 2: modelo de deformación de 'piel fina'. 3: modelo en rampa. 4: modelo estratificado. Explicación en el texto.

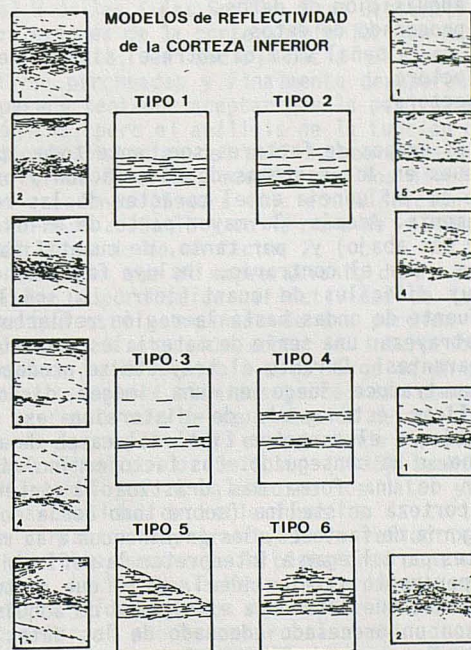


Figura 3. Modelos de reflectividad de la corteza inferior en áreas extensionales recientes y ejemplos de perfiles BIRPS representativos de cada una de las fábricas (según McGearry & Warner, 1987). Explicación en el texto. 1: perfil WINCH. 2: perfil SWAT. 3: perfil SALT. 4: perfil SHET. 5: perfil DRUM.

Por su lado, Wever et al. (1987) analizan los perfiles DEKORP adquiridos sobre los orógenos caledónico y hercínico del centro y NW de Europa y distinguen tres modelos de reflectividad (en áreas con un flujo térmico constante de 60-70 mW/m²):

Tipo 1 o 'acampanado', con una corteza inferior reflectora en toda su extensión.

Tipo 2 o 'multidigitado', con dos o más bandas de reflectividad, separadas por zonas transparentes.

Tipo 3 o 'digitado simple', con una sola banda de reflexiones, normalmente de poco espesor y situada a nivel del Moho de reflexión o a techo de la corteza inferior reflectiva.

Se puede observar su similitud con la clasificación de McGeary (1987), a pesar de que utilizan datos diferentes y de áreas que han sufrido historias geológicas también diferentes.

FACTORES QUE CONDICIONAN EL CARÁCTER DE UNA REFLEXIÓN.

En el aspecto final de las reflexiones en los perfiles de sismica de reflexión profunda influyen una gran cantidad de factores, muchos de los cuales son todavía mal conocidos. A efectos de la exposición, podemos agrupar estos factores en cuatro categorías:

1. Sistema de adquisición de datos.
2. Sistema de procesado de datos.
3. Trayectoria de la señal sísmica entre el sistema de adquisición y la región reflectora.
4. Región reflectora.

Los dos primeros grupos de factores son, ante todo, de índole técnico, y las mejoras continuas en los sistemas de adquisición y procesado de datos permiten suponer que su influencia en el carácter de las reflexiones se irá reduciendo paulatinamente. Además, la mayor parte de ellos son susceptibles de modelización (ver más abajo) y, por tanto, de cuantificación.

El tercer grupo, por el contrario, incluye factores de índole físico, que son, además, muy difíciles de cuantificar. La señal sísmica, en su recorrido desde la fuente de ondas hasta la región reflectora y de regreso a la superficie, debe atravesar una serie de materiales y estructuras que distan mucho de ser 'transparentes'. Durante el trayecto se produce una deformación de la señal que se traduce luego en una imagen distorsionada de los reflectores. Cuantificar este grado de distorsión es fundamental para reconocer su influencia en el carácter final del campo de reflexiones, pero es algo que todavía no se ha conseguido. Los factores incluidos en este grupo son los que limitan de una forma más drástica la interpretación de la reflectividad de la corteza cristalina (sobre todo de la corteza inferior).

La última categoría de factores, los intrínsecos a la región reflectora, son los más importantes para llegar a interpretar la reflectividad en términos geológicos. En un experimento sísmico ideal, en el que se tuviera un control total de las distorsiones que introduce el sistema de adquisición (que luego podrían eliminarse con un procesado adecuado de los datos) y en el que el espacio comprendido entre la superficie del terreno y la región reflectora estuviera ocupado por un material completamente transparente a las ondas sísmicas, el carácter de las reflexiones en el perfil final sería función

únicamente de las características físicas, químicas y geométricas de la región reflectora.

Sistema de adquisición de datos.

El sistema de adquisición de datos (instrumentos y técnicas empleadas) tiene una serie de limitaciones que en su conjunto producen una distorsión de la señal sísmica; esta distorsión afecta sobre todo a la forma (amplitud y frecuencia) de las reflexiones registradas y a la relación señal/ruido.

Podemos agrupar las características del sistema de adquisición que influyen en el carácter de las reflexiones en dos grandes apartados: las que son propias del instrumental utilizado y las que dependen de las condiciones superficiales de registro. Entre las primeras incluimos el rango de frecuencias registrable por los receptores, la máxima frecuencia registrable, el intervalo de muestreo, la longitud de la palabra de máquina del sistema de grabación digital, la sensibilidad de los receptores, la direccionalidad de la fuente y los receptores y el tipo y longitud de la traza utilizada. Entre las segundas cabe citar el nivel de ruido ambiental, el ruido generado por la fuente, el acoplamiento fuente/terreno y receptores/terreno, la cobertura del perfil, la topografía del área, etc.

Dado un instrumental fijo, la única manera de mejorar la calidad de los datos sísmicos adquiridos es modificando las condiciones superficiales de registro, en especial el acoplamiento de la fuente y los receptores.

Peddy & Hobbs (1987) hacen un estudio muy interesante del efecto del nivel de ruido en la visibilidad de los reflectores subhorizontales de la corteza inferior en el perfil WAM de BIRPS realizado en la plataforma y el borde continental al W de las islas Scilly, Gran Bretaña (fig 4a). En la parte E del perfil las reflexiones de la corteza inferior son muy prominentes, pero al avanzar hacia el oeste, es decir, hacia el borde continental, las reflexiones se vuelven parcheadas y finalmente desaparecen (figura 4b). En primera instancia parece tentador aceptar que la pérdida de reflectividad se debe a causas geológicas, pero el análisis de la función de atenuación de la amplitud de dos regiones distintas del perfil, una de la zona con reflexiones y la otra de la que carece de ellas, muestra que los niveles de ruido son muy diferentes en ambas secciones (figura 4c): en la sección con reflexiones el nivel de ruido es de -65dB a 15s TD y en la otra de -50dB a 15s TD, es decir, un aumento del nivel de ruido de un factor 2.

Los autores concluyen que el empeoramiento de la relación señal/ruido puede ser la causa de la pérdida de reflectividad de la corteza inferior, aunque no descartan que esté acompañado de una pérdida real. Hobbs *et al.* (1987) hacen una llamada de atención sobre la interpretación de los datos sísmicos con un nivel de ruido variable a lo largo de un mismo perfil, sobre todo debido a que las secciones sísmicas procesadas comercialmente incluyen entre los pasos de procesado uno de equalización de amplitudes, que tiende a promediar lateralmente las amplitudes, enmascarando por completo las áreas con diferentes relaciones señal/ruido. Smithson (1986) también previene de los posibles efectos del cambio en la calidad de los datos y en las condiciones de registro sobre la cantidad de reflexiones de la corteza inferior. Este autor encuentra un variación muy grande en la reflectividad de la corteza inferior entre los perfiles tomados en la plataforma continental y los tomados en tierra en el norte de los Apalaches: los primeros muestran una corteza inferior muy reflectiva y los segundos una corteza inferior transparente. Sin embargo, como hace notar el propio autor, el cambio de reflectividad puede deberse en parte a la diferente orientación de los perfiles con respecto a la dirección estructural dominante, ya que unos están tomados según esta dirección y los otros perpendicularmente a ella.

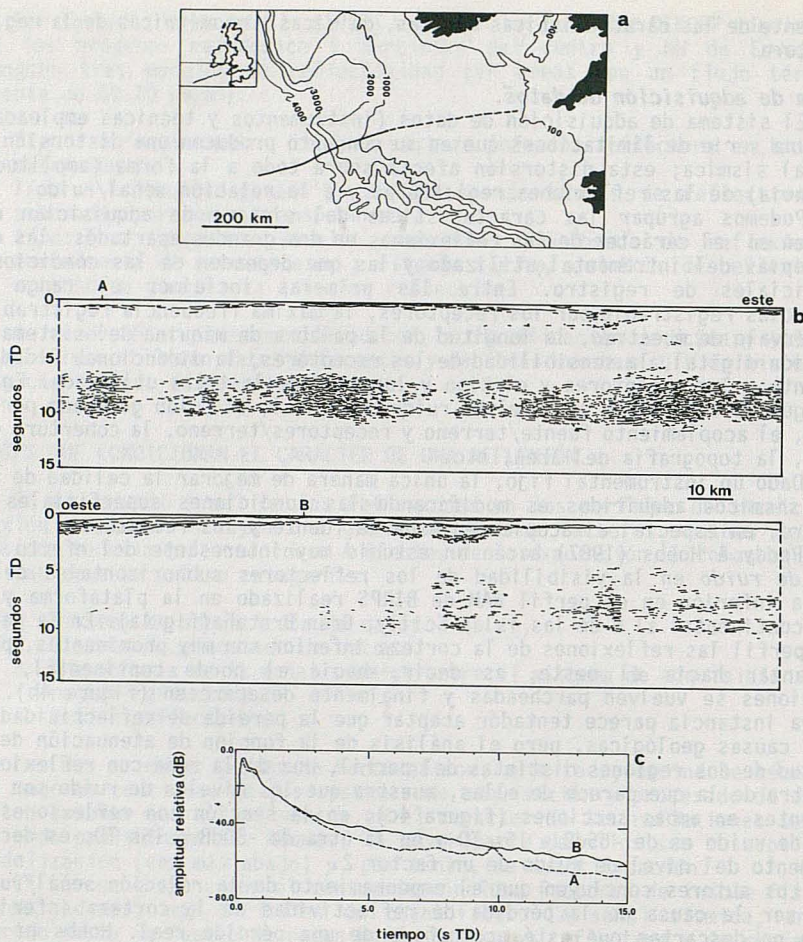


Figura 4. (a) Situación del perfil WAM (Western Approaches Margin) de BIRPS en el Mar Céltico, el talud continental y la Llanura abisal de Porcupino. La parte punteada del perfil corresponde a la sección mostrada en la parte (b). (b) Line drawing de la parte oriental del perfil WAM, desde el punto de inicio en el este hasta el comienzo del talud continental en el oeste. Obsérvese la disminución de la reflectividad de la corteza inferior conforme se aproxima al talud continental. (c) Curvas de amplitud relativa frente a tiempos dobles para dos puntos diferentes del perfil (A y B en la parte (b)). La curva A corresponde a un punto con gran cantidad de reflexiones en la corteza inferior, mientras que la curva B corresponde a un punto con una corteza inferior transparente. ((a), (b), (c), según Peddy & Hobbs, 1987).

Sistema de procesado de datos.

El procesado de un perfil sísmico comprende una serie de pasos que tienden a eliminar de los registros sísmicos todas las reflexiones no primarias y a presentar los reflectores en sus posiciones reales en el subsuelo. Casi nunca se consigue este objetivo en los perfiles convencionales

poco profundos, mucho menos en los de sísmica profunda, y cada uno de los pasos de la secuencia de procesado influye positiva o negativamente en la continuidad, la amplitud y la orientación de las reflexiones (figura 5).



Figura 5. Factores que introducen ruido en los perfiles sísmicos y pasos de la secuencia de procesado que tienden a eliminarlo (según Sheriff, 1977).

Una secuencia básica de tratamiento comprende los pasos de desmultiplexado, etiquetado de trazas, recuperación de amplitudes, eliminación de trazas defectuosas, deconvolución (y mute 1), correcciones estáticas, análisis de velocidades, correcciones dinámicas (y mute 2), ajuste automático de estáticas residuales, stack, filtrado de frecuencias, igualación de amplitudes, correcciones estáticas regionales, aumento de la coherencia espacial y migración.

Claves para el carácter de las reflexiones en el perfil final son las correcciones estáticas, el análisis de velocidades y la migración.

En los últimos años se han desarrollado nuevos algoritmos de procesado, encaminados precisamente a mejorar la fidelidad del perfil sísmico final. Entre las líneas con más futuro podemos citar las técnicas especiales de reducción de ruido, como por ejemplo el filtrado de coherencia espacial, que ha dado lugar a nuevos esquemas para suprimir el ruido no correlacionable en presencia de una señal coherente (Leven & Roy-Chowdhury, 1984, Kong *et al.*, 1985); la deconvolución pre-stack; el filtrado FK; las nuevas técnicas de inversión de velocidad (Zhu & Brown, 1986); la aplicación de 'megaestáticas' (Smithson *et al.*, 1987), que consiste en aplicar correcciones estáticas de gran magnitud con el objetivo de eliminar las diferencias de tiempo entre trayectorias cercanas pero que atraviesan zonas con distinta estructura de velocidad; o técnicas de migración pre-stack de las trazas asociadas con un

mismo punto de tiro (Peshef & Kosloff, 1986, Milkereit, 1987, Louie et al., 1988), que tienen la atractiva característica de permitir visualizar en el perfil final incluso las discontinuidades que forman ángulos muy grandes con la horizontal (Milkereit et al., 1990).

Trayectoria entre el reflector y la superficie.

Si quisiéramos ver un objeto a través de un cristal, la fidelidad con que lo describiésemos dependería en gran medida de la calidad del vidrio interpuesto. Un cristal delgado, homogéneo internamente y bien pulido en superficie nos daría una imagen casi perfecta del objeto. Por el contrario, un cristal grueso, no homogéneo (con variaciones internas de densidad y opacidad) y superficies deficientemente pulidas produciría una imagen distorsionada del objeto que está detrás, más distorsionada cuanto mayores fuesen las imperfecciones. Este es el caso de la sísmica de reflexión profunda: intentamos 'ver' unas estructuras en la corteza a través de un 'cristal' (la parte de la corteza situada por encima) que posee una gran cantidad de imperfecciones y que ni siquiera es 'transparente' en muchos casos. La imagen que obtenemos de los reflectores profundos está distorsionada, tanto que puede llegar a hacerlos irreconocibles.

El conjunto de factores que distorsionan la imagen de un reflector como consecuencia de la variación de las propiedades de los materiales interpuestos entre él y el sistema de adquisición se denominan pérdidas por transmisión y ruido generado por la fuente. Las primeras aparecen por la mera presencia de materiales entre el sistema de adquisición y el reflector (aunque los materiales sean perfectamente homogéneos) y afecta sólo a la forma de las reflexiones registradas: son la atenuación anelástica y la dispersión geométrica. Las segundas tienen que ver directamente con las inhomogeneidades de los materiales interpuestos, en especial la de objetos menores que la primera zona de Fresnel, e incluyen la difracción, los efectos de enfoque y desenfoque debidos a la variación lateral de la velocidad de propagación de las ondas sísmicas (Reston, 1987, Klemperer & BIRPS, 1987, Warner, 1990a), los fenómenos de interferencia entre reflectores y la dispersión elástica en reflectores especulares o en inhomogeneidades aleatorias (Warner, 1990a).

CAUSAS DE LA REFLECTIVIDAD

El papel de la región reflectora.

Para que se produzca una reflexión sísmica debe existir una superficie de separación entre dos medios de diferente *impedancia acústica*. La impedancia acústica (definida como el producto de la velocidad sísmica por la densidad) depende de las propiedades de los materiales, entre ellas (Blundell, 1990): la composición mineralógica, la fábrica cristalina, la textura y la estructura de la roca, la porosidad, el contenido en fluidos y el diaclasado (o cualquier otra dislocación mecánica, como fracturas microcracks, etc.). La amplitud relativa de la reflexión de una discontinuidad recibe el nombre de *coeficiente de reflexión* y depende del cambio fraccional de la impedancia acústica a través de la superficie de separación. Un contraste de impedancia acústica produce una reflexión, pero para que la reflexión sea *coherente* la interfase debe poseer una superficie mínima, que se conoce con el nombre de primera zona de Fresnel. Esta superficie mínima depende de la longitud de onda de la señal sísmica y de la distancia de la interfase a la fuente de las ondas. Las superficies reflectoras con un área inferior a la primera zona de Fresnel, o con geometrías cuya longitud de onda sea menor que la primera zona de

Fresnel, no producen reflexiones coherentes, sino *difracciones*, debido a su capacidad para dispersar la energía incidente.

El contraste de impedancias acústicas es, hablando en términos matemáticos, la condición necesaria para que se produzca una reflexión, pero no es condición suficiente para ello. Existen otras causas que se encargan de modular las reflexiones originadas en un contraste de impedancia acústica (e.g. fenómenos de interferencia, anisotropía, etc.), y que dependen en gran medida de las características de la región reflectora.

Podemos agrupar convenientemente los *factores intrínsecos a la región reflectora* que influyen en el carácter de las reflexiones en tres grandes apartados: (i) contraste de impedancias acústicas, (iii) orientación del reflector, y (iii) geometría del reflector.

El *contraste de impedancias acústicas* es, como acabamos de explicar, el fenómeno físico básico que debe existir para que se produzca cualquier tipo de reflexión sísmica, pero no todo contraste de impedancias acústicas producirá una reflexión en el perfil sísmico final. Existe un límite de resolución acústica por debajo del cual no es posible registrar la reflexión de una discontinuidad. Este límite depende de la profundidad del reflector, del ángulo de incidencia de la señal sísmica y de la sensibilidad del sistema de adquisición. Todo ello hace que un reflector continuo pueda aparecer discontinuo en un perfil sísmico, sobre todo si el propio reflector tiene variaciones laterales en el contraste de impedancias acústicas.

La *orientación* del reflector influye en gran medida en la intensidad de la reflexión final. La técnica de reflexión sísmica habitual de tiro a punto de profundidad común y el proceso de stack posterior tienden a resaltar las reflexiones horizontales y a anular las reflexiones de reflectores muy inclinados, de tal modo que no es posible con una técnica de adquisición y procesado convencional caracterizar en una sección sísmica los reflectores con buzamientos superiores a unos 40 grados (Milkereit *et al*, 1990). Esta puede ser una de las razones de que la corteza superior aparezca casi siempre transparente en los perfiles sísmicos: a pesar de que existen contrastes de impedancia acústica capaces de producir reflexiones, la complejidad estructural que caracteriza a los materiales cristalinos de la corteza superior, con orientaciones muy variables de las interfases que separan rocas con diferentes propiedades acústicas, hace que el resultado final sea el de una región transparente a las ondas sísmicas.

La *geometría* de la región reflectora es el otro factor clave que determina su reflectividad. Podemos distinguir, por un lado, las dimensiones generales del cuerpo reflector (superficie horizontal y espesor vertical) y, por otro, la forma. El primero determina el fenómeno de interferencia, el segundo el de enfoque/desenfoque y uno y otro el de difracción.

No se puede, sin embargo, desligar completamente la región reflectora del resto de factores que influyen en el carácter de las reflexiones, en especial de los que dependen del sistema de adquisición de datos y de la trayectoria de la señal sísmica. Existe un concepto muy importante que acopla precisamente las características de la región reflectora (más concretamente sus dimensiones) y las características del sistema de adquisición (longitud de onda de la señal sísmica) y es el concepto de *resolución*. La resolución espacial es una medida del tamaño mínimo de los objetos que una señal sísmica es capaz de detectar en forma de reflexiones coherentes. Este tamaño mínimo no es igual en la vertical que en la horizontal y por ello se distingue entre resolución horizontal y resolución vertical.

La *resolución horizontal* es la superficie mínima que debe tener un objeto para que de lugar a una reflexión coherente y se conoce con el nombre de primera zona de Fresnel (Sheriff, 1977); depende de la longitud de onda de

la señal sísmica y de la distancia de la interfase a la superficie. Para un reflector situado en la corteza inferior y una frecuencia de 20 Hz, la primera zona de Fresnel es del orden de 12 km² (equivalente a una longitud de 4 km para el caso bidimensional). Las estructuras del subsuelo con áreas menores que ésta o con geometrías cuya longitud de onda sea menor que ésta no producen reflexiones coherentes, sino difracciones.

La *resolución vertical* mide la distancia mínima entre dos interfaces para que aparezcan como dos reflexiones distintas; depende de la longitud de onda de la señal sísmica. Debido al proceso de atenuación anelástica, que como hemos comentado más arriba afecta en mayor medida a las frecuencias altas (longitudes de onda pequeñas), conforme una señal sísmica penetra en el subsuelo va cambiando su contenido frecuencial por pérdida de las frecuencias altas, lo que se traduce en una disminución de la resolución vertical con la profundidad (a éllo contribuye también el aumento de la velocidad sísmica con la profundidad). A la profundidad de la corteza inferior, la frecuencia más alta que permanece en la señal es del orden de 20 Hz (Sheriff, 1977), que para una velocidad de propagación de unos 5 km/s implica una longitud de onda de unos 250 m (Lavergne, 1989, p. 62). La experiencia en formaciones sedimentarias muestra que el espesor mínimo de una capa en la que el techo y el muro se ven como dos reflexiones distinguibles es del orden de 1/4 de la longitud de onda dominante si la relación señal/ruido es alta y de 1/2 si es baja. Esto significa que en la corteza inferior no es posible distinguir como reflexiones individualizadas discontinuidades que estén a menos de 100 m de distancia vertical. Por debajo de este umbral empiezan a jugar un papel predominante los fenómenos de interferencia constructiva y destructiva.

La *detección* es otro concepto relacionado y se define como la posibilidad de obtener reflexiones de capas delgadas (i.e. de discontinuidades muy próximas) y el límite de detección es el espesor mínimo de una capa visible en sísmica de reflexión. Depende de la longitud de onda, la relación señal/ruido y el contraste de velocidades (Lavergne, 1989). Teóricamente, con relaciones señal/ruido altas y contrastes de velocidad fuertes es posible detectar capas con un espesor mayor o igual a 1/30 de la longitud de onda dominante (Sheriff, 1980). En la práctica, sin embargo, este límite es de 1/4 de la longitud de onda. A profundidades del orden de 20 km, donde predominan longitudes de onda de unos 250 m, las interfaces separadas menos de 60 m interfieren destructivamente y no se detectan, mientras que si están separadas por más de 60 m interfieren constructivamente y es posible que se traduzcan en reflexiones en el perfil sísmico final.

El origen de cualquier reflexión es un contraste de impedancia acústica entre dos medios con distintas propiedades. Esto es cierto tanto para las rocas sedimentarias como para las rocas cristalinas. Para las primeras conocemos además su disposición estructural gracias a la presencia de *planos de estratificación*, que sirven a la vez de marcadores estructurales y de discontinuidades entre materiales con diferentes propiedades acústicas. La separación entre planos de estratificación, junto con la longitud de onda de la señal sísmica, son las responsables de los fenómenos de interferencia constructiva y destructiva que modulan el carácter de las reflexiones de las rocas sedimentarias. Son los propios fenómenos de interferencia los que hacen que no exista una correspondencia unívoca entre las reflexiones de un perfil y los reflectores del subsuelo. Es decir, las reflexiones de las capas sedimentarias son en realidad *reflexiones compuestas*, consecuencia de los fenómenos de interferencia en sucesiones de capas delgadas (Sheriff, 1977). Los sondeos mecánicos, las diagráfias y los sondeos sísmicos verticales (SSV) han permitido precisar este cuadro de reflectividad en las rocas sedimentarias,

por correlación de la información proporcionada por estas técnicas con la información sísmica de superficie.

Vemos pues que ha sido el conocimiento conjunto de los *coeficientes de reflexión* y de las *características geométricas de la región reflectora* las que han permitido profundizar en las causas de la reflectividad de las rocas sedimentarias. Sin embargo, para las rocas cristalinas de la corteza, y en especial para las de la corteza inferior, carecemos de cualquier información acerca de las características de la región reflectora. Esto plantea una serie de interrogantes acerca de las reflexiones que observamos en los perfiles: ¿son estas reflexiones primarias, debidas únicamente a contrastes en la impedancia acústica?, ¿son importantes, y en qué medida, los efectos de interferencia en el comportamiento reflector de la corteza cristalina?, ¿influye decisivamente la geometría de la región reflectora en el carácter de la reflectividad?, ¿qué escalas de longitud vertical y horizontal caracterizan a los materiales de la corteza inferior?, ¿existen orientaciones preferentes que aumentan la reflectividad (anisotropía)?, ¿cuál es la causa de que la mayor parte de las reflexiones de la corteza inferior tengan longitudes menores que la primera zona de Fresnel?, ¿se debe la reflectividad de la corteza inferior a causas físicas situables en la propia corteza inferior, o son consecuencia de otros factores ajenos a ella?, ¿pueden producirse reflexiones en ausencia de contrastes importantes en la impedancia acústica de los materiales?, etc.

Tenemos ante nosotros una tarea doble, o mejor, una tarea en dos niveles. En primer lugar, hay que identificar y aislar las características de la región reflectora que son importantes a la hora de influir en su reflectividad; en segundo lugar, hay que relacionar dichas características con estructuras geológicas determinadas. Para conseguir el primer objetivo el arma más eficaz es la *modelización sísmica*; para lograr el segundo no hay otra opción que extrapolar hasta la superficie (o hasta algún sondeo) las reflexiones que queremos caracterizar. Aquí sólo vamos a desarrollar el primer aspecto del problema, es decir, identificar las causas (físicas) de la reflectividad. La segunda parte del problema, el descubrir las estructuras geológicas que actúan de reflectores en la corteza inferior, el porqué de su reflectividad en cada caso, y los procesos (geoquímicos, petrológicos, tectónicos y geofísicos) que han originado dichas estructuras, se reserva para un próximo artículo.

Modelización sísmica.

En sísmica de reflexión, modelizar consiste en construir sismogramas sintéticos a partir de un modelo geológico hipotético y compararlo luego con los sismogramas reales obtenidos sobre el terreno.

La construcción de un sismograma sintético requiere de una serie temporal de reflectividades que se convolucionan con una onda en fase cero o fase mínima de la forma apropiada (Rickert, Butterwarth etc., o modelizada a su vez) y con una banda de frecuencias similar a la de la sección sísmica. La serie temporal de reflectividades se puede obtener a partir de las diagrfías sónicas en un sondeo mecánico, por medio de un sondeo sísmica vertical, o puede calcularse a partir de un modelo geológico hipotético. Este último caso es el más común en la modelización sísmica de la corteza inferior y la construcción del modelo geológico se puede considerar el primer paso en la modelización. El nivel de detalle del modelo es función de los rasgos que se quieran modelizar y de la información geológica disponible. Puede variar desde una subdivisión del espacio del modelo en regiones simples de velocidad y densidad constantes, separadas por discontinuidades planas (e.g. Peddy, 1984, fig.3), hasta reproducciones más o menos fieles de la geología de un área

(e.g. Hurich & Smithson, 1987, fig.11), con límites multicapa de geometrías complicadas separando regiones con variaciones laterales de la velocidad y/o la densidad.

Una vez construido el modelo geológico, éste se resuelve utilizando para ello un método de cálculo numérico basado bien en la teoría de rayos sísmicos o en la resolución completa de la ecuación de onda. El método de los rayos consiste en el trazado de rayos sísmicos (mediante una técnica de ordenador apropiada) a través del modelo, que se reflejan en los contactos reflectores según una incidencia normal (Zawislak & Smithson, 1981). Los problemas resueltos de este modo parten necesariamente de un modelo subdividido por límites continuos en regiones cerradas de impedancia acústica constante (Blundell *et al.*, 1985). Los modelos basados en la resolución de la ecuación de onda no tienen esta limitación. También se puede optar por una solución intermedia: definir primero, mediante el trazado de rayos, un número elevado de puntos de control en las superficies reflectoras y aplicar luego a estos puntos una simulación de la teoría de ondas (trazado de los rayos de difracción que parten de los puntos de control) (Reston, 1987). La resolución de la ecuación de onda puede abordarse mediante diferencias finitas (e.g. Gibbson & Lavender, 1987), elementos finitos o técnicas espectrales (e.g. Juhlin, 1990, Sandmeier *et al.*, 1987), e incluir los efectos de la anelasticidad, heterogeneidad y anisotropía en una, dos o tres dimensiones.

La forma de construir los sismogramas sintéticos dependerá en cada caso de las características que se deseen reproducir: puede optarse por reproducir fielmente la forma completa de una sola reflexión (e.g. Warner, 1990a, Hobbs, 1990), la amplitud de una o varias reflexiones (e.g. Hurich & Smithson, 1987), la respuesta frecuencial de determinadas reflexiones (Juhlin, 1990), las relaciones geométricas y de corte entre varios reflectores (e.g. McGeary, 1989, Peddy, 1984) o la 'apariencia' general de toda una sección sísmica (e.g. Blundell *et al.*, 1985, Sandmeier *et al.*, 1987, Fountain *et al.*, 1987, Reston, 1987, Hurich & Smithson, 1987).

Causas de la reflectividad.

De los ejemplos descritos y de muchos otros que pueden encontrarse en la literatura, se puede extraer la conclusión de que no existe un acuerdo sobre las causas que provocan la reflectividad de la corteza inferior. A pesar de que la modelización sísmica es el arma más potente para restringir las posibles causas, los estudios realizados hasta la fecha no han hecho más que exponer una lista de posibles causas, sin indicar la importancia relativa de cada una de ellas en cada caso particular.

Podemos resumir de la siguiente forma las causas propuestas para la reflectividad de la corteza inferior (figura 6):

1. Efectos geométricos (en ausencia de contrastes importantes de impedancia acústica). En especial:
 - 1.1. Fenómenos de interferencia. Se han descrito tres modalidades principales:
 - 1.1.1. en una capa delgada limitada superior e inferiormente por materiales distintos (Blundell *et al.*, 1985) (figura 6a).
 - 1.1.2. en una serie de capas delgadas (Hurich & Smithson, 1987) (figura 6b).
 - 1.1.3. según modelos más complejos (Sandmeier *et al.*, 1987, Reston, 1987) (figura 6c,d).
 - 1.2. Fenómenos de difracción (Reston, 1987).
 - 1.3. Continuidad lateral de los reflectores (Juhlin, 1990).

2. Contrastes de impedancia acústica. Este contraste puede tener dos orígenes:

2.1. Contrastes composicionales (Juhlin, 1990, Warner, 1990b, Potter et al., 1987) (figura 6e).

2.2. Contrastes físicos. Se han expuesto principalmente dos:

2.2.1. debidos a la anisotropía generada por la presencia de fábricas deformacionales (Green et al., 1990) (figura 6f).

2.2.2. debidos a la presencia de fluidos (figura 6g), ya sean acuosos (Hall, 1986, Matthews, 1986, Gough, 1986, Hyndman, 1988) o magmas (Mayer & Brown, 1986, Potter et al., 1987, deVoogd et al., 1986).

2.2.3. debidos a otras causas, como por ejemplo, transformaciones polimórficas (Bois et al., 1987).

CAUSAS DE LA REFLECTIVIDAD DE LA CORTEZA INFERIOR

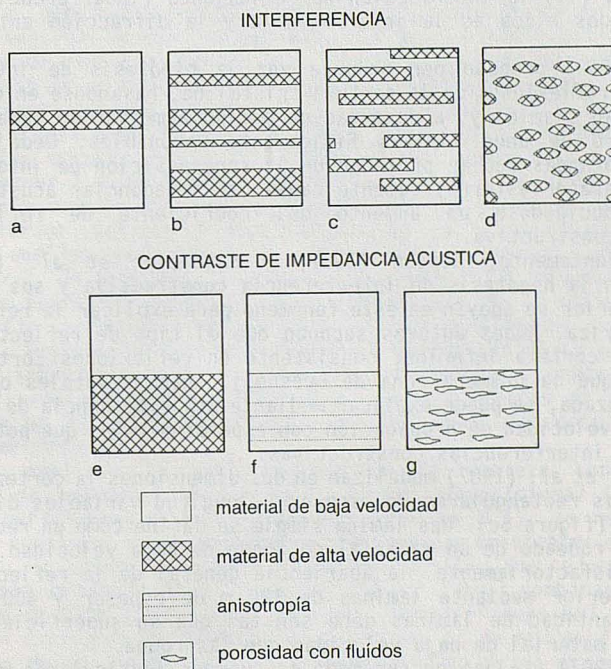


Figura 6. Causas de la reflectividad de la corteza inferior propuestas en la literatura. a-d: reflectividad debida a fenómenos de interferencia en ausencia de contrastes de impedancia acústica importantes. e-g: reflectividad debida sólo a contrastes de impedancia acústica. Explicación en el texto.

En un epígrafe anterior ya se ha hablado ampliamente de la posibilidad de que los *efectos geométricos*, en ausencia de contrastes acústicos importantes, puedan ser la causa principal de la reflectividad de la corteza

inferior. Es una hipótesis atractiva, ya que los efectos geométricos son capaces de amplificar la intensidad de una reflexión por encima del valor correspondiente al contraste de impedancias acústicas entre los materiales que forman el reflector. Debido a que los coeficientes de reflexión de la corteza inferior son muy altos, del orden de 0.1 (superiores a la mayor parte de los coeficientes de reflexión de las rocas sedimentarias) y son difíciles de explicar por simples contrastes litológicos, se ha acudido a esta hipótesis para no necesitar de contrastes tan bruscos para poder explicar la reflectividad. De este modo, con contrastes de impedancia acústica más normales (como los existentes, por ejemplo, entre un gneiss ácido y otro básico), pero amplificadas por interferencia constructiva, se pueden explicar las reflexiones de la corteza inferior.

Otro de los motivos por los que la hipótesis de la interferencia es atractiva, es la observación sistemática de que la mayor parte de las reflexiones de la corteza inferior tienen una longitud inferior a la primera zona de Fresnel. Reston (1987) cita cuatro factores que pueden acortar la longitud de las reflexiones en un perfil no migrado: (i) la presencia de ruido, (ii) el efecto de enfoque, (iii) la reducción de la coherencia de una reflexión por variaciones laterales de la velocidad en la corteza suprayacente, y (iv) la intersección de reflexiones (interferencia). El más factible de todos ellos es la interferencia y la difracción entre cuerpos complejos.

Fuchs (1969) propuso por primera vez la hipótesis de interferencia aplicada a las reflexiones de la corteza cristalina, basándose en el hecho de que la velocidad sísmica y la densidad no podían aumentar en cada reflector y seguir dentro de unos límites físicamente razonables. Dedujo que las reflexiones profundas debían proceder de la superposición de interfases muy juntas que separasen alternativamente capas con impedancias acústicas altas y bajas, produciéndose un aumento del coeficiente de reflexión por interferencia constructiva.

Más recientemente, Reston (1987) y Sandmeier *et al.* (1987) han profundizado en la hipótesis de interferencia constructiva y sus modelos de la corteza inferior se apoyan en este fenómeno para explicar la reflectividad que la caracteriza. Ambos autores suponen que el tipo de reflectividad más habitual de la corteza inferior, consistente en reflexiones cortas (muchas veces menores que la primera zona de Fresnel), subhorizontales o según una disposición cruzada, se puede explicar mediante una alternancia de materiales de alta y baja velocidad de propagación con espesores tales que potencien los efectos de las interferencias constructivas.

Sandmeier *et al.* (1987) modelizan en dos dimensiones la corteza inferior mediante láminas rectangulares de espesor y longitud variables distribuidas aleatoriamente (figura 6c). Una lámina simple se define como un rectángulo de alta velocidad rodeado de un material de fondo de baja velocidad. Consiguen reproducir satisfactoriamente la apariencia general de la reflectividad de la corteza inferior mediante láminas de 120 m de espesor y 400-1200 m de longitud. La cantidad de láminas debe ser tal que su superficie total sea igual a la del material de baja velocidad que las rodea.

Reston (1987) utiliza en sus modelos cuerpos lenticulares en lugar de láminas rectangulares (figura 6d). Variando el empaquetamiento (dimensiones y separación) y la forma (relación longitud/espesor) de los cuerpos lenticulares, consigue reproducir la apariencia de los perfiles de reflexión de la corteza inferior.

Hurich & Smithson (1987), en un estudio de los coeficientes absolutos de reflexión, ponen de manifiesto, mediante sismogramas sintéticos, que una serie de capas de espesor adecuado en la que alternen dos materiales de

diferentes propiedades acústicas, puede amplificar hasta en un factor 3 su coeficiente de reflexión con respecto al de una discontinuidad simple entre esos mismos materiales (figura 6b). Conceden mucha importancia a los efectos de interferencia constructiva a la hora de hacer visibles en los perfiles sísmicos reflexiones que de otra forma estarían por debajo del umbral acústico de detección.

Sin embargo, existe un número creciente de autores que rechazan la hipótesis de que los efectos geométricos (enfoque/desenfoque, difracción e interferencia) sean capaces de explicar la reflectividad de la corteza inferior. En este sentido, Warner (1990a) critica los resultados de Hurich & Smithson (1987) por no tener en cuenta las características de fase de la onda incidente a la hora de calcular el coeficiente de amplificación. Normalmente las modelizaciones se llevan a cabo con una onda en fase cero (e.g. Hurich & Smithson, op.cit.) y sin embargo, el tipo de señal usado en sísmica de reflexión se aproxima a una onda en fase mínima. Utilizando este tipo de fase, la máxima amplificación que se consigue del coeficiente de reflexión cuando la interferencia constructiva es máxima es de 1.5X. Raynaud (1986) ha mostrado que los efectos de interferencia constructiva en una secuencia estratificada de capas puede producir, como máximo, un aumento en un factor 2 en la amplitud de reflexión, pero que los modelos más realistas sólo producen amplificaciones del 50 % o menores. Warner (1990a) concluye que la interferencia constructiva no es capaz de explicar por sí sola la intensidad de las reflexiones de la corteza inferior y que cualquier modelo geológico factible de la corteza inferior tiene que ser capaz de explicar la existencia de contrastes de impedancia acústica del 20 %, que producen coeficientes de reflexión del orden de 0.1, mucho mayores que en las secuencias sedimentarias.

Klemperer & BIRPS (1987) también se oponen a la conclusión de que la interferencia constructiva es la responsable del carácter reflectivo de la corteza inferior. Dan por válidos factores de amplificación de hasta 2X, pero encuentran que los coeficientes de reflexión de la corteza inferior varían en un factor $\gg 2$, de modo que la presencia o ausencia de interferencia en capas delgadas no puede ser la única causa responsable de la reflectividad y que deben existir contrastes sustanciales en las propiedades acústicas de los materiales. Según estos autores, cualquier modelo geológico satisfactorio de la corteza inferior debe incorporar capas de decenas a centenares de metros de espesor, con extensiones laterales a menudo mayores de 10 km y que se diferencien en las propiedades acústicas de capas adyacentes en unas pocas unidades porcentuales.

Para explicar la reflectividad de la corteza inferior mediante *contrastos de impedancia acústica de origen composicional* (es decir, para explicar coeficientes de reflexión del orden de 0.1), se requiere que los contrastes acústicos sean aproximadamente como los existentes entre un granito y un gabro, o entre un gabro y una peridotita. Por ejemplo, se pueden conseguir con un cambio de velocidad de 6.3 a 7.1 km/s combinado con un cambio de densidad de 2700 a 3000 kg/m³ (Warner, 1990a) (figura 6e).

Si el contraste de impedancias acústicas es de origen físico, debido a los efectos de la *anisotropía*, por superposición de un mismo material pero isótropo a un lado de la discontinuidad y anisótropo al otro (figura 6f), entonces los coeficientes de reflexión medidos requieren un 40% de anisotropía en el material inferior. Así, sería necesaria una roca isótropa con una velocidad de 6.6 km/s contra una roca anisótropa de la misma composición pero con velocidades de 5.4 km/s en la dirección lenta y de 8.0 km/s en la dirección rápida (Warner, 1990a). Esta es una anisotropía poco realista para una roca deformada en condiciones naturales. En este sentido, Hall (1986) cita los resultados del análisis de velocidad sísmica según direcciones ortogonales

en un espécimen de afibolita del Lewisian escocés y encuentra un valor de 6.64 km/s para la velocidad más baja y uno de 7.46 km/s para la más alta, lo que equivale a una anisotropía algo superior al 10%. Christensen (1989) recoge muchos valores de anisotropía sísmica en rocas ígneas y metamórficas y, salvo una excepción (una labradorita con un 43% de anisotropía), todas las rocas listadas tienen anisotropías muy inferiores al 40%. Así por ejemplo, los valores para esquistos y anfibolitas (que en promedio son los más elevados), quedan siempre por debajo del 20% y para el resto de las rocas ígneas y metamórficas la anisotropía es todavía mucho menor.

Por último, si el contraste de impedancia acústica es debido a la presencia de fluidos acuosos en rocas que han alcanzado el equilibrio textural (algo muy factible a las presiones y temperaturas de la corteza inferior), sería necesario tener en contacto una roca sin porosidad con otra con un 7% de porosidad (suponiendo una composición global igual para ambas rocas, figura 6g) (Warner, 1990a). Este valor de porosidad no es muy probable a las profundidades de la corteza inferior.

CONSIDERACIONES FINALES

A modo de conclusión, en este apartado se pretenden subrayar una serie de puntos discutidos ya a través del artículo, pero que por tratarse de conceptos emparentados pero distintos es importante matizar sus diferencias.

1. Un perfil sísmico no es una representación exacta de la geología profunda de un área. Es muy importante apreciar la diferencia entre el carácter de las reflexiones y las características de los reflectores. Entre ambos conceptos no existe una relación del tipo $CR = f(RR)$, es decir, el carácter de las reflexiones, CR, no es sólo función de las características de la región reflectora, RR, sino que entran en juego otras variables, de modo que podemos expresar dicha relación en la forma:

$$CR = f(\alpha SA, BSP, \gamma T, \delta RR)$$

donde SA son los factores que dependen del sistema de adquisición de datos, SP los que dependen del sistemas de procesado de datos, T los que dependen de la trayectoria seguida por la señal sísmica y RR los factores que dependen de las características de la región reflectora. α , B, γ y δ son pesos que pueden tener cualquier valor entre 0 y 1 sujetos a la restricción $\alpha + \beta + \gamma + \delta = 1$. En cada perfil (e incluso en cada porción de un mismo perfil) los pesos α , B, γ y δ son diferentes y, por tanto, también es diferente la importancia de cada una de las variables en el resultado final.

Esta es la razón por la que la prudencia ha sido y es la norma general a la hora de interpretar las reflexiones de la corteza cristalina en los perfiles de sísmica de reflexión profunda. Y ésta es también la razón de la intensa labor que se ha llevado a cabo para reducir al mínimo la importancia de los factores extrínsecos a la región reflectora (SA, SP y T en la ec.5) en el carácter final de las reflexiones. En la terminología de la ec. 5, esto se traduce en hacer que α , B y γ tiendan a cero, mientras que δ tiende a 1.

2. También se ha querido resaltar a lo largo de este artículo la diferencia que existe entre la interpretación geométrico-descriptiva y la interpretación físico-geológica de los perfiles. La primera se lleva a cabo en tres niveles: elementos de las reflexiones aisladas, fábricas sísmicas y modelos de reflectividad, que conducen a una clasificación de la corteza en función de su comportamiento sísmico. Para pasar de la descripción geométrica a la

interpretación físico-geológica es necesario un profundo conocimiento de todos los factores que afectan al carácter del campo de reflexiones, así como de otra serie de parámetros geológicos, geoquímicos y geofísicos que entran en la interpretación a modo de restricciones adicionales a los posibles modelos factibles.

3. El otro par de conceptos relacionados del que se han querido resaltar sus diferencias es el de *causas físicas* de la reflectividad y *estructuras geológicas* responsables de dicha reflectividad. Se ha querido evitar en todo momento una secuencia del tipo

caracter de las reflexiones → *estructura geológica responsable*

y en cambio se ha utilizado otra del tipo

caracter de las reflexiones →^(a) *causas de la reflectividad* →^(b) *características de la región reflectora* →^(c) *estructura geológica responsable*.

El paso (a), comentado ya en el punto 1 de este resumen final, presupone que la relación funcional expresada por la ec. (5), cumple la restricción adicional $\delta \gg \alpha + \beta + \gamma$, es decir, que, en primera aproximación, el carácter de las reflexiones se puede correlacionar directamente con las características de la región reflectora. El paso (b) implica reconocer, a partir de las causas de la reflectividad, las dimensiones, forma, orientación y contraste de impedancias acústicas de la región reflectora responsables del carácter de las reflexiones. El paso (c) es el más complicado, ya que requiere del conocimiento de la historia geológica del área, deducida, en su mayor parte, de los estudios geológicos de superficie (geoquímicos, petrológicos, tectónicos, geodinámicos, etc) y los estudios geofísicos complementarios (gravimétricos, geomagnéticos, geotérmicos, reológicos, etc.). Los datos aportados por todas estas ramas de las Ciencias de la Tierra, en conjunción con los aportados por la sísmica de reflexión profunda, permiten restringir los posibles modelos geológicos compatibles con la geometría, la orientación y los contrastes de impedancia acústica de la región reflectora.

Se puede extraer una conclusión inmediata de este razonamiento: identificar las causas de la reflectividad en un caso particular no significa reconocer la estructura geológica responsable de dicho campo de reflexiones. Una misma estructura geológica puede deber su reflectividad a causas físicas distintas en cada caso; y, al contrario, dos estructuras geológicas distintas pueden ser reflectoras por la misma causa física. Por ejemplo, una zona de falla puede ser reflectiva por una o varias de las siguientes causas: (i) el contraste de impedancias acústicas que supone poner en contacto dos materiales distintos como consecuencia del desplazamiento de la falla; (ii) la formación en el plano de falla de una banda estrecha de materiales con propiedades acústicas diferentes como consecuencia de los procesos de deformación cataclástica que tienen lugar en las inmediaciones del plano de falla; (iii) la presencia de una película de fluidos que se desplazan por el plano de falla y la banda deformada adyacente aprovechando sus características de zona de debilidad. La estructura geológica es en todos los casos la misma (una zona de falla), pero la causa física que la hace reflectiva varía en cada caso. También podemos poner un ejemplo del caso contrario: un contraste de impedancias acústicas de origen composicional puede encontrarse en una zona de falla, en el contacto entre un cuerpo intrusivo y la roca encajante, entre dos formaciones sedimentarias, etc. En todos los casos la causa de la reflectividad es la misma, pero las estructuras geológicas responsables de dicha reflectividad son muy distintas. Hurich & Smithson (1987) están de acuerdo en que identificar las causas de las reflexiones corticales es

esencial para obtener más información de carácter geológico de los perfiles de sísmica de reflexión profunda. La interpretación de estos perfiles es difícil precisamente porque se carece del calibrado geológico que se tiene en las rocas sedimentarias. El siguiente paso que debe dar la sísmica de reflexión profunda es el que lleve a la identificación, no sólo de las estructuras geológicas que caracterizan a la corteza inferior, sino de los procesos responsables de la génesis de dichas estructuras.

REFERENCIAS

- ALLMENDINGER, R., HAUGE, T., HAUSER, E., POTTER, C., KLEMPERER, S., NELSON, K., KNUEPFER, P. & OLIVIER, J. (1987a). *Geol. Soc. Am. Bull.*, **98**, 308-319.
- ALLMENDINGER, R., NELSON, K., POTTER, C., BROWN, L., BARAZANGI, M. & OLIVIER, J. (1987b). *Geology*, **15**, 304-310.
- BARAZANGI, M. & BROWN, L. (1986a). *Am. Geophys. Union Geodyn. Ser.*, **13**, 1-311.
- BARAZANGI, M. & BROWN, L. (1986b). *Am. Geophys. Union Geodyn. Ser.*, **14**, 1-339.
- BEHRENDT, J., HUTCHINSON, D., LEE, M., THORNER, C., TREHU, A., CANNON, W. & GREEN, A. (1990). *Tectonophysics*, **173**: 595-616.
- BLUNDELL, D. (1990). *J. Geol. Soc. London*, **147**, 895-913.
- BLUNDELL, D., HURICH, C. & SMITHSON, S. (1985). *J. Geol. Soc. London*, **142**, 245-258.
- BOIS, C., DAMOTTE, B., MASCLE, A., CAZES, M., HIRN, A. & BIJU-DUVAL, B. (1987). *Geophys. J. R. Astr. Soc.*, **89**, 279-286.
- BOIS, C. & E.C.O.R.S. SCIENTIFIC PARTY (1990). *Tectonophysics*, **173**, 397-410.
- CHRISTENSEN, N. (1989). In: Carmichael (ed). *Practical handbook of physical properties of rocks and minerals*. Florida, CRC Press, 429-546.
- D.E.K.O.R.P. RESEARCH GROUP (1990). *Tectonophysics*, **173**, 199-206.
- DeVOOGD, B., SERPA, L., BROWN, L., HAUSER, E., KAUFMAN, S., OLIVIER, J., TROXEL, B., WILLEMIN, J. & WRIGHT, L. (1986). *Geology*, **14**, 64-67.
- DOHR, G. (1957). *Geol. Rund.*, **46**, 17-26.
- FOUNTAIN, D., McDONAGHT, D. & GORHAM, J. (1987). *Geophys. J. R. Astr. Soc.*, **89**, 61-66.
- FUCHS, K. (1969). *Zeits. Geophys.*, **35**, 133-149.
- GIBSON, B. & LEVANDER, A. (1987). *Geophys. J. R. Astr. Soc.*, **89**, 19-20.
- GOUGH, D. (1986). *Nature*, **323**, 143-144.
- GREEN, A., MILKEREIT, B., PERCIVAL, J., DAVIDSON, A., PARRISH, R., COOK, F., GEIS, W., CANNON, W., HUTCHINSON, D., WEST, G. & CLOWES, R. (1990). *Tectonophysics*, **173**, 627-638.
- HALL, J. (1986). *Geol. Soc. Spec. Publ.*, **24**, 51-62.
- HOBBS, R. (1990). *Tectonophysics*, **173**, 25-30.
- HOBBS, R., PEDDY, C. & B.I.R.P.S. GROUP (1987). *Geophys. J. R. Astr. Soc.*, **89**, 239-242.
- HURICH, C. & SMITHSON, S. (1987). *Earth Planet. Sci. Lett.*, **85**, 416-426.
- HYNDMAN, R. (1988). *J. Geophys. Res.*, **93**, 3391-3405.
- JUHLIN, C. (1990). *Tectonophysics*, **173**, 345-360.
- JUNGER, A. (1951). *Geophysics*, **16**, 499-505.
- KLEMPERER, S. & B.I.R.P.S. GROUP (1987). *Geophys. J. R. Astr. Soc.*, **89**, 217-222.
- KONG, S., PHINNEY, R. & ROY-CHOWDHURY, K. (1985). *Geophysics*, **50**, 539-550.
- LAVERGNE, M. (1989). *Seismic methods*. Graham & Trotman-Technip.
- LEVEN, J., FINLAYSON, D., WRIGHT, C., DOOLEY, J. & KENNETT, B. (eds) (1990). *Tectonophysics*, **173**, 1-641.
- LEVEN, J. & ROY-CHOWDHURY, K. (1984). *54th. Annu. Int. Meet., Soc. Explr. Geophys.*, 432-436.
- LOUIE, J., CLAYTON, R. & LeBRAS, R. (1988). *Geophysics*, **53**, 176-185.
- MATTHEWS, D. (1986). *Geol. Soc. Spec. Publ.*, **24**, 11-22.

- MATTHEWS, D. (1987). *Ann. Geophys.*, 5B: 353-356.
- MATTHEWS, D. & SMITH, C. (eds) (1987). *Geophys. J. R. Astr. Soc.*, **89**, 1-447.
- MAYER, J. & BROWN, L. (1986). *Geophysics*, **51**, 1050-1054.
- McGEARY, S. (1987). *Geophys. J. R. Astr. Soc.*, **89**, 231-238.
- McGEARY, S. (1989). *J. Geol. Soc. London*, **146**, 261-269.
- McGEARY, S. & WARNWER, M. (1987). *Nature*, **317**: 795-797.
- MEISSNER, R., WEVER, Th. & FLUEH, E. (1987). *Ann. Geophys.*, 5B: 357-364.
- MILKEREIT, B. (1987). *J. Geophys. Res.*, **92**, 7916-7930.
- MILKEREIT, B., GREEN, A., LEE, M., AGENA, W. & SPENCER, C. (1990). *Tectonophysics*, **173**, 1-14.
- O'DOHERTY, R. & ANSTEY, N. (1971). *Geophys. Prosp.*, **19**, 430-458.
- PEDDY, C. (1984). *Nature*, **312**, 628-630.
- PEDDY, C. & HOBBS, R. (1987). *Ann. Geophys.*, 5B, 331-338.
- PESHEF, M. & KOSLOFF, D. (1986). *Geophysics*, **51**, 324-331.
- POTTER, C., LIU, C., HUANG, J., ZHENG, L., HAUGE, T., HAUSER, E., ALLMENDINGER, R., OLIVIER, J., KAUFMAN, S. & BROWN, L. (1987). *Geol. Soc. Am. Bull.*, **98**, 330-337.
- RAYNAUD, B. (1986). *Seismic studies of the lower crust*. Ph.D. Thesis, Univ. Cambridge, 128pp.
- RESTON, T. (1987). *Ann. Geophys.*, 5B, 339-348.
- SANDMEIER, K., WAELDE, W. & WENZEL, F. (1987). *Geophys. J. R. Astr. Soc.*, **89**, 339-344.
- SHERIFF, R. (1977). *AAPG Mem.*, **26**, 3-14.
- SHERIFF, R. (1980). *Seismic Stratigraphy*. IHRDC, Boston (USA).
- SHOR, G. (1955). *Trans. Am. Geophys. Union*, **36**, 133-138.
- SMITHSON, S. (1986). *Geol. Soc. Spec. Publ.*, **24**, 23-34.
- SMITHSON, S., JOHNSON, R., HURICH, C., VALASEK, P. & BRANCH, C. (1987). *Geophys. J. R. Astr. Soc.*, **89**, 67-72.
- TOMEK, C., DVORAKOVA, L., IBRMAJER, I., YIRICEK, R. & KORAB, T. (1987). *Geophys. J. R. Astr. Soc.*, **89**: 383-388.
- WARNER, M. (1990a). *Tectonophysics*, **173**, 15-23.
- WARNER, M. (1990b). *Tectonophysics*, **173**, 163-174.
- WEVER, Th., TRAPPE, H. & MEISSNER, R. (1987). *Ann. Geophys.*, 5B, 255-266.
- ZAWISLAK, R. & SMITHSON, S. (1981). *Geophysics*, **46**, 1684-1701.
- ZHU, T. & BROWN, L. (1986). *J. Geophys. Res.*, **91**, 11477-11495.

LAS SURGENCIAS TERMALES DE LOS BAÑOS (TERUEL). II. CALCULOS GEOTERMOMETRICOS.

Tena, J.; Auqué, L.F.; Mandado, J. y Gimeno, M.J.

Departamento de Geología. Facultad de Ciencias, Universidad de Zaragoza. 50009 ZARAGOZA.

ABSTRACT

Springs of *Los Baños* (Teruel) are low enthalpy manifestations associated with a carbonatic-evaporitic reservoir. Calculations of deep temperature and solution conditions by means of empirical geothermometers are very difficult because of their equilibrium basic assumptions, or their calibration conditions, are not fulfilled in these geothermal system types.

Geochemical modeling increasing water temperature, in closed system simulations (using PHREEQE code), allows a better determination and discussion of the solution deep conditions. Results appear to indicate the existence, at depth, of equilibrium situations among calcite, quartz (and even dolomite) and solution at 40-50 °C. Problems with thermodynamic data, mainly referred to dolomite, and related to some assumptions in modeling calculations are also shown.

1. INTRODUCCION

La aplicación de técnicas geotermométricas clásicas para la evaluación de la temperatura en profundidad de sistemas hidrotermales de baja-muy baja entalpía, en especial si están asociados a materiales carbonatado-evaporíticos, presenta una serie de problemas específicos dentro de la investigación de sistemas geotermales.

Estas dificultades, referidas al estudio del potencial geotérmico de la provincia de Teruel, han sido puestas de manifiesto en distintos trabajos previos (p. ej. Fernández, 1987; Fernández et al., 1988) en los que se ha constatado la inaplicabilidad de muchos de los geotermómetros actualmente existentes. Esta situación ha conducido, lógicamente, a indeterminaciones en la delimitación de la posible temperatura en profundidad de las aguas consideradas.

La actual existencia de técnicas de modelización fisicoquímica permite el planteamiento de cálculos geotermométricos considerando la solución acuosa como sistema multicomponente, a la vez que facilita la evaluación de los posibles equilibrios minerales responsables del quimismo de las aguas. Estas técnicas, aunque no necesitan la realización de supuestos iniciales sobre la existencia de determinados equilibrios (a diferencia de lo que ocurre con los geotermómetros clásicos), requieren el planteamiento de un estudio fisicoquímico más detallado e, ineludiblemente, la introducción de otro tipo de supuestos.

Las características geoquímicas y fisicoquímicas de Los Baños de Teruel (Fernández, 1987; Fernández et al., 1988; Auqué et al., 1991) hacen de estos manantiales un adecuado marco para la realización de una evaluación y discusión preliminar de estas metodologías, incluyendo algunos geotermómetros no utilizados en los trabajos previos mencionados.

2. SITUACION Y ANTECEDENTES

Los manantiales de Los Baños de Teruel se sitúan a 6 km de la ciudad de Teruel, por la carretera de Alfambra, en las ruinas de lo que fue un antiguo balneario. Geológicamente estos manantiales se sitúan en la convergencia de las fosas del Jiloca y Alfambra, asociadas a la falla de Conclud, de dirección NO-SE. Las surgencias aparecen en el aluvial del río Alfambra, a escasos metros de unos materiales carbonatados jurásicos que afloran a favor de la citada falla.

Composicionalmente se trata de aguas de características sulfatado-cálcicas, con una temperatura de surgencia y quimismo muy homogéneo en todos los manantiales analizados (los análisis químicos, así como la exposición de la metodología analítica y de muestreo seguida pueden verse en Auqué et al., 1991, en este mismo volumen). Estos caracteres geoquímicos generales ponen de manifiesto la circulación preferente de estas aguas a través de materiales predominantemente sulfatado-carbonatados, litologías presentes en la zona y correspondientes a materiales jurásicos.

Estudios isotópicos realizados por IGME (1982) indican un tiempo de residencia de estas aguas en el acuífero superior a 30 años, siendo poco factible la presencia de procesos de mezcla con aguas más frías durante el ascenso de las soluciones termales.

Un análisis fisicoquímico detallado pone de manifiesto la existencia de ligeros estados de subsaturación de las soluciones respecto a calcita y una mayor subsaturación respecto a dolomita, yeso y anhidrita. El cuarzo constituye la única forma alotrópica de la sílice que presenta una situación de sobresaturación en estos manantiales.

Estos resultados son coherentes con la existencia de procesos de disolución de yeso y/o anhidrita y de disolución incongruente de dolomita; procesos que implican, a su vez, la necesidad de una precipitación (saturación) de calcita durante el transcurso de los fenómenos de interacción heterogénea de estas soluciones.

3. APLICACION DE TECNICAS GEOTERMOMETRICAS CLASICAS. RESULTADOS

En la tabla 1 se indican los resultados de algunos de los geotermómetros más ampliamente utilizados en trabajos de prospección geotérmica. Los geotermómetros catiónicos Na-K y Na-K-Ca proporcionan temperaturas considerablemente elevadas (en el caso de este último geotermómetro se ha utilizado el coeficiente $\beta=1/3$. Si se emplea el otro coeficiente propuesto por los autores de este geotermómetro, $\beta=4/3$, las temperaturas

obtenidas todavía son mayores) y, obviamente, carentes de cualquier significado realista en el contexto geológico en el que se encuadra el sistema estudiado.

Tabla 1: Resultados de temperatura (en °C) obtenidos por distintos geotermómetros para cuatro muestras de los manantiales de Los Baños (Teruel).

MUESTRAS		T-1	T-2	T-3	T-4
SiO _{2(aq)} -cuarzo (Fournier & Potter, 1982)		36.2	37.3	38.4	37.3
Na-K	White (1965)	164	195	167	180
	Ellis (1970)	175	205	178	191
	Michard (1990)	153	174	155	164
Na-K-Ca (Fournier & Truesdell, 1973)*		81	81	82	82
Na-K-Ca-Mg (Fournier & Potter, 1979)		79	72	72	72
K-Mg (Giggenbach et al., 1983)		16	21	16	18
Na-Li (Fouillac & Michard, 1981)		81	81	82	82

* Calculado con $\beta = 1/3$ (Fournier & Truesdell, 1973).

El resto de geotermómetros catiónicos empleados proporciona una temperatura de 80 °C en el caso del Na-Li, entre 70 y 80 °C en el del geotermómetro Na-K-Ca-Mg e inferiores a la propia temperatura de surgencia en el caso del geotermómetro K-Mg.

De los geotermómetros basados en la solubilidad de las distintas formas alotrópicas de la sílice, únicamente el SiO_{2(aq)}-cuarzo proporciona temperaturas en el reservorio por encima de la de surgencia (tabla 1). Los geotermómetros SiO_{2(aq)}-calcedonia y SiO_{2(aq)}-sílice amorfa señalan temperaturas por debajo de las de surgencia o negativas, hecho perfectamente lógico si consideramos los estados de subsaturación de estas aguas (Auqué et al., 1991) respecto a estos dos minerales en condiciones de surgencia.

Por último, se debe señalar que la aplicación del geotermómetro calcita-dolomita, basado en la existencia de una situación de equilibrio en profundidad entre estos dos minerales (no incluido en la tabla 1) y desarrollado por Marini et al. (1987), proporciona valores en torno a los 80-90 °C.

Como puede apreciarse, los resultados obtenidos por este tipo de técnicas son tremendamente variables y con frecuencia aberrantes (por ejemplo cuando la temperatura obtenida en el reservorio es menor que en surgencia) o irreales por excesivamente elevados. En este último caso se encuentran los obtenidos por el geotermómetro Na-K (mediante cualquiera de los calibrados presentados en la tabla 1). El supuesto de aplicación de este geotermómetro se basa en la existencia de un equilibrio entre feldespatos en profundidad (o al menos a una asignación conceptual de los calibrados obtenidos empíricamente con este equilibrio¹); y este supuesto no parece adecuarse a las condiciones del sistema analizado, en el que las concentraciones de Na^+ y K^+ no están tamponadas por un equilibrio mineral de este tipo. La obtención de temperaturas excesivamente elevadas al aplicar este geotermómetro a sistemas de baja entalpía es un hecho frecuentemente constatado (p. ej. Henley et al., 1984) en la bibliografía sobre el tema.

Similares problemas presenta el geotermómetro Na-K-Ca, que además ha sido calibrado con aguas de temperaturas comprendidas entre 100 y 300 °C (Fournier & Truesdell, 1973) y, por tanto, de difícil extrapolación a sistemas con temperaturas inferiores a 100 °C. Por otra parte, es un geotermómetro especialmente sensible a los contenidos de Mg^{2+} de las aguas.

Eliminando los resultados de los geotermómetros ya señalados, quedan los definidos por el $\text{SiO}_{2(\text{aq})}$ -cuarzo, Na-Li, Na-K-Ca-Mg y calcita-dolomita, en los que el rango de temperaturas obtenido oscila entre los 40 °C definidos por el primero y los 80 °C proporcionados por los otros tres. De todos estos geotermómetros, únicamente el $\text{SiO}_{2(\text{aq})}$ -cuarzo y calcita-dolomita tienen una base termodinámica perfectamente definida por una situación de equilibrio concreto; los otros dos son geotermómetros empíricos, no estrictamente definidos en cuanto a las reacciones de equilibrio descriptivas de su comportamiento. E incluso el geotermómetro Na-Li posiblemente responda a reacciones de intercambio catiónico con minerales arcillosos (Fouillac & Michard, 1981).

Los resultados del geotermómetro $\text{SiO}_{2(\text{aq})}$ -cuarzo se encuentran limitados por las bajas concentraciones de SiO_2 de las aguas, hecho que constituye uno de los principales problemas de este geotermómetro (p. ej. D'Amore et al., 1987). Por otro lado, el geotermómetro calcita-dolomita presenta una excesiva dependencia de los datos termodinámicos utilizados para su calibrado, tal como indican sus propios autores (Marini et al., 1987).

En definitiva, los rangos de incertidumbre obtenidos al aplicar las técnicas geotermométricas clásicas a sistemas de baja entalpía como el que nos ocupa son considerables. En nuestro caso, el rango de temperaturas que proporcionan los geotermómetros que pueden considerarse como más plausibles oscila entre 40 y 80 °C, y aunque

¹ Los calibrados de este geotermómetro son normalmente empíricos y las aguas utilizadas para su realización no suelen corresponder a sistemas de características análogas a las del aquí estudiado. Por ejemplo, el calibrado de Michard (1990) ha sido realizado a partir de sistemas geotermales instalados exclusivamente en sistemas graníticos y de aguas alcalinas.

existe una mayoría de geotermómetros que apuntan hacia el extremo superior de este rango, también es cierto que la mayoría de ellos presentan importantes incertidumbres acerca de la fiabilidad de sus resultados.

4. MODELIZACION GEOTERMOMETRICA

4.1. Supuestos básicos.

La utilización de metodologías de modelización geoquímica para el cálculo de la temperatura en profundidad constituye uno de los tipos de simulación termodinámica realizable mediante códigos de pautas de reacción. Este tipo de cálculo supone considerar el aumento de temperatura como proceso irreversible desencadenante de la evolución (pauta de reacción) del estado fisicoquímico de la solución.

El planteamiento consiste en que, a partir de la especiación de la solución a la temperatura y pH de surgencia y en pasos sucesivos de incremento de T, se calcule una nueva especiación y el cambio de pH (producido por la interdependencia entre las constantes de equilibrio de disociación de ácidos y bases débiles y la temperatura; Merino, 1975; Michard & Fouillac, 1980; Auqué et al., 1989). Los resultados de estos cálculos pueden expresarse como una evolución de los índices de saturación (I.S.) respecto a distintos minerales, definiendo para cada uno de ellos la temperatura a la que se alcanza un estado de equilibrio (I.S. = 0).

Una descripción más detallada de los fundamentos de este tipo de técnicas puede verse en los trabajos de los autores que acabamos de mencionar. No obstante, conviene tener presente que un cálculo en las condiciones que acabamos de definir está implícitamente asumiendo una serie de supuestos sobre el funcionamiento del sistema. Estos supuestos son comunes al resto de técnicas geotermométricas utilizadas y asumen la inexistencia de procesos de reequilibrio durante el ascenso de las soluciones (procesos de reequilibrio que hacen referencia tanto a fenómenos de disolución/precipitación como gasificación/desgasificación de las soluciones). Esta serie de asunciones implican la realización del cálculo de modelización en condiciones de sistema cerrado según el concepto de Michard & Fouillac (1980).

Respecto al sistema natural estudiado, estos supuestos implican que no existen cambios de quimismo desde el reservorio a la surgencia y tampoco hay modificaciones sustanciales del pH debido a procesos de pérdida de CO₂.

4.2. Aplicación a los manantiales de Los Baños.

Según estos supuestos y utilizando el código PHREEQE (Parkhurst et al., 1980), se ha procedido al cálculo de la evolución de las soluciones acuosas analizadas frente a aumentos progresivos de temperatura. En la figura 1 A se representa la evolución de

distintos I.S. frente a la temperatura para la muestra T-1 (Auqué et al., 1991). Y en la tabla 2 se presentan los valores numéricos correspondientes a este cálculo.

Como puede constatarse, la calcedonia, mineral con I.S. negativo en condiciones de surgencia, evoluciona hacia valores todavía más subsaturados al aumentar la temperatura. Tanto el yeso como la anhidrita, también subsaturados en condiciones de surgencia, presentan una tendencia opuesta disminuyendo el estado de subsaturación al aumentar la temperatura; no obstante, a 100 °C todavía presentan importantes grados de subsaturación, siendo necesario aumentar la temperatura hasta valores irreales para alcanzar una situación de equilibrio.

Tabla 2: Modelización geotermométrica de una de las muestras tomadas en los manantiales de Los Baños de Teruel. Evolución de los índices de saturación frente a la temperatura de distintos minerales.

Temp (°C)	20.7	40.0	60.0	80.0	100.0
pH	6.720	6.614	6.611	6.638	6.709
Calcita	-0.271	-0.108	0.114	0.376	0.664
Aragonito	-0.418	-0.242	-0.007	0.266	0.563
Dolomita	-1.011	-0.495	-0.013	0.403	0.735
Yeso	-0.700	-0.719	-0.691	-0.625	-0.535
Anhidrita	-0.966	-0.803	-0.657	-0.532	-0.427
Fluorita	-0.837	-1.099	-1.352	-1.591	-1.823
Cuarzo	0.279	-0.006	-0.267	-0.499	-0.707
Calcedonia	-0.221	-0.433	-0.627	-0.799	-0.954

Los únicos minerales que alcanzan una situación de equilibrio en el rango de temperaturas de 20 a 100 °C son calcita, aragonito, dolomita y cuarzo. Y a ellos nos referiremos específicamente al comentar los resultados de la modelización geotermométrica (los correspondientes resultados pueden verse en las figuras 1 y 2).

Las temperaturas a las que las soluciones consideradas alcanzan el equilibrio con calcita oscilan entre 40 y 50 °C, la dolomita entre 50 y 55 °C y el cuarzo en torno a los 40 °C. Los rangos obtenidos son, a priori, más estrechos que los definidos por las técnicas geotermométricas clásicas, coincidiendo además con los resultados del geotermómetro SiO₂-cuarzo.

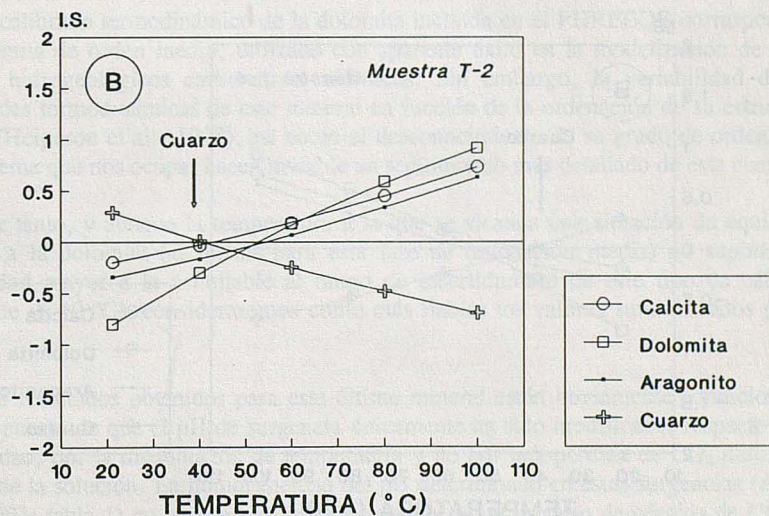
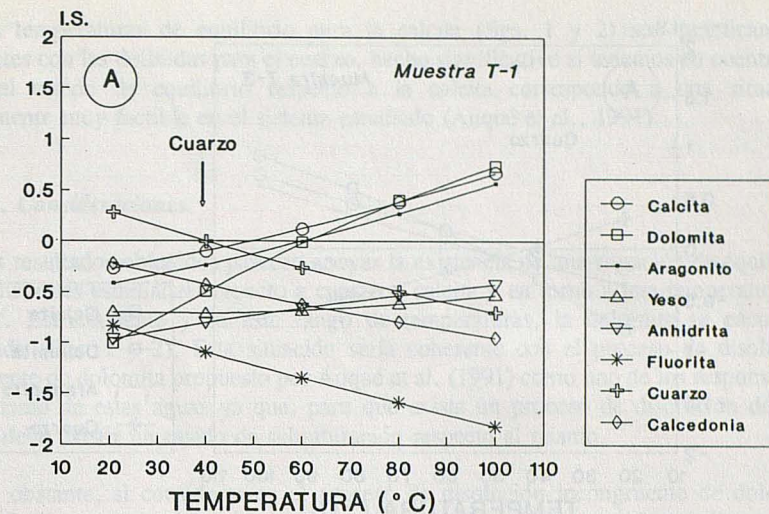


Figura 1. Evolución de los índices de saturación frente a la temperatura para las aguas de los manantiales de Teruel y correspondientes a las muestras presentadas por Auqué et al. (1991). A.- Muestra T-1. B.-Muestra T-2.

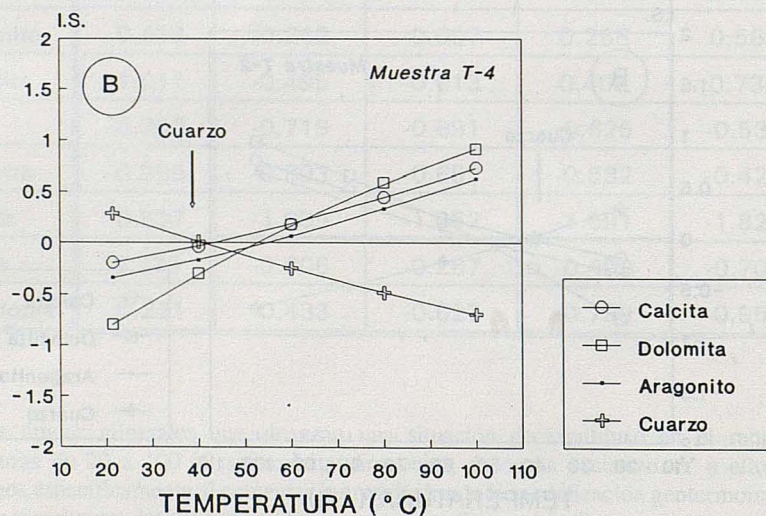
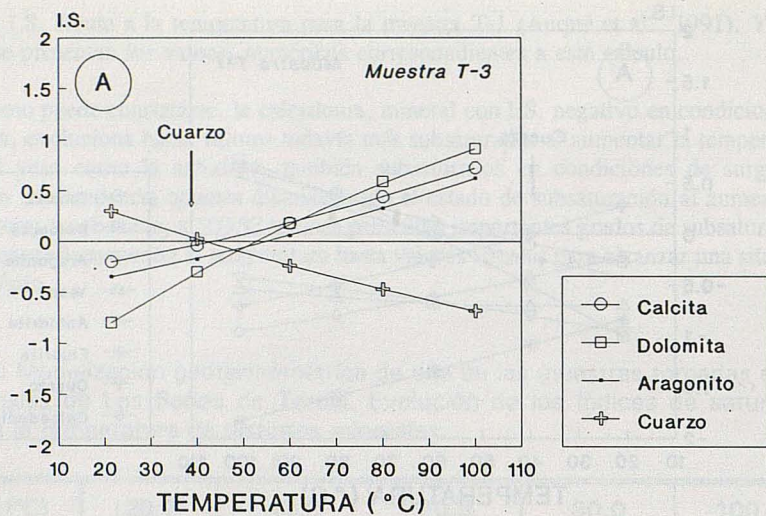


Figura 2. Evolución de los índices de saturación frente a la temperatura para las aguas de los manantiales de Teruel y correspondientes a las muestras presentadas por Auqué et al. (1991). A.- Muestra T-3. B.-Muestra T-4.

Las temperaturas de equilibrio para la calcita (figs. 1 y 2) son prácticamente coincidentes con las definidas para el cuarzo, hecho significativo si tenemos en cuenta que además el estado de equilibrio respecto a la calcita corresponde a una situación aparentemente muy factible en el sistema estudiado (Auqué et al., 1991).

4.3. Consideraciones.

Los resultados obtenidos parecen apoyar la existencia de una situación de equilibrio de las soluciones estudiadas respecto a cuarzo y calcita y en torno a una temperatura de 40-50 °C. Estrictamente y en este rango de temperaturas, la dolomita se encuentra subsaturada (figs. 1 y 2). Esta situación sería coherente con el proceso de disolución incongruente de dolomita propuesto por Auqué et al. (1991) como uno de los responsables del quimismo de estas aguas ya que, para que exista un proceso de disolución de este mineral, debe existir un estado de subsaturación respecto al mismo.

No obstante, si consideramos el proceso de disolución incongruente de dolomita como un fenómeno de dedolomitización catalizado por la adición de sulfato cálcico a las aguas, esta situación también sería coherente con la existencia de un cuasi-equilibrio calcita-dolomita en profundidad (equilibrio en estado estacionario).

El calibrado termodinámico de la dolomita incluida en el PHREEQE corresponde a una dolomita de orden medio, utilizado con aparente éxito en la modelización de otros sistemas hidrogeológicos carbonatado-sulfatados. Sin embargo, la variabilidad de las propiedades termodinámicas de este mineral en función de la ordenación de su estructura mineral (Helgeson et al., 1978), así como el desconocimiento de su grado de ordenación en el sistema que nos ocupa, hacen inviable un seguimiento más detallado de esta cuestión.

Por tanto, y aunque la temperatura a la que se alcanza una situación de equilibrio respecto a la dolomita (al menos para esta fase de ordenación media) no supone una variabilidad mayor a la asimilable al rango de incertidumbre de este tipo de cálculos (rangos de ± 10 °C), consideraremos como más fiables los valores suministrados por la calcita.

Los resultados obtenidos para este último mineral están obviamente condicionados por el supuesto de que el pH de surgencia únicamente ha sido modificado, respecto al de profundidad, por la disminución de temperatura y no por una pérdida de CO₂ durante el ascenso de la solución. La homogeneidad del pH determinado en estas surgencias (Auqué et al., 1991; tabla 1) no parece indicar la existencia de un proceso de pérdida de CO₂, al menos de manera selectiva para las distintas surgencias. En cualquier caso, pérdidas de CO₂ que supusieran un aumento del pH en surgencia de 0.2 unidades sobre el existente en profundidad supondría un aumento de 12 °C sobre los 40 °C inicialmente obtenidos para una situación de equilibrio con calcita, lo que no supone una excesiva variación sobre el rango de 40-50 °C propuestos al principio.

CONCLUSIONES

La aplicación de las técnicas geotermométricas clásicas a sistemas geotermales de baja-muy baja entalpía presenta numerosos problemas, fundamentalmente referidos a la inexistencia de los equilibrios solución-mineral apropiados en estos sistemas de baja temperatura, y/o a la inadecuación de los calibrados planteados para los distintos geotermómetros en este tipo de sistemas.

La utilización de técnicas de modelización geotermométrica permite una mejor definición en el análisis de las posibles condiciones de la solución en profundidad. En el caso de Los Baños de Teruel se obtiene una situación factible de equilibrio entre calcita, cuarzo e incluso dolomita, a una temperatura de 40-50 °C.

No obstante, también se han puesto de manifiesto las dificultades producidas por los calibrados termodinámicos, especialmente en el caso de la dolomita; dificultades que se reflejan igualmente en el propio calibrado del geotermómetro calcita-dolomita clásico.

La definición de las condiciones de temperatura en profundidad para estos manantiales descansa en una serie de supuestos factibles pero necesitados de una más amplia testificación. Son necesarios análisis de modelización más detallados (actualmente en fase de realización) incluyendo elementos y sistemas composicionales no utilizados en este trabajo.

BIBLIOGRAFIA

AUQUE, L.F.; FERNANDEZ, J.; TENA, J.; MANDADO, J.; GIMENO, M.J. y TOLOSA, P. (1989). Análisis de los estados de equilibrio termodinámico en el reservorio de las surgencias termales de Fitero (Navarra) y Arnedillo (Rioja). *Rev. Soc. Geol. Esp.*, 2, 125-132.

AUQUE, L.F.; TENA, J.; GIMENO, M.J.; MANDADO, J. y FERNANDEZ, J. (1991). Las surgencias termales de los Baños (Teruel). II. Cálculos geotermométricos. *Revista de la Academia de Ciencias de Zaragoza*, en este mismo volumen.

D'AMORE, F.; FANCELLI, R. & CABOI, R. (1987). Observations on the application of chemical geothermometers to some hydrothermal systems in Sardinia. *Geothermics*, 16, 271-282.

ELLIS, A.J. (1970). Quantitative interpretation of chemical characteristics of geothermal systems. *Geothermics*, 2, 516-528.

FERNÁNDEZ, J. (1987). *Ambiente geológico y caracterización geoquímica de las aguas termales en la provincia de Teruel*. Tesis de Licenciatura, Fac. Ciencias, Universidad de Zaragoza, 163 pp. (No publicado).

FERNÁNDEZ, J.; AUQUÉ, L.F. Y MANDADO, J. (1988). Quimismo de las aguas termales y mineromedicinas de la provincia de Teruel (España). *Rev. Geociencias Aveiro*, 3, 1-2, 249-264.

FOUILLAC, C. & MICHARD, G. (1981). Sodium-lithium ratio in water applied to the geothermometry of geothermal waters. *Geothermics*, 10, 55-70.

FOURNIER, R.O. & TRUESDELL, A.H. (1973). An empirical Na-Ca-K geothermometer for natural waters. *Geochim. Cosmochim. Acta*, 37, 1255-1275.

FOURNIER, R.O. & POTTER, R.W. (1979). Magnesium correction to the Na-K-Ca chemical geothermometer. *Geochim. Cosmochim. Acta*, 43, 1543-1550.

FOURNIER, R.O. & POTTER, R. W. (1982). A revised expanded (quartz) geothermometer. *Geoth. Res. Council Bull.*, 11, 3-9.

GIGGENBACH, W.F.; GONFIANTINI, R.; JANGI, B.L. & TRUESDELL, A.H. (1983). Isotopic and chemical composition of Parbati Valley geothermal discharges, N.W. Himalaya, India. *Geothermics*, 12, 199-222.

HELGESON, H.C.; DELANY, J.M.; NESBITT, H.W. & BIRD, D.K. (1978). Summary and critique of the thermodynamic properties of rock-forming minerals. *Am. J. Sci.*, 278, 1-229.

HENLEY, R.W.; TRUESDELL, A.H. and BARTON, P.B. with a contribution by WHITNEY, J.A. (1984). *Fluid Mineral Equilibria in Hydrothermal Systems. Reviews in Economic Geology I*. Published by the Soc. of Economic Geologists.

IGME (1982). *Estudio de las manifestaciones termales de Extremadura-Salamanca-Aragón-Rioja, orientadas a su posible explotación como recursos geotérmicos*. Memoria Tomo II, Aragón-Rioja. (No publicado).

MARINI, L.; CHIODINI, G. & CIONI, R. (1986). New geothermometers for carbonate evaporite geothermal reservoirs. *Geothermics*, 5, 77-86.

MERINO, E. (1975). Diagenesis in Tertiary sandstones from Kettleman Noth Dome, California. II. Interstitial solutions: distribution of aqueous species at 100 °C and chemical reaction to diagenetic mineralogy. *Geochim. Cosmochim. Acta*, 39, 1629-1645.

MICHARD, G. (1990). Behaviour of major elements and some trace elements (Li, Rb, Cs, Fe, Mn, W, F) in deep hot waters from granitic areas. *Chem. Geol.*, 89, 117-134.

MICHARD, G. et FOUILLAC, C. (1980). Contrôles de la composition chimique des eaux thermales sulfurées sodiques du sud de la France. In: *Géochimie des interactions entre les eaux, les minéraux et les roches* (Y. Tardy, ed.). Masson, Paris. pp. 147-166.

PARKHURST, D.L.; THORSTENSON, D.C. & PLUMMER, L.N. (1980). *PHREEQE. A computer program for geochemical calculations*. U.S. Geological Survey Water Resources Investigations, 80-96, 157 pp.

PLUMMER, L.N. & BUSENBERG, E. (1982). The solubilities of calcite, aragonite and vaterite in CO₂-H₂O solutions between 0 and 90 °C, and an evaluation of the aqueous model for the system CaCO₃-CO₂-H₂O. *Geochim. Cosmochim. Acta*, 46, 1011-1040.

WHITE, D.E. (1965). Saline waters from sedimentary rocks. *Mem. Am. Assoc. Pet. Geol.*, 4, 342-366.

AGRADECIMIENTOS

Este trabajo fue financiado en parte por una Beca de la Diputación General de Aragón (L.F. Auqué).