

UNIVERSIDAD NACIONAL DEL LITORAL
FACULTAD DE INGENIERÍA QUÍMICA

TESIS PRESENTADA COMO PARTE DE LOS REQUISITOS DE LA UNIVERSIDAD
NACIONAL DEL LITORAL PARA LA OBTENCIÓN DEL GRADO ACADÉMICO DE

Doctor en Matemática

**Ecuaciones de Evolución Nolineales, Problemas
Inversos Asociados y Aplicaciones a un Modelo para
Materiales con Memoria de Forma**

Departamento de Matemática – FIQ
Programa Especial de Matemática Aplicada – INTEC

Pedro Morin

Dirigida por Rubén D. Spies

1998

Resumen

Para muchos problemas de interés práctico que involucran un fenómeno físico es posible formular un modelo matemático que aproxime satisfactoriamente la dinámica del mismo. En algunas oportunidades este modelo puede escribirse como una ecuación de evolución en un espacio de estados de dimensión, tal vez, infinita. Sucede con frecuencia que esta ecuación posee ciertos parámetros desconocidos que se desean estimar a partir de observaciones del fenómeno físico en cuestión. Estos problemas, llamados *de identificación de parámetros* han recibido mucha atención en los últimos años. A menudo pueden formularse como problemas de optimización en espacios de estados de dimensión infinita, y consisten en determinar el parámetro que minimice una función de costo apropiada sobre algún conjunto de parámetros *admisibles*.

En este trabajo se desarrolla un algoritmo de identificación de parámetros utilizando cuasilinealización para un problema de Cauchy abstracto no lineal general, se prueban teoremas de convergencia y el mismo se aplica a la identificación de los parámetros no físicos que aparecen en las ecuaciones que modelan el comportamiento termomecánico de un sólido unidimensional con memoria de forma.

La organización de la tesis es la siguiente: En el Capítulo 1 se hace una breve introducción a las ecuaciones de evolución en espacios de Banach y en las últimas dos secciones del mismo se presentan los problemas que se atacarán en los capítulos posteriores.

En el Capítulo 2 se propone el algoritmo basado en cuasilinealización para resolver el problema de identificación de parámetros que se presenta en la Sección 1.7, y se obtienen teoremas de convergencia bajo la hipótesis de existencia y suavidad de la derivada de Fréchet del estado con respecto al vector de parámetros.

En el Capítulo 3 se presentan teoremas que dan condiciones suficientes sobre los operadores que definen la ecuación de evolución involucrada en el problema de identificación, para la existencia y suavidad de la derivada de Fréchet mencionada arriba.

En el Capítulo 4 se introduce un sistema de ecuaciones que modela el comportamiento de un sólido unidimensional con memoria de forma, se presenta una formulación abstracta como ecuación de evolución en un espacio de Banach determinado, y se enuncian los resultados conocidos hasta el momento para la misma.

En el Capítulo 5 se presenta un algoritmo basado en aproximaciones modales para la resolución numérica de las ecuaciones del modelo del Capítulo 4, con resultados de convergencia y experimentos numéricos ilustrativos. La inclusión de este capítulo es necesaria, dado que para aplicar el algoritmo de cuasilinealización al sistema se deben poder resolver numéricamente las ecuaciones, y los algoritmos conocidos hasta el momento para realizar esta tarea no eran suficientemente satisfactorios con respecto al tiempo de ejecución.

En el Capítulo 6 se aplica el algoritmo de cuasilinealización para identificar los parámetros no físicos del modelo presentado en el Capítulo 4, haciendo previamente un análisis de la identificabilidad de los mismos.

Para finalizar, en el Capítulo 7 se aplica el algoritmo desarrollado para la identificación de parámetros a la resolución numérica de un problema de control terminal asociado al modelo para materiales con memoria de forma presentado en el Capítulo 4. Este problema se presenta previamente en la Sección 1.8.

Índice General

1	Introducción y Preliminares	1
1.1	Ecuaciones de Evolución Lineales	1
1.2	Semigrupos Fuertemente Continuos	2
1.3	Caracterización de Generadores Infinitesimales	6
1.4	Semigrupos Diferenciables y Analíticos	9
1.5	Potencias Fraccionarias de Operadores Cerrados	14
1.6	Ecuaciones de Evolución Nolineales	17
1.7	Identificación de Parámetros	20
1.8	Un Problema de Control	22
2	Un Método Directo para la Identificación de Parámetros	25
2.1	Presentación del Algoritmo de Cuasilinealización	27
2.2	Convergencia del Algoritmo de Cuasilinealización	33
2.3	Presentación Clásica del AC	38
3	Diferenciabilidad con Respecto a Parámetros	41
3.1	Existencia de la Derivada de Fréchet de $z(t;q)$	42
3.2	Continuidad Lipschitz de la Derivada de Fréchet	48
4	Materiales con Memoria de Forma	51
4.1	Introducción	51
4.2	Modelos Matemáticos	53
4.3	Formulación Abstracta del Problema	59

4.4	Resultados Previos	61
5	Aproximaciones Numéricas	65
5.1	Aproximaciones Modales	65
5.2	Discretización Temporal	71
5.3	Resultados Numéricos	74
6	Una Aplicación del Algoritmo de Cuasilinealización	83
6.1	Identificabilidad de Parámetros	83
6.2	No Identificabilidad de Parámetros	85
6.3	Estimación de los Parámetros	86
7	Resolución Numérica del Problema de Control	95
7.1	Hipótesis para la Convergencia del Algoritmo	96
7.2	Experimentos Numéricos	97
8	Conclusiones y Planes para el Futuro	103

Notación

La notación que utilizaremos a lo largo de este trabajo es, en su mayor parte, estándar.

Si X es un espacio de Banach, denotaremos la norma de ese espacio con $\|\cdot\|_X$, ó simplemente con $\|\cdot\|$, cuando quede claro del contexto, y denotaremos con X^* al espacio dual topológico de X , es decir al espacio de todos los funcionales lineales y continuos sobre X . Análogamente, si X es un espacio de Hilbert $\langle \cdot, \cdot \rangle_X$ ó $\langle \cdot, \cdot \rangle$ denotarán el producto interno en X . Si X e Y son espacios de Banach, $\mathcal{L}(X, Y)$ denotará el espacio de Banach de todos los operadores lineales y acotados de X en Y . Para $A \in \mathcal{L}(X, Y)$, $\|A\|_{\mathcal{L}(X, Y)}$ denotará la norma de operadores en $\mathcal{L}(X, Y)$. Si $X = Y$, se denotará con $\mathcal{L}(X)$ al espacio $\mathcal{L}(X, Y)$. Si A es un operador lineal de un espacio de Banach X en X , entonces denotaremos con $D(A)$, $R(A)$, $\sigma(A)$, $\sigma_p(A)$, y $\rho(A)$ al dominio, rango, espectro, espectro puntual y resolvente de A , respectivamente, donde $\rho(A) = \{\lambda \in \mathbb{C} : (\lambda I - A)^{-1} \in \mathcal{L}(X)\}$, $\sigma(A) = \mathbb{C} \setminus \rho(A)$ y $\sigma_p(A) = \{\lambda \in \mathbb{C} : \lambda I - A \text{ no es inyectivo}\}$. Si M es un operador o una matriz, M^* denotará el operador adjunto o la matriz conjugada transpuesta, respectivamente. Para $\Omega \subset \mathbb{R}^n$ y $1 < p < \infty$, si X es un espacio de Banach, $L^p(\Omega; X)$ denotará el espacio de todas las funciones fuertemente medibles $f : \Omega \rightarrow X$ tal que $\|f\|_{L^p(\Omega; X)} = \left(\int_{\Omega} \|f(x)\|_X^p dx\right)^{1/p} < \infty$ (ver Sección 1.2). Cuando $X = \mathbb{R}$ denotaremos con $L^p(\Omega)$ a $L^p(\Omega; X)$. Para un entero no negativo k , $C^k(\Omega; X)$ denotará al espacio de todas las funciones continuas de Ω en X con k derivadas continuas en Ω . Los espacios $H^k(\Omega)$ y $H_0^k(\Omega)$ denotarán los espacios estándar de Sobolev definidos por $H^k(\Omega) = \{f \in L^2(\Omega) : f^{(j)} \in L^2(\Omega), j = 1, 2, \dots, k\}$ y $H_0^k(\Omega) = \{f \in H^k(\Omega) : f^{(j)}|_{\partial\Omega} = 0, j = 0, 1, \dots, k-1\}$, respectivamente.

Capítulo 1

Introducción y Preliminares

Varios problemas de áreas diversas como en el estudio de materiales y estructuras inteligentes, en Física, en Geología, Ecología, Ingeniería Química, Biología, etc. (ver [7]) originan problemas cuya formulación abstracta resulta en una ecuación diferencial ordinaria no lineal en un espacio de Banach, donde además en el término no lineal aparece un vector de parámetros desconocidos.

En este capítulo introduciremos brevemente algunos resultados sobre ecuaciones de evolución en espacios de Banach. Comenzaremos estudiando problemas lineales, lo que nos conducirá inmediatamente al estudio de la teoría de *Semigrupos Fuertemente Continuos*. Luego consideraremos problemas no lineales. Concluiremos el capítulo con una sección donde definiremos explícitamente un problema de identificación de parámetros donde el parámetro a identificar aparece en la parte no lineal de la ecuación, y con otra sección donde plantearemos un problema de control asociado a esta clase de ecuaciones.

1.1 Ecuaciones de Evolución Lineales

Consideremos una ecuación de evolución de la forma

$$\begin{cases} \frac{dz(t)}{dt} = Az(t) & t > 0 \\ z(0) = z_0, \end{cases} \quad (1.1)$$

donde se $z(t)$ es un elemento de un espacio de Banach $(Z, \|\cdot\|)$, A es un operador lineal de $D(A) \subset Z$ en Z , y $z_0 \in Z$. Esta ecuación de evolución se denomina *Problema de Cauchy*

Abstracto para A con valor inicial $z_0 \in Z$. Diremos que $z(t)$, $t \geq 0$ es una *solución clásica* o simplemente una solución de (1.1) si $z(t)$ es una función con valores en Z , continua para $t \geq 0$, y tal que para $t > 0$ es continuamente diferenciable, $z(t) \in D(A)$, y satisface (1.1).

Como ejemplo podemos mencionar el caso en que $Z = L^2(0, 1)$, $Az = k \Delta z$, $k \in \mathbb{R}$ para $z \in D(A) = \{z \in H^2(0, 1) : z'(0) = z'(1) = 0\}$. El problema (1.1) correspondiente a este operador equivale formalmente al problema de valores iniciales y de frontera

$$\begin{cases} \frac{\partial u}{\partial t}(x, t) = k \frac{\partial^2 u}{\partial x^2}(x, t), & (x, t) \in (0, 1) \times (0, \infty) \\ \frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(1, t) = 0, & t \in (0, \infty) \\ u(x, 0) = u_0(x), & x \in [0, 1] \end{cases} \quad (1.2)$$

donde la ecuación diferencial involucrada es conocida como ecuación del calor.

El primer cuestionamiento en torno al problema de Cauchy (1.1) es, naturalmente, el de la existencia y unicidad de soluciones.

Podemos observar que si Z es de dimensión finita o si A es un operador acotado, entonces el problema (1.1) tiene solución única para todo $z_0 \in Z$ y ésta viene dada por $z(t) = e^{At}z_0$, donde $e^{At} = I + tA + \frac{t^2}{2}A^2 + \frac{t^3}{3!}A^3 + \dots$. La convergencia de esta serie se da en la topología uniforme de $\mathcal{L}(Z)$.

La rama de la Matemática que estudia la existencia, unicidad, regularidad, etc. de las soluciones del problema de Cauchy (1.1) en espacios de Banach se conoce como Teoría de Semigrupos.

1.2 Semigrupos Fuertemente Continuos

Supongamos que para un operador dado A , $z(t)$ es una solución de (1.1) con valor inicial z_0 . Entonces, como el problema es autónomo (pues estamos suponiendo que el operador A es independiente de t), se debe cumplir que $z(t) = T(t)z_0$, donde $T(t)$ es cierto operador lineal, es decir, el valor de la solución en t sólo depende de la condición inicial y lo hace en forma lineal. Además, el estado $z(t + s)$ alcanzado en tiempo $t + s$ a partir de z_0 es el mismo que el estado alcanzado en tiempo s partiendo del estado inicial $z(t)$, es

decir $T(t+s)z_0 = z(t+s) = T(s)z(t) = T(s)T(t)z_0$. Para que el problema sea “bien planteado” (well posed) en el sentido de que la solución dependa continuamente de la condición inicial, pediremos a $T(t)$ que sea un operador acotado en Z .

Esta disgresión sobre la forma de las soluciones del problema (1.1) nos conduce al concepto de C_0 -Semigrupo o Semigrupo Fuertemente Continuo.

Definición 1.1 *Un semigrupo fuertemente continuo sobre un espacio de Banach Z o simplemente un C_0 -semigrupo sobre Z es una familia $\{T(t)\}_{t \geq 0}$ de operadores lineales y acotados sobre Z que satisface las siguientes propiedades:*

- (i) $T(t+s) = T(t)T(s)$, para $t, s \geq 0$;
- (ii) $T(0) = I$;
- (iii) $\|T(t)z_0 - z_0\| \rightarrow 0$, cuando $t \rightarrow 0^+$, para todo $z_0 \in Z$.

Notemos que si $A \in \mathcal{L}(Z)$, $T(t) = e^{At}$ es un C_0 -semigrupo.

Asociado a todo C_0 -semigrupo existe lo que llamamos su *generador infinitesimal*.

Definición 1.2 *El generador infinitesimal de un C_0 -semigrupo $T(t)$ sobre Z es el operador lineal definido por*

$$D(A) = \left\{ z \in Z : \text{existe } \lim_{t \rightarrow 0^+} \frac{T(t)z - z}{t} \right\}$$

y para $z \in D(A)$, $Az = \lim_{t \rightarrow 0^+} \frac{T(t)z - z}{t}$.

La mayoría de los resultados de aquí al final del presente capítulo pueden encontrarse en [42]. Los próximos tres teoremas enuncian las propiedades básicas de los C_0 -semigrupos y sus generadores infinitesimales.

Teorema 1.3 *Sean $T(t)$ y $S(t)$ C_0 -semigrupos de operadores lineales acotados sobre un espacio de Banach Z con generadores infinitesimales A y B , respectivamente. Si $A = B$, entonces $T(t) = S(t)$ para todo $t \geq 0$.*

Teorema 1.4 Sea $T(t)$ un C_0 -semigrupo sobre un espacio de Banach Z con generador infinitesimal A . Entonces:

(a) Existen constantes $\omega \geq 0$ y $M \geq 1$ tales que

$$\|T(t)\| \leq Me^{\omega t} \quad \text{para } 0 \leq t < \infty.$$

(b) Para todo $z \in Z$ y $t \geq 0$, se tiene que

$$\lim_{\substack{h \rightarrow 0 \\ t+h \geq 0}} \frac{1}{h} \int_t^{t+h} T(s)z \, ds = T(t)z.$$

(c) Si $z \in Z$ y $t \geq 0$, entonces $\int_0^t T(s)z \, ds \in D(A)$ y además

$$A \left(\int_0^t T(s)z \, ds \right) = T(t)z - z.$$

(d) Si $z \in D(A)$ se tiene que $T(t)z \in D(A)$ para todo $t \geq 0$ y

$$AT(t)z = \frac{d}{dt}T(t)z = T(t)Az.$$

(e) Para cada $z \in D(A)$ y $s, t \geq 0$,

$$T(t)z - T(s)z = \int_s^t T(\tau)Az \, d\tau = \int_s^t AT(\tau)z \, d\tau.$$

Teorema 1.5 Si A es el generador infinitesimal de un C_0 -semigrupo sobre un espacio de Banach Z , entonces A es un operador cerrado y $\bigcap_{n=1}^{\infty} D(A^n)$ es un subespacio denso en Z .

Observación: Notemos que en el enunciado del teorema 1.4 hemos escrito integrales de funciones con valores en un espacio de Banach. Estas integrales se denominan integrales de Bochner y no son más que una generalización de la integral de Lebesgue para funciones con valores en espacios de Banach. Brevemente, si $(\Omega, \mathcal{P}, \mu)$ es un espacio de medida y Z un espacio de Banach, entonces se definen las funciones simples y la integral de Bochner de funciones simples del siguiente modo (ver [21, Capítulo 6]):

Definición 1.6 Sea $(\Omega, \mathcal{P}, \mu)$ un espacio de medida sigma-finita y sea Z un espacio de Banach, entonces:

- (a) Una función $u : \Omega \rightarrow Z$ es simple si existen constantes nonulas $u_i \in Z$ y conjuntos medibles disjuntos E_i , $i = 1, 2, \dots, n$ con $\mu(E_i) < \infty$ tales que

$$u(\omega) = \sum_{i=1}^n u_i \chi_{E_i}(\omega).$$

Donde χ_{E_i} es la función característica de E_i :

$$\chi_{E_i}(\omega) = \begin{cases} 1 & \text{si } \omega \in E_i \\ 0 & \text{si } \omega \notin E_i. \end{cases}$$

- (b) La integral de Bochner de una función simple $u : \Omega \rightarrow Z$ sobre el conjunto $E \in \mathcal{P}$ se define por

$$\int_E u d\mu = \sum_{i=1}^n u_i \mu(E \cap E_i).$$

Luego se definen las funciones fuertemente medibles y las integrales de ellas del siguiente modo:

Definición 1.7 Sea $(\Omega, \mathcal{P}, \mu)$ un espacio de medida sigma-finita y sea Z un espacio de Banach, entonces:

- (a) Una función $u : \Omega \rightarrow Z$ es una función fuertemente medible si existe una sucesión $\{u_n\}$ de funciones simples tales que $\|u_n(\omega) - u(\omega)\| \rightarrow 0$ para casi todo $\omega \in \Omega$ cuando $n \rightarrow \infty$.

- (b) La integral de Bochner de una función fuertemente medible $u : \Omega \rightarrow Z$ es el límite (si existe) de las integrales de Bochner de las funciones simples u_n que aproximan a u . Es decir

$$\int_E u d\mu = \lim_{n \rightarrow \infty} \int_E u_n d\mu.$$

Puede mostrarse que la integral está bien definida y que es independiente de la sucesión aproximante $\{u_n\}$ elegida, y que si u es fuertemente medible, entonces u es integrable Bochner si y sólo si $\|u(\cdot)\|$ es integrable como función a valores reales.

Los resultados de linealidad conocidos para la integral de Lebesgue permanecen válidos para la integral de Bochner. También podemos definir al espacio $L^p(\Omega : Z)$ como el conjunto de funciones de Ω en Z fuertemente medibles que satisfacen $\int_{\Omega} \|u\|^p d\mu < \infty$, que dotado de la norma $\|u\|_{L^p} \doteq (\int_{\Omega} \|u\|^p d\mu)^{1/p}$ es un espacio de Banach.

No existe para las integrales de Bochner un resultado como el *Teorema de la Convergencia Monótona* dado que no siempre puede definirse un orden en un espacio de Banach. Aunque sí existe un resultado similar al *Teorema de la Convergencia Dominada* y dice así:

Teorema 1.8 ([21]) *Si $\{u_n\}$ es una sucesión de funciones integrables Bochner tales que $\|u_n(\omega)\| \leq g(\omega)$ para todo $n \in \mathbb{N}$, donde g es una función a valores reales integrable en $E \in \mathcal{P}$, y $u_n(\omega) \rightarrow u(\omega)$ para casi todo $\omega \in \Omega$, entonces u es integrable Bochner y*

$$\int_E u d\mu = \lim_{n \rightarrow \infty} \int_E u_n d\mu.$$

En virtud de los teoremas 1.4 y 1.5 resulta claro que si A es el generador infinitesimal de un C_0 -semigrupo $T(t)$ y $z_0 \in D(A)$, entonces el problema (1.1) tiene solución única y ésta viene dada por $z(t) = T(t)z_0$. En la siguiente sección veremos cómo podemos determinar si un operador A es el generador infinitesimal de un C_0 -semigrupo.

1.3 Caracterización de Generadores Infinitesimales de C_0 -semigrupos

Dado un operador cerrado A con dominio denso en un espacio de Banach Z (condiciones necesarias para que genere un C_0 -semigrupo), nos preguntamos: ¿qué condiciones **suficientes** hay que impliquen que A genera un C_0 -semigrupo? La primera respuesta a este interrogante para semigrupos contractivos la da el teorema de Hille–Yosida.

Teorema 1.9 (Hille–Yosida) *Un operador lineal A es el generador infinitesimal de un C_0 -semigrupo de contracciones $T(t)$ ($\|T(t)\| \leq 1, t \geq 0$), si y sólo si se cumplen las dos condiciones siguientes:*

- (i) *A es cerrado y $D(A)$ es denso en Z .*
- (ii) *El conjunto resolvente de A , $\rho(A)$ contiene al intervalo real $(0, \infty)$ y para todo $\lambda \in (0, \infty)$,*

$$\|(\lambda I - A)^{-1}\| \leq \frac{1}{\lambda}.$$

El operador $(\lambda I - A)^{-1}$ se denomina resolvente de A en λ y se denota usualmente por $R(\lambda; A)$.

Por otro lado, resulta que si $\|T(t)\| \leq Me^{\omega t}$ para todo $t \geq 0$ y si $\lambda > \omega$, entonces $e^{-\lambda t}T(t)z$ es integrable en el sentido de Bochner para todo $z \in Z$, y más aún, se tiene que $\lambda \in \rho(A)$ y

$$R(\lambda; A)z = \int_0^{\infty} e^{-\lambda t}T(t)z dt,$$

es decir, la resolvente de A en λ es la transformada unilateral de Laplace del semigrupo generado por A . De este comentario y el teorema anterior se obtiene el siguiente

Corolario 1.10 *Sea A el generador infinitesimal de un C_0 -semigrupo de contracciones $T(t)$. Entonces el conjunto resolvente de A contiene al semiplano derecho del plano complejo, es decir, $\rho(A) \supset \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda > 0\}$ y para tales λ*

$$\|R(\lambda; A)\| \leq \frac{1}{\operatorname{Re} \lambda}.$$

El siguiente corolario nos da una caracterización de los generadores infinitesimales de semigrupos que pueden no estar acotados por 1, pero que están acotados por $e^{\omega t}$ ($M = 1$) para algún $\omega \geq 0$.

Corolario 1.11 *Un operador lineal A es el generador infinitesimal de un C_0 -semigrupo que satisface $\|T(t)\| \leq e^{\omega t}$ si y sólo si se cumplen las dos condiciones siguientes:*

- (i) A es cerrado y $D(A)$ es denso en Z .
- (ii) El conjunto resolvente $\rho(A)$ de A contiene a la semirrecta $\{\lambda : \text{Im } \lambda = 0, \text{Re } \lambda > \omega\}$ y para tales λ

$$\|R(\lambda; A)\| \leq \frac{1}{\lambda - \omega}.$$

Hemos visto hasta aquí una caracterización de los generadores infinitesimales de semigrupos en términos de la norma de la resolvente de A . A continuación veremos otra caracterización que tiene en cuenta propiedades como la disipatividad del operador A . Antes necesitaremos las siguientes dos definiciones:

Definición 1.12 Sea Z un espacio de Banach y sea Z^* su dual, para $z \in Z$, definimos el conjunto de dualidad $F(z)$ de z por

$$F(z) \doteq \{z^* \in Z^* : \langle z^*, z \rangle = \|z\|^2 = \|z^*\|^2\},$$

donde $\langle z^*, z \rangle$ denota el valor del funcional $z^* \in Z^*$ en $z \in Z$ (por el teorema de Hahn-Banach $F(z)$ es no vacío).

Definición 1.13 Un operador lineal A es disipativo si para todo $z \in D(A)$ existe un $z^* \in F(z)$ tal que $\text{Re}\langle Az, z^* \rangle \leq 0$.

Notemos que si Z es un espacio de Hilbert, A es disipativo si y sólo si $\text{Re}\langle Az, z \rangle \leq 0$ para todo $z \in D(A)$, pues en este caso $F(z) = \{z\}$, $\forall z \in Z$. El siguiente resultado da una caracterización de los operadores disipativos en espacios de Banach.

Teorema 1.14 Un operador lineal A sobre un espacio de Banach Z es disipativo si y sólo si $\|(\lambda I - A)z\| \geq \lambda\|z\|$ para todo $z \in D(A)$ y para todo $\lambda > 0$.

El siguiente teorema, conocido como teorema de Lumer-Phillips provee una condición necesaria y una suficiente en términos de la disipatividad de un operador lineal para que

el mismo genere un semigrupo de contracciones.

Teorema 1.15 (Lumer–Phillips) *Sea A un operador lineal con dominio denso en un espacio de Banach Z .*

- (a) *Si A es disipativo y existe un $\lambda_0 > 0$ tal que el rango $R(\lambda_0 I - A)$ de $(\lambda_0 I - A)$ es Z , es decir, si $(\lambda_0 I - A)$ es sobreyectivo, entonces A es el generador infinitesimal de un C_0 -semigrupo de contracciones en Z .*
- (b) *Recíprocamente, si A es el generador infinitesimal de un C_0 -semigrupo de contracciones en Z entonces $R(\lambda I - A) = Z$ para todo $\lambda > 0$ y A es disipativo. Más aún, para todo $z \in D(A)$ y todo $z^* \in F(z)$, $\operatorname{Re}\langle Az, z^* \rangle \leq 0$.*

Corolario 1.16 *Sea A un operador lineal con dominio denso en un espacio de Banach Z . Si A y A^* son ambos disipativos, entonces A es el generador infinitesimal de un C_0 -semigrupo de contracciones en Z .*

El siguiente teorema es una generalización del teorema de Hille–Yosida (Teorema 1.9) y da una caracterización del generador infinitesimal de un C_0 -semigrupo en general.

Teorema 1.17 *Un operador lineal A es el generador infinitesimal de un C_0 -semigrupo $T(t)$ que satisface $\|T(t)\| \leq Me^{\omega t}$, si y sólo si se cumplen las dos condiciones siguientes:*

- (i) *A es cerrado y $D(A)$ es denso en Z .*
- (ii) *El conjunto resolvente $\rho(A)$ de A contiene al intervalo real (ω, ∞) y*

$$\|R(\lambda; A)^n\| \leq \frac{M}{(\lambda - \omega)^n} \quad \text{para } \lambda > \omega, \quad n = 1, 2, \dots$$

1.4 Semigrupos Diferenciables y Analíticos

Hemos visto hasta ahora (Teorema 1.4) que si A genera un C_0 -semigrupo en un espacio de Banach Z , entonces el problema de Cauchy abstracto (1.1) tiene una única solución

para todo $z_0 \in D(A)$, y hemos visto también condiciones necesarias y suficientes para que un operador A genere un C_0 -semigrupo. Sin embargo, en el ejemplo sencillo de la ecuación del calor, sabemos que el problema tiene solución para **todo** $z_0 \in Z$ y no sólo para $z \in D(A)$. Nos preguntamos entonces: ¿Qué condiciones sobre el operador A aseguran que el problema tenga solución para todo $z_0 \in Z$, o al menos en un conjunto más grande que $D(A)$? Para contestar a esta pregunta necesitaremos previamente algunas definiciones y resultados sobre semigrupos diferenciables.

Definición 1.18 *Se dice que un semigrupo $T(t)$ en un espacio de Banach Z es diferenciable para $t > t_0$ si para todo $z \in Z$, el mapeo $t \rightarrow T(t)z$ es diferenciable en el intervalo (t_0, ∞) . El semigrupo es diferenciable si es diferenciable para $t > 0$.*

Según el Teorema 1.4, si $T(t)$ es un C_0 -semigrupo con generador infinitesimal A y $z \in D(A)$, entonces el mapeo $t \rightarrow T(t)z$ es diferenciable para todo $t \geq 0$. Si $T(t)$ fuera más aún diferenciable, entonces para **todo** $z \in Z$ y no sólo para $z \in D(A)$, el mapeo $t \rightarrow T(t)z$ sería diferenciable para $t > 0$, aunque no para $t = 0$. Notemos por otro lado que si $t \rightarrow T(t)z$ es diferenciable para todo $z \in Z$ y $t \geq 0$ (en $t = 0$ consideramos la derivada lateral derecha), entonces $D(A) = X$ (por definición de $D(A)$) y como A es cerrado, sería necesariamente acotado en virtud del teorema del gráfico cerrado. Por lo que acabamos de mencionar, vemos que no tiene sentido definir como semigrupos diferenciables a aquellos que lo son para $t \geq 0$, puesto que estaríamos volviendo al caso de operadores acotados.

El siguiente lema junto con su corolario constituyen un resultado importante sobre los semigrupos diferenciables, conocido como *propiedad suavizante de los semigrupos diferenciables* (smoothing property of differentiable semigroups).

Lema 1.19 *Sea $T(t)$ un C_0 -semigrupo diferenciable para $t > t_0$ y sea A su generador infinitesimal, entonces para todo $n \in \mathbb{N}$ y $t > nt_0$, el operador $T(t)$ manda Z a $D(A^n)$. En particular, si $t > t_0$, $T(t)z \in D(A)$ para todo $z \in Z$.*

Corolario 1.20 *Si $T(t)$ es un C_0 -semigrupo diferenciable, entonces $T(t)z \in D(A^n)$ para todo $t > 0$, $z \in Z$ y $n \in \mathbb{N}$.*

Como consecuencia de estos dos resultados, vemos que si $T(t)$ es un semigrupo diferenciable, entonces el problema de Cauchy (1.1) tiene solución para todo $z \in Z$.

A continuación daremos una caracterización de los generadores infinitesimales de C_0 -semigrupos diferenciables para $t > t_0$.

Teorema 1.21 *Sea $T(t)$ un C_0 -semigrupo con generador infinitesimal A . Si $\|T(t)\| \leq Me^{\omega t}$, entonces las dos afirmaciones siguientes son equivalentes:*

- (i) *Existe $t_0 > 0$ tal que $T(t)$ es diferenciable para $t > t_0$.*
- (ii) *Existen constantes reales $a, b > 0$ y $C > 0$ tales que (ver Figura 1.1a)*

$$\rho(A) \supset \Sigma \doteq \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda \geq a - b \log |\operatorname{Im} \lambda|\} \quad (1.3)$$

y

$$\|R(\lambda; A)\| \leq C |\operatorname{Im} \lambda| \quad \text{para } \lambda \in \Sigma, \quad \operatorname{Re} \lambda \leq \omega. \quad (1.4)$$

De la demostración de este teorema [42, Teorema 2.4.7] se sigue la siguiente caracterización del generador infinitesimal de un semigrupo diferenciable ($t_0 = 0$).

Teorema 1.22 *Sea $T(t)$ un C_0 -semigrupo que satisface $\|T(t)\| \leq Me^{\omega t}$ y cuyo generador infinitesimal es A . Entonces el semigrupo $T(t)$ es diferenciable si y sólo si para todo $b > 0$ existen constantes reales a_b y C_b con $C_b > 0$ tales que*

$$\rho(A) \supset \Sigma_b = \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda > a_b - b \log |\operatorname{Im} \lambda|\}$$

y

$$\|R(\lambda; A)\| \leq C_b |\operatorname{Im} \lambda| \quad \lambda \in \Sigma_b, \quad \operatorname{Re} \lambda \leq \omega.$$

A continuación introduciremos el concepto de *Semigrupo Analítico* que será de gran utilidad a la hora de definir potencias fraccionarias de operadores cerrados y para trabajar con problemas de Cauchy abstractos no lineales.

Definición 1.23 Sea $\Delta = \{\gamma \in \mathbb{C} : \phi_1 < \arg \gamma < \phi_2, \phi_1 < 0 < \phi_2\}$ y para $\gamma \in \Delta$, sea $T(\gamma)$ un operador lineal y acotado en el espacio de Banach Z . Decimos que la familia $T(\gamma)$, $\gamma \in \Delta$, es un semigrupo analítico en Δ si se cumplen las siguientes condiciones:

- (i) El mapeo $\gamma \rightarrow T(\gamma)$ es analítico en Δ .
- (ii) $T(0) = I$ y $\lim_{\substack{t \rightarrow 0 \\ t \in \mathbb{R}^+}} T(t)z = z$ para todo $z \in Z$.
- (iii) $T(\gamma_1 + \gamma_2) = T(\gamma_1)T(\gamma_2)$ para $\gamma_1, \gamma_2 \in \Delta$.

Claramente la restricción de un semigrupo analítico al eje real es un C_0 -semigrupo. En general, diremos que un semigrupo $T(t)$, $t \geq 0$ es un semigrupo analítico si es la restricción al eje real de un semigrupo analítico definido en un sector angular Δ , o equivalentemente, si puede extenderse en forma analítica a un sector angular Δ en la forma mencionada en la Definición 1.23. El siguiente teorema da una caracterización del generador infinitesimal de un semigrupo analítico en términos del conjunto y el operador resolvente.

Teorema 1.24 Sea $T(t)$ un C_0 -semigrupo con generador infinitesimal A que satisface $\|T(t)\| \leq Me^{\omega t}$ y tal que $\omega \in \rho(A)$ ($0 \in \rho(A - \omega I)$). Las siguientes afirmaciones son equivalentes:

- (a) $T(t)$ puede extenderse a un semigrupo analítico en un sector $\Delta_\delta = \{\gamma \in \mathbb{C} : |\arg \gamma| < \delta\}$ para algún $\delta > 0$ y para cada $0 < \delta' < \delta$ existe una constante $M_{\delta'} \geq 1$ tal que $\|T(\gamma)\| \leq M_{\delta'} e^{\omega \operatorname{Re} \gamma}$ en el subsector cerrado $\overline{\Delta_{\delta'}}$ de Δ_δ .
- (b) Existe una constante C tal que para todo $\sigma > 0$, $\tau \neq 0$

$$\|R((\omega + \sigma) + i\tau; A)\| \leq \frac{C}{|\tau|}$$

(c) Existen $0 < \delta < \pi/2$ y $M > 0$ tales que (ver Figura 1.1b)

$$\rho(A) \supset \Sigma = \left\{ \lambda \in \mathbb{C} : |\arg(\lambda - \omega)| < \frac{\pi}{2} + \delta \right\} \cup \{\omega\}$$

y

$$\|R(\lambda; A)\| \leq \frac{M}{|\lambda - \omega|} \quad \text{para todo } \lambda \in \Sigma \setminus \{\omega\}.$$

(d) $T(t)$ es diferenciable para $t > 0$ y existe una constante C tal que

$$\|AT(t)\| \leq \frac{C}{t} e^{\omega t} \quad \text{para todo } t > 0.$$

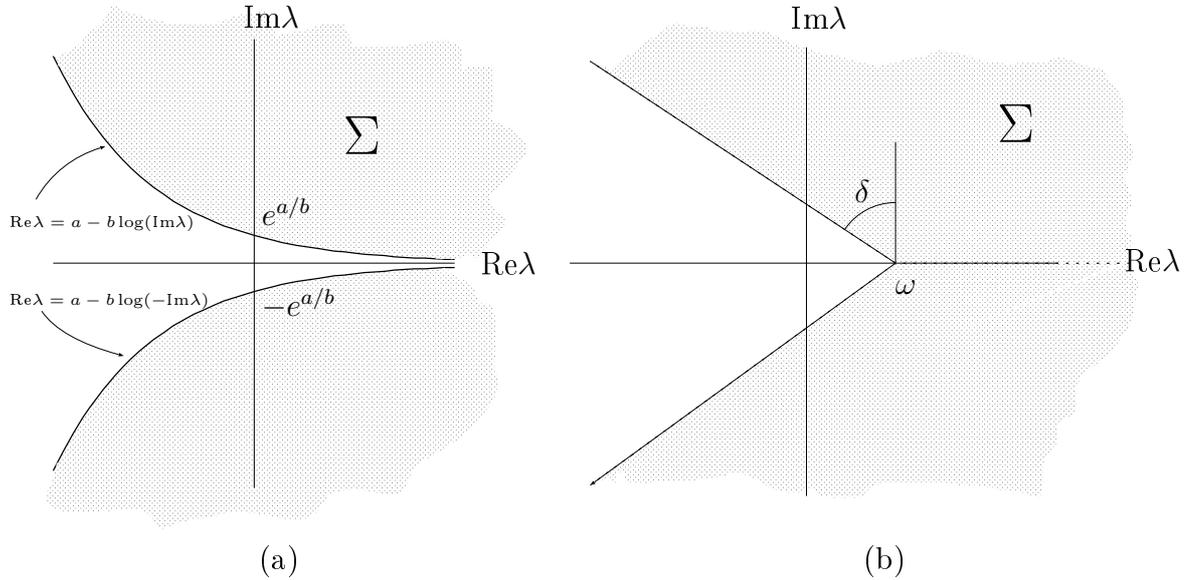


Figura 1.1: Forma del conjunto $\Sigma \subset \mathbb{C}$ para los semigrupos Diferenciables (a) y Analíticos (b).

Existen otras caracterizaciones de generadores infinitesimales de semigrupos analíticos, pero no las enunciamos aquí dado que no son necesarias para el desarrollo del presente trabajo y no es nuestro objetivo hacer un estudio exhaustivo de la teoría de semigrupos, sino solamente exponer algunos resultados básicos de la misma.

Podemos concluir entonces que si A es el generador infinitesimal de un semigrupo diferenciable o analítico, el problema de valores iniciales (1.1) tiene una solución única

para todo $z_0 \in Z$. Si A es el generador infinitesimal de un C_0 -semigrupo que no es diferenciable, entonces, en general, si $z_0 \notin D(A)$, el problema de valores iniciales (1.1) no tiene solución. Sin embargo, la función $t \rightarrow T(t)z_0$ está bien definida aún si $z_0 \notin D(A)$ y es una “solución generalizada” a la que llamaremos *solución débil* de (1.1).

1.5 Potencias Fraccionarias de Operadores Cerrados

En esta sección daremos una definición de potencia α de un operador cerrado, para valores reales de α . Esta definición nos conducirá a la definición de los dominios de las potencias del operador, que resultarán ser espacios que se encuentran “entre” el dominio del operador original y el espacio de Banach Z sobre el que está definido. Estos espacios aparecerán de forma natural más adelante al estudiar ecuaciones de evolución no lineales.

Supondremos a lo largo de esta sección que el operador A es el generador infinitesimal de un semigrupo analítico. Por el teorema 1.24 y por el hecho que la resolvente de un operador cerrado es un conjunto abierto, existen un número real $\omega \geq 0$, un entorno V del origen en \mathbb{C} y una constante $\delta \in (0, \pi/2)$ tales que

$$\rho(\omega I - A) \supset \Sigma^+ \doteq \{\lambda \in \mathbb{C} : 0 < \delta < |\arg \lambda| \leq \pi\} \cup V$$

y para todo $\lambda \in \Sigma^+ \setminus \{0\}$, $\|R(\lambda; A - \omega I)\| \leq \frac{M}{|\lambda|}$. Para simplificar la notación llamaremos A_ω al operador $A - \omega I$ que satisface $0 \in \rho(A_\omega)$ y que genera el semigrupo acotado $T(t)e^{-\omega t}$.

Definimos, para $\alpha > 0$,

$$(-A_\omega)^{-\alpha} = \int_C \gamma^{-\alpha} (-\gamma I - A_\omega)^{-1} d\gamma \quad (1.5)$$

donde C es un camino que se encuentra en Σ^+ de $\infty e^{-i\theta}$ a $\infty e^{i\theta}$, con $\delta < \theta < \pi$, evitando el eje real negativo y el origen (ver Figura 1.2). (Nota: tomamos $\gamma^{-\alpha} > 0$ para valores positivos de γ). Puede probarse que la integral (1.5) converge en la topología uniforme de operadores y define un operador lineal acotado $(-A_\omega)^{-\alpha} \in \mathcal{L}(Z)$. Si $\alpha = n \in \mathbb{N}_0$ el integrando es analítico en Σ^+ y es fácil verificar que el camino de integración puede

transformarse a un pequeño círculo alrededor del origen. Luego, utilizando el teorema de los residuos se sigue que la integral es $(-A_\omega)^{-n}$ y para valores de α enteros no negativos, la definición de $(-A_\omega)^{-\alpha}$ coincide con la definición clásica de $((-A_\omega)^{-1})^\alpha$, en particular $(-A_\omega)^0 = I$.

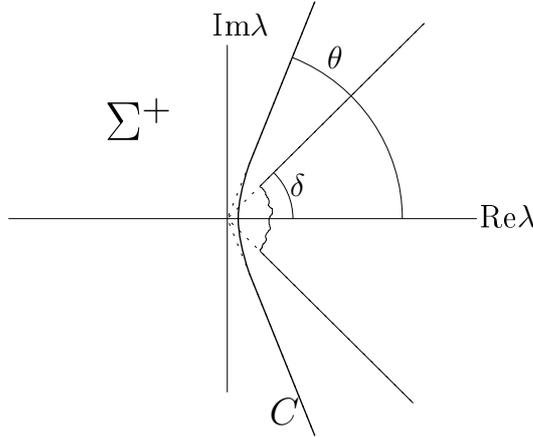


Figura 1.2: Región Σ^+ y forma que debe tener la curva C para definir las potencias fraccionarias de un operador cerrado

Para $0 < \alpha < 1$ el camino puede deformarse al lado superior y al lado inferior del semieje real negativo y se obtiene de este modo

$$(-A_\omega)^{-\alpha} = \frac{\text{sen } \pi\alpha}{\pi} \int_0^\infty t^{-\alpha} (tI - A_\omega)^{-1} dt \quad 0 < \alpha < 1. \quad (1.6)$$

De esta última expresión y utilizando el hecho que $((t + \omega)I - A)^{-1} = \int_0^\infty e^{-s(t+\omega)} T(s) ds$ para todo $t > 0$, puede deducirse la fórmula

$$(-A_\omega)^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-\omega t} T(t) dt \quad (1.7)$$

donde $\Gamma(\alpha)$ es la función Gamma, $\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1} dx$.

Utilizando esta última expresión, que nos da una fórmula para definir $(-A_\omega)^{-\alpha}$ para $0 < \alpha < 1$, puede demostrarse el siguiente lema, el cual lista propiedades de $(-A_\omega)^{-\alpha}$ que eran de esperarse de las potencias de un operador.

Lema 1.25 (i) Si $\alpha \geq 0$ y $\beta \geq 0$, entonces $(-A_\omega)^{-(\alpha+\beta)} = (-A_\omega)^{-\alpha} (-A_\omega)^{-\beta}$.

- (ii) Existe una constante C tal que $\|(-A_\omega)^{-\alpha}\| \leq C$ para $0 \leq \alpha \leq 1$.
- (iii) Para todo $z \in Z$ se tiene que $\lim_{\alpha \rightarrow 0} (-A_\omega)^{-\alpha} z = z$.
- (iv) La familia de operadores $S(t) = (-A_\omega)^{-t}$ es un C_0 -semigrupo de operadores lineales y acotados sobre Z .
- (v) $(-A_\omega)^{-\alpha}$ es uno a uno para todo $\alpha > 0$.

En virtud de este lema, podemos establecer la siguiente

Definición 1.26 Para todo $\alpha > 0$ se define la potencia α de $(-A_\omega)$ por

$$D(-A_\omega)^\alpha = R(-A_\omega)^{-\alpha}, \quad (-A_\omega)^\alpha = ((-A_\omega)^{-\alpha})^{-1}.$$

Teorema 1.27 Sea $(-A_\omega)^\alpha$ según la Definición 1.26, entonces

- (a) El operador $(-A_\omega)^\alpha$ es cerrado en el dominio $D(-A_\omega)^\alpha = R(-A_\omega)^{-\alpha}$, el rango de $(-A_\omega)^{-\alpha}$.
- (b) Si $\alpha \geq \beta \geq 0$, entonces $D(-A_\omega)^\alpha \subset D(-A_\omega)^\beta$, donde la inclusión es continua y densa.
- (c) $D(-A_\omega)^\alpha$ es denso en Z para todo $\alpha \geq 0$.
- (d) Si α y β son reales, entonces

$$(-A_\omega)^{\alpha+\beta} z = (-A_\omega)^\alpha (-A_\omega)^\beta z$$

para todo $z \in D(-A_\omega)^\gamma$, donde $\gamma = \max(\alpha, \beta, \alpha + \beta)$.

A continuación se da una expresión explícita para $(-A_\omega)^\alpha$, si $\alpha \in (0, 1)$.

Teorema 1.28 Sea $0 < \alpha < 1$. Si $z \in D(A)$ entonces

$$(-A_\omega)^\alpha z = -\frac{\text{sen } \pi\alpha}{\pi} \int_0^\infty t^{\alpha-1} A_\omega (tI - A_\omega)^{-1} z dt.$$

Por último, mencionamos a continuación algunos resultados que relacionan $(-A_\omega)^\alpha$ con el semigrupo $T(t)$ generado por A .

Teorema 1.29 *Sea A el generador infinitesimal de un semigrupo analítico $T(t)$, y sea $\omega \in \rho(A) \cap \mathbb{R}$, entonces*

(a) $T(t) : Z \rightarrow D(-A_\omega)^\alpha$ para todo $\alpha \geq 0$ y $t > 0$ (notar que no vale para $t = 0$).

(b) Para todo $\alpha \geq 0$, $t \geq 0$ y $z \in D(-A_\omega)^\alpha$ se tiene que $T(t)(-A_\omega)^\alpha z = (-A_\omega)^\alpha T(t)z$.

(c) Para todo $\alpha \geq 0$ y $t > 0$, el operador $(-A_\omega)^\alpha T(t)$ es un operador acotado de Z en Z . Más aún, para todo $\alpha \geq 0$ existe una constante $M_\alpha \geq 1$ tal que

$$\|(-A_\omega)^\alpha T(t)\|_{\mathcal{L}(Z)} \leq M_\alpha e^{\omega t} t^{-\alpha}.$$

(d) Para todo $0 < \alpha \leq 1$ existe una constante $C_\alpha > 0$ tal que para todo $z \in D(-A_\omega)^\alpha$, se tiene que

$$\|T(t)z - z\| \leq C_\alpha t^\alpha \|(-A_\omega)^\alpha z\|.$$

1.6 Una Clase de Ecuaciones de Evolución No lineales

En esta sección consideramos el problema de valores iniciales

$$\begin{cases} \frac{dz(t)}{dt} = Az(t) + f(t, z(t)), & t \in (0, T) \\ z(0) = z_0, \end{cases} \quad (1.8)$$

donde A es el generador infinitesimal de un C_0 -semigrupo $T(t)$ en un espacio de Banach Z , y $f(t, z)$ es una función de $[0, T] \times D(F)$ en Z para algún conjunto $D(F) \subset Z$.

Al final de la Sección 1.4 vimos la noción de solución débil para el problema de valores iniciales lineal homogéneo. Análogamente, puede definirse, para el problema de valores iniciales más general (1.8) el concepto de solución débil.

Definición 1.30 *Llamaremos solución débil de la ecuación (1.8) a toda función $z(t)$ continua en $[0, T)$ que satisfaga la ecuación integral*

$$z(t) = T(t)z_0 + \int_0^t T(t-s)f(s, z(s)) ds, \quad 0 \leq t < T. \quad (1.9)$$

Es claro que si $f(t, z) = h(t)$, es decir, si $f(t, z)$ es independiente de z , el problema de valores iniciales (1.8) tiene solución débil en $[0, T)$ siempre que $h \in L^1(0, T : Z)$. En este caso, la ecuación no es no lineal, sino que es lineal no homogénea. Para estudiar la existencia de soluciones *clásicas* de la ecuación lineal no homogénea basta con estudiar la regularidad de $z(t)$ dada por

$$z(t) = T(t)z_0 + \int_0^t T(t-s)h(s) ds,$$

dado que si ésta es diferenciable en $(0, T)$ y continua en $[0, T)$, entonces $z'(t) = Az(t) + h(t)$ en $(0, T)$ y $z(t)$ es solución clásica. Resultados en esta línea pueden encontrarse en [42, Capítulos 4 a 6].

En general, el problema de la existencia de soluciones del PVI (1.8) se reduce al problema de existencia y *regularidad* de soluciones de la ecuación integral (1.9).

Cuando la función $f(t, z)$ *depende* de la variable z puede suceder que ni siquiera la ecuación integral (1.9) tenga solución, lo que implicaría que la ecuación (1.8) no tiene siquiera solución débil y, mucho menos, solución clásica. Esta característica hace más interesante el problema matemático de la existencia de soluciones en este caso. El primer resultado relacionado con la existencia de soluciones del problema (1.8) es el teorema que sigue, cuya demostración puede encontrarse en [42, Teorema 6.1.7].

Teorema 1.31 *Sea $f : [0, T] \times Z \rightarrow Z$ uniformemente Lipschitz en Z y tal que para cada $z \in Z$, $f(\cdot, z)$ es una función continua de $[0, T] \rightarrow Z$. Entonces, si $z_0 \in D(A)$, el problema de valores iniciales (1.8) tiene una única solución clásica en $[0, T)$.*

El siguiente resultado [42, Teorema 6.3.1] sobre la existencia de soluciones para el

PVI (1.8) requiere que A genere un semigrupo analítico, pero permite que la parte no lineal $f(t, z)$ sea menos regular, a la vez que asegura la existencia de soluciones para el dato inicial z_0 en un conjunto más grande que $D(A)$. Sabemos, por la Sección 1.5, que si A genera un semigrupo analítico, entonces existe un $\omega \in \mathbb{R}_0^+$ tal que pueden definirse las potencias α de $(\omega I - A) = -A_\omega$, para cualquier $\alpha \in \mathbb{R}$. Además, si $\alpha > 0$, el operador $(-A_\omega)^\alpha$ es cerrado y luego el espacio $D(-A_\omega)^\alpha$ es de Banach con la norma del gráfico de $(-A_\omega)^\alpha$. Llamaremos Z_α al espacio $D(-A_\omega)^\alpha$ con la norma $\|z\|_\alpha \doteq \|(-A_\omega)^\alpha z\|_Z$ que es equivalente a la norma del gráfico de $(-A_\omega)^\alpha$ pues $(-A_\omega)^\alpha$ es inversible con inversa acotada.

Teorema 1.32 *Sea A el generador infinitesimal de un semigrupo analítico $T(t)$ que satisface $\|T(t)\| \leq M$. Supongamos que U es un conjunto abierto de $\mathbb{R}_0^+ \times Z_\alpha$ y que la función $f : U \rightarrow Z$ satisface que para todo par $(t, z) \in U$ existe un entorno $V \subset U$ y constantes $L \geq 0$, $0 < \theta \leq 1$ tales que*

$$\|f(t_1, z_1) - f(t_2, z_2)\| \leq L(|t_1 - t_2|^\theta + \|z_1 - z_2\|_\alpha) \quad (1.10)$$

para todo $(t_i, z_i) \in V$. Entonces para todo dato inicial $(0, z_0) \in U$, el problema de valores iniciales (1.8) tiene una solución local única $z \in C([0, t_1] : Z) \cap C^1((0, t_1) : Z)$ donde $t_1 = t_1(z_0) > 0$.

Observando la demostración de este resultado [42, Teorema 6.3.1] podemos concluir que la solución no sólo es continua en la norma del espacio Z , sino que, más aún, es continua en la norma más fuerte de Z_α , para α satisfaciendo (1.10).

Cuando se estudian ecuaciones de evolución en espacios de estado de dimensión finita juega un papel esencial la llamada *Desigualdad de Gronwall*. Al considerar espacios de dimensión infinita, y específicamente ecuaciones de evolución como las presentadas en esta sección, al mismo rol lo cumple una generalización de la Desigualdad de Gronwall para integrandos singulares. Ésta se presenta en el siguiente lema cuya demostración puede encontrarse en [29, Teorema 7.1.1].

Lema 1.33 Sean $T > 0$, $a(t)$ una función no negativa, localmente integrable en $[0, T]$, $L \geq 0$ y $0 < \delta < 1$. Entonces existe una constante $K = K(\delta, L, T)$ independiente de la función $a(t)$, tal que para toda función $u(t)$ que satisfice

$$u(t) \leq a(t) + L \int_0^t \frac{1}{(t-s)^\delta} u(s) ds \quad \text{en } [0, T],$$

se tiene que $u(t)$ también verifica

$$u(t) \leq a(t) + K \int_0^t \frac{a(s)}{(t-s)^\delta} ds, \quad \text{para todo } t \in [0, T].$$

1.7 Identificación de Parámetros

En esta sección introduciremos el problema de identificación de parámetros cuya resolución atacaremos en los capítulos siguientes.

Sean Z y $\tilde{\mathcal{Q}}$ dos espacios de Banach, A el generador infinitesimal de un semigrupo analítico $T(t)$ en Z , D un subconjunto de Z , \mathcal{Q} un subconjunto de $\tilde{\mathcal{Q}}$ y $F : \mathcal{Q} \times [0, T] \times D \rightarrow Z$. Para cada $q \in \mathcal{Q}$ y $z_0 \in Z$, consideraremos el siguiente problema de Cauchy no lineal en Z

$$\begin{cases} \frac{dz(t)}{dt} = Az(t) + F(q, t, z(t)), & t \in (0, T) \\ z(0) = z_0. \end{cases} \quad (1.11)$$

Llamaremos espacio de parámetros a $\tilde{\mathcal{Q}}$, mientras que a \mathcal{Q} lo llamaremos *conjunto de parámetros admisibles*. En síntesis, consideraremos un PVI del tipo (1.8) pero la parte no lineal depende de un parámetro q que toma valores en un subconjunto de un espacio de Banach.

Supongamos por un momento que el problema (1.11) está bien planteado en el sentido de que para cada $q \in \mathcal{Q}$, existe una única solución en $[0, T]$ a la que llamaremos $z(t; q)$. Supongamos también que poseemos observaciones \hat{z}_i del sistema físico descrito por (1.11) en tiempos discretos t_i , $i = 1, 2, \dots, m$ y deseamos determinar, a partir de ellas, el parámetro $q \in \mathcal{Q}$ tal que la parte observada $\mathcal{C}z(t_i; q)$ del estado $z(t_i; q)$ (dependiente del parámetro q) aproxime lo mejor posible, en algún sentido, a las observaciones $\{\hat{z}_i\}$. Este

tipo de problema recibe el nombre de *problema de identificación de parámetros* dado que uno quiere estimar el valor del parámetro q del sistema físico descrito por (1.11) para el cual el estado coincide con las observaciones.

De manera más precisa, definimos el *problema de identificación de parámetros* (ID) como sigue:

(ID): encontrar $q \in \mathcal{Q}$ que minimice el criterio de error

$$J(q) = \frac{1}{2} \sum_{i=1}^m \|\mathcal{C}z(t_i; q) - \hat{z}_i\|_Y^2$$

donde \mathcal{C} es un operador lineal y acotado del espacio de estados Z en un espacio de Hilbert real Y . El espacio Y representa el espacio de las observaciones y el operador \mathcal{C} es el operador de observación, es decir, $\mathcal{C}z(t; q)$ es la parte observada del estado en tiempo t . En principio, teniendo en cuenta la descripción del problema, bastaría pedir que Y fuera un espacio métrico para poder definir la distancia de $\mathcal{C}z(t; q)$ a \hat{z}_i , pero por conveniencia y dado que en general $Y = \mathbb{R}^n$, para algún n , pedimos que Y sea un espacio de Hilbert real.

Para el tratamiento de estos problemas de identificación existen dos clases de métodos. La primera clase, frecuentemente utilizada en problemas lineales, la constituyen los *métodos indirectos*. Estos consisten en aproximar primero las ecuaciones del modelo mediante la utilización de elementos finitos, diferencias finitas, aproximaciones espectrales, etc., y luego utilizar algoritmos de optimización sobre el problema en dimensión finita resultante.

La segunda clase está compuesta por los llamados *métodos directos*, que se basan en la aplicación **directa** de un algoritmo de optimización (quizás infinito-dimensional) y en la aplicación de aproximaciones numéricas en cada paso del algoritmo para aproximar las soluciones del sistema que sean necesarias.

Ambos métodos tienen ventajas y desventajas, y dependiendo del problema que se esté considerando, un método de una clase puede resultar mejor que otro de la otra clase.

En general, los métodos indirectos son más sencillos de implementar computacional-

mente pero, necesitan calcular una mayor cantidad de veces (que los métodos directos) las soluciones de las ecuaciones dinámicas involucradas. Por esta razón, en la práctica, su aplicación se restringe mayormente a problemas lineales, o a aquellos en los que los métodos de aproximación resultan computacionalmente muy baratos.

Por otro lado, para los métodos indirectos, no puede demostrarse más que convergencia subsecuencial, mientras que para ciertos métodos directos se puede probar convergencia “total”, y en algunos casos, hasta puede establecerse un orden de convergencia.

1.8 Un Problema de Control

En esta sección introduciremos un problema de control asociado a una ecuación de evolución no lineal en un espacio de Banach. En el capítulo 7 abordaremos la resolución numérica del problema que planteamos en esta sección, para el sistema particular que resulta de considerar las ecuaciones que modelan el comportamiento termomecánico de un sólido unidimensional con memoria de forma (ver Capítulo 4). A pesar de que no haremos un estudio exhaustivo del problema de control que aquí planteamos, consideramos conveniente incluirlo, dado que los problemas de control surgen frecuentemente en la práctica y también forman parte de la gran familia de problemas inversos asociados a ecuaciones diferenciales.

Es nuestro interés estudiar no sólo el problema de control que veremos en la presente sección, sino también otros problemas de control asociados a este tipo de ecuaciones. Estos otros problemas serán motivo de trabajos posteriores a la finalización de esta tesis.

Al igual que en la sección anterior, sea Z un espacio de Banach, A el generador infinitesimal de un semigrupo analítico $T(t)$ en Z , D un subconjunto de Z , B un operador lineal y acotado del espacio de Banach U en Z , $F : [0, T] \times D \rightarrow Z$. Para cada función (control) $u : [0, T] \rightarrow U$ y $z_0 \in Z$, consideraremos el siguiente problema de Cauchy

nolineal en Z

$$\begin{cases} \frac{dz(t)}{dt} = Az(t) + F(t, z(t)) + Bu(t), & t \in (0, T) \\ z(0) = z_0. \end{cases} \quad (1.12)$$

Supongamos ahora que el problema (1.12) tiene una única solución $z_u(t)$, $t \in [0, T)$ para cada $u(t)$ en algún conjunto \mathcal{U} de controles admisibles. Supongamos también que tenemos un objetivo \hat{z} a lograr en un tiempo dado $t = t_1$, $0 < t_1 < T$. En otras palabras, supongamos que queremos lograr que una parte observada $\mathcal{C}z_u(t_1)$ del estado z dependiente de u , en el tiempo $t = t_1$, sea igual al objetivo \hat{z} . Este problema se denomina *Problema de Control Terminal* dado que uno quiere determinar el control $u(t)$ que logre que el estado $z_u(t)$ satisfaga cierta condición solamente en un tiempo final $t = t_1$.

Podría suceder que no exista un control u que logre el objetivo deseado ($\mathcal{C}z_u(t_1) = \hat{z}$), y entonces en ese caso deseamos encontrar un control que logre la mejor aproximación, en algún sentido, de $\mathcal{C}z_u(t_1)$ a \hat{z} . De manera más precisa, planteamos el problema de control terminal del siguiente modo

(C): encontrar $u \in \mathcal{U}$ que minimice el funcional

$$J(u) = \frac{1}{2} \|\mathcal{C}z_u(t_1) - \hat{z}\|_Y^2$$

donde \mathcal{C} es nuevamente un operador lineal y acotado de Z en Y al que llamaremos operador de observación, y Y es un espacio de Hilbert real al que llamaremos espacio de observaciones.

Notemos que surgen inmediatamente muchos interrogantes en torno a este problema, como ejemplos podemos dar los siguientes:

- (a) ¿Existe siempre una función $u \in \mathcal{U}$ que realiza el mínimo de $J(u)$? (*Existencia de solución* del problema recién planteado)
- (b) Si ese mínimo existe, ¿se cumple que $J(u) = 0$? Es decir ¿se logra $\mathcal{C}z_u(t_1) = \hat{z}$? (*Controlabilidad*)

(c) ¿Existe un operador $K : Z \rightarrow U$ tal que el control de retroalimentación del estado $u(t) = Kz(t)$ logre el objetivo? (*Controlabilidad por retroalimentación*)

Estos interrogantes y muchos otros constituyen el núcleo de lo que se conoce como Teoría de Control. En el volumen 2 del libro [10] se hace un tratamiento exhaustivo de estos problemas, entre otros, para el caso en que la ecuación de evolución es lineal ($F(t, z) \equiv 0$). En un artículo reciente, J. Burns y S. Kang [17] estudian el problema de estabilización ($\hat{z} = 0$) por retroalimentación para el caso particular de la ecuación de Burgers, tomándola como ejemplo de un problema no lineal.

Por otro lado, si observamos la definición del problema de control (C), podemos concluir inmediatamente que se trata de un caso particular del problema de identificación (ID). En efecto, si definimos $\mathcal{Q} = \mathcal{U}$, y $\tilde{F}(u, t, z) = F(t, z) + Bu(t)$, el problema (1.12) es un caso particular del problema (1.11) donde además la parte no lineal de la ecuación depende linealmente del parámetro. Y el funcional para el problema (C) es el caso particular en que las observaciones se toman en un solo instante de tiempo $t = t_1$. Nosotros tomaremos ventaja de esta similitud entre los dos problemas. Utilizaremos todo el andamiaje construido para la resolución del problema de identificación (ID), en la resolución numérica del problema de control (C) (ver Capítulo 7).

Capítulo 2

Un Método Directo para la Identificación de Parámetros

En este capítulo propondremos un método para la resolución del problema de identificación de parámetros presentado en la Sección 1.7. La derivación del método se hará de una forma completamente intuitiva y sencilla, y se probarán resultados de convergencia.

El método para la identificación que utilizaremos en este trabajo es un método directo usando cuasilinealización. Básicamente, desarrollaremos el método como sigue: dada una estimación inicial del parámetro que se quiere determinar, reemplazamos la variable de estado por una aproximación lineal de la misma alrededor del parámetro estimado inicialmente, y resolvemos este problema de optimización simplificado, obteniendo una nueva estimación del parámetro. Repitiendo este proceso basado en cada nueva aproximación, determinaremos una sucesión de iteraciones que bajo ciertas condiciones converge a un parámetro óptimo.

El uso de cuasilinealización en problemas de identificación de parámetros ha sido muy exitoso. El concepto fue introducido por primera vez por Bellman y Kalaba [9] en 1965 como una herramienta general para la resolución numérica de problemas de valores en la frontera. En 1969, Buell y Kalaba [15] utilizaron un algoritmo basado en cuasilinealización para ajustar modelos no lineales para el metabolismo de drogas a datos experimentales. En 1973, Banks y Groome [6] desarrollaron un algoritmo usando cuasilinealización y probaron teoremas de convergencia para problemas de identificación de

parámetros asociados a ecuaciones diferenciales no lineales de la forma $\dot{x}(t) = g(t, x(t), \beta)$, pero su trabajo es válido solamente cuando el espacio de estado es de dimensión finita, es decir, $x(t) \in \mathbb{R}^n$, y no puede extenderse inmediatamente al contexto de dimensión infinita. En 1980, Burns y Cliff [16] introdujeron un algoritmo para identificar coeficientes constantes en una ecuación lineal de retardo, aunque no dieron resultados de convergencia. Más recientemente, en 1989, Brewer, Burns y Cliff [12] trabajaron sobre varios aspectos de identificación que aparecen en el estudio y en la aplicación de métodos de cuasilinealización para sistemas lineales no homogéneos del tipo $\frac{d}{dt}z(t) = A(q)z(t) + u(t)$, donde la dependencia del parámetro q aparece en el operador $A(q)$. En este trabajo sí se establecieron teoremas de convergencia, los cuales estuvieron basados en resultados de diferenciabilidad de las soluciones con respecto a parámetros. Estos últimos resultados habían sido obtenidos por Brewer [11] en 1982. Al año siguiente, en 1990, Hammer [28] utilizó estas ideas en problemas no lineales del tipo $\frac{d}{dt}z(t) = A(q)z(t) + f(t, z(t))$ donde $f(t, z)$ es no lineal, pero el parámetro a identificar sólo aparece en el operador $A(q)$ que define la parte lineal de la ecuación.

En los trabajos mencionados arriba, el desarrollo del método para la identificación, como así también la demostración de los resultados de convergencia del algoritmo, se hace trabajando sobre las ecuaciones particulares y utilizando las propiedades de regularidad específicas de las soluciones de las mismas. Además, el algoritmo o método de identificación se presenta en una forma poco intuitiva que no permite percibir rápidamente en qué consiste, en la Sección 2.3 haremos una breve presentación de este enfoque clásico.

En este capítulo estudiaremos el problema de identificación de parámetros (ID) para ecuaciones del tipo (1.11) donde el espacio de estados es de dimensión infinita y el parámetro aparece en la parte no lineal de la ecuación. Esta clase de problemas no ha sido analizada previamente.

La organización de este capítulo es la siguiente. En la Sección 2.1 propondremos un algoritmo basado en cuasilinealización para el problema de identificación de parámetros planteado en la Sección 1.7, y en la Sección 2.2 expondremos resultados de convergencia

para el algoritmo propuesto. En la Sección 2.3 mostraremos la forma en que se presenta en la literatura clásica el algoritmo de cuasilinealización, y veremos qué relación existe con la presentación que llevaremos a cabo en la Sección 2.1.

2.1 Presentación del Algoritmo de Cuasilinealización

En esta sección introduciremos un algoritmo basado en cuasilinealización para la identificación de parámetros en el problema (1.11). La forma en que introduciremos el algoritmo no es la forma clásica en que aparece en la bibliografía mencionada en la introducción de este capítulo, pero es una forma que creemos mucho más intuitiva y no hace uso de las llamadas “ecuaciones de sensibilidad” sobre las cuales se basa la presentación clásica del algoritmo.

Recordemos en primer lugar que, dado que A es el generador infinitesimal de un semigrupo analítico, $\omega \doteq \sup \{\operatorname{Re}(\lambda) : \lambda \in \sigma(A)\} < \infty$ y para todo $\lambda > \omega$, y $0 < \delta < 1$, se puede definir el espacio de Banach $Z_\delta \doteq (D((\lambda I - A)^\delta), \|\cdot\|_\delta)$, donde $\|z\|_\delta \doteq \|(\lambda I - A)^\delta z\|_Z$ (ver Sección 1.5). En lo que resta del capítulo dejaremos $\lambda > \omega$ fijo.

Consideremos la siguiente hipótesis sobre la parte no lineal de (1.11).

(F) *Existen $\delta \in (0, 1)$ y $\theta \in (0, 1]$ tal que $F : \mathcal{Q} \times [0, T] \times Z_\delta \rightarrow Z$ es localmente Lipschitz continua en t y z , es decir, para todo $q \in \mathcal{Q}$ y todo conjunto acotado U de $[0, T] \times Z_\delta$ existe una constante $L = L(q, U)$ tal que*

$$\|F(q, t_1, z_1) - F(q, t_2, z_2)\|_Z \leq L \left(|t_1 - t_2|^\theta + \|z_1 - z_2\|_\delta \right),$$

para todo $(t_i, z_i) \in U$, donde la constante L puede elegirse independiente de q en cualquier compacto \mathcal{Q}_C de \mathcal{Q} .

El siguiente teorema de existencia de soluciones para la ecuación de evolución (1.11) se sigue inmediatamente del Teorema 1.32 y es en realidad una reescritura del mismo adaptada a la situación actual.

Teorema 2.1 *Supongamos que F satisface la hipótesis (F) para algún $\delta \in (0, 1)$ y $\theta \in (0, 1]$. Entonces, para $q \in \mathcal{Q}$ y $z_0 \in Z_\delta$, existe $t_1 = t_1(q, z_0) > 0$ tal que (1.11) tiene una única solución en $[0, t_1)$. Es decir, existe una función $z(\cdot; q) \in C([0, t_1) : Z_\delta) \cap C^1((0, t_1) : Z)$ que satisface la ecuación integral*

$$z(t; q) = T(t)z_0 + \int_0^t T(t-s)F(q, s, z(s; q)) ds, \quad \text{para todo } t \in [0, t_1).$$

Además, $t_1(q, z_0) > 0$ puede elegirse independiente de q en subconjuntos compactos de \mathcal{Q} .

El algoritmo de cuasilinealización que presentaremos a continuación es una especie de algoritmo de Newton, en el sentido que en un paso del mismo se reemplaza $z(t; q)$ por una “aproximación de Taylor” de primer orden con respecto al parámetro o variable q . Para lo cual será necesario definir primero qué se entiende por “la derivada de $z(t; q)$ con respecto a q ”, dado que q es una variable en un espacio de Banach y $z(t; q)$ es un elemento de otro espacio de Banach. Es por eso que antes de presentar el algoritmo, introducimos a continuación la siguiente definición de derivada de Fréchet.

Definición 2.2 *Supongamos que X e Y son espacios de Banach, B es un subconjunto abierto de X y $F : B \rightarrow Y$. Decimos que F es derivable Fréchet en $x_0 \in B$ si existe $\Lambda \in \mathcal{L}(X, Y)$ tal que*

$$\lim_{h \rightarrow 0} \frac{\|F(x_0 + h) - F(x_0) - \Lambda h\|_Y}{\|h\|_X} = 0.$$

El operador Λ se llama Derivada de Fréchet de F en x_0 y se denota con $D_x F(x_0)$ o bien $F_x(x_0)$.

Consideremos ahora el problema de identificación de parámetros (ID) planteado en la Sección 1.7. Es decir, supongamos que poseemos mediciones \hat{z}_i del sistema físico descrito por el PVI (1.11) en tiempos discretos t_i , $i = 1, 2, \dots, m$. Supongamos también que el conjunto de parámetros admisibles \mathcal{Q} es un subconjunto abierto del espacio de parámetros $\tilde{\mathcal{Q}}$, que existe $q^* \in \mathcal{Q}$ que minimiza $J(q)$ sobre \mathcal{Q} , y para todo $q \in \mathcal{Q}$, la solución $z(t; q)$

del PVI (1.11) existe en el intervalo $[0, \tilde{T}]$, donde $\tilde{T} \doteq \max_{1 \leq i \leq m} t_i$.

Para derivar el algoritmo de identificación supongamos por el momento que para cada $t \in [0, \tilde{T}]$ fijo, el mapeo $q \rightarrow z(t; q)$ es diferenciable Fréchet (en el Capítulo 3 se presentarán condiciones suficientes sobre F para que se cumpla esta hipótesis).

Proponemos el siguiente algoritmo: Elegir un estimador inicial $q^0 \in \mathcal{Q}$, hacer $k = 0$ y realizar los siguientes pasos:

Paso 1: Dado el estimador q^k de q^* , aproximar $z(t; q)$ por su expansión de Taylor de primer orden en q alrededor de q^k , es decir, definir $z^{k+1}(t; q) = z(t; q^k) + z_q(t; q^k)(q - q^k)$ donde $z_q(t; q^k)$ denota la derivada de Fréchet de $z(t; q)$ con respecto a q en q^k .

Paso 2: Definir el “criterio de error modificado” $J^k(q)$ como:

$$J^k(q) \doteq \frac{1}{2} \sum_{i=1}^m \|C z^{k+1}(t_i; q) - \hat{z}_i\|_Y^2 \quad (2.1)$$

$$= \frac{1}{2} \sum_{i=1}^m \|C [z(t_i; q^k) + z_q(t_i; q^k)(q - q^k)] - \hat{z}_i\|_Y^2. \quad (2.2)$$

Paso 3: Definir q^{k+1} como el valor de q para el que se alcanza el mínimo del criterio de error modificado $J^k(q)$. Para encontrar efectivamente q^{k+1} , diferenciar (2.2), igualar el resultado a cero y resolver para q . Finalmente, llamar q^{k+1} a esta solución, reemplazar k por $k + 1$ y repetir el Paso 1.

El algoritmo descripto arriba presenta una forma de obtener, a partir de un parámetro $q^k \in \mathcal{Q}$, un parámetro q^{k+1} que en principio podría no pertenecer al conjunto de parámetros admisibles \mathcal{Q} . Para evitar este problema, hemos supuesto que el parámetro óptimo q^* es un punto interior de \mathcal{Q} diciendo que \mathcal{Q} es un subconjunto abierto de $\tilde{\mathcal{Q}}$. Bajo esta hipótesis puede probarse (ver Sección 2.2) que si comenzamos el algoritmo iterativo con un parámetro q^0 suficientemente cercano a q^* , entonces todas las iteraciones producirán parámetros q^k en un entorno de q^* contenido en \mathcal{Q} .

Observar que, a menos que $z_q(t_i; q^k) = 0$ para todo $i = 1, 2, \dots, m$, el funcional $J^k(q)$ es estrictamente convexo, por lo tanto existe una única solución a la ecuación $D_q J^k(q) = 0$ y esta es minimizante global de $J^k(q)$. Si derivamos la expresión (2.2) obtenemos que para todo $h \in \tilde{\mathcal{Q}}$,

$$D_q J^k(q)h = \sum_{i=1}^m \langle D_q \{ \mathcal{C} [z(t_i; q^k) + z_q(t_i; q^k)(q - q^k)] - \hat{z}_i \} h, \mathcal{C} [z(t_i; q^k) + z_q(t_i; q^k)(q - q^k)] - \hat{z}_i \rangle_Y$$

pues hemos supuesto que Y es un espacio de Hilbert real. Luego, para todo $h \in \tilde{\mathcal{Q}}$

$$D_q J^k(q)h = \sum_{i=1}^m \langle \mathcal{C} z_q(t_i; q^k)h, [\mathcal{C} z(t_i; q^k) + z_q(t_i; q^k)(q - q^k)] - \hat{z}_i \rangle_Y.$$

Por lo tanto, la condición $D_q J^k(q) = 0$ se satisface si y sólo si

$$\sum_{i=1}^m \langle \mathcal{C} z_q(t_i; q^k)h, \mathcal{C} z_q(t_i; q^k)(q - q^k) \rangle_Y = - \sum_{i=1}^m \langle \mathcal{C} z_q(t_i; q^k)h, \mathcal{C} z(t_i; q^k) - \hat{z}_i \rangle_Y \quad (2.3)$$

para todo $h \in \tilde{\mathcal{Q}}$.

Según la ecuación (2.3), para determinar q^{k+1} debemos resolver un número infinito de ecuaciones (una para cada $h \in \tilde{\mathcal{Q}}$). Aunque esta ecuación determina q^{k+1} , no es posible calcular q^{k+1} a partir de la misma. Sin embargo, si suponemos que $\tilde{\mathcal{Q}}$ es un espacio de Hilbert separable, éste posee una base ortonormal numerable $\{g_j : j = 1, 2, \dots\}$ y entonces la ecuación (2.3) equivale a

$$\sum_{i=1}^m \langle \mathcal{C} z_q(t_i; q^k)g_j, \mathcal{C} z_q(t_i; q^k)(q - q^k) \rangle_Y = - \sum_{i=1}^m \langle \mathcal{C} z_q(t_i; q^k)g_j, \mathcal{C} z(t_i; q^k) - \hat{z}_i \rangle_Y$$

o, lo que es lo mismo,

$$\sum_{i=1}^m \left\langle g_j, [\mathcal{C} z_q(t_i; q^k)]^* \mathcal{C} z_q(t_i; q^k)(q - q^k) \right\rangle_Y = - \sum_{i=1}^m \left\langle g_j, [\mathcal{C} z_q(t_i; q^k)]^* (\mathcal{C} z(t_i; q^k) - \hat{z}_i) \right\rangle_Y \quad (2.4)$$

para todo $j = 1, 2, \dots$, lo que constituye un conjunto *numerable* de ecuaciones.

Como $\{g_j\}$ es una base ortonormal, el parámetro $q \in \tilde{\mathcal{Q}}$ si y sólo si existe una única sucesión $\alpha = \{\alpha_i\}_{i=1}^{\infty} \in \ell^2$ tal que $q = \sum_{j=1}^{\infty} \alpha_j g_j$, y además $\|q\|_{\tilde{\mathcal{Q}}} = \|\alpha\|_{\ell^2}$. Por lo tanto,

el algoritmo para la identificación de parámetros puede ser reformulado en términos de las sucesiones α de coeficientes de q como sigue. Definamos $\ell^2(\mathcal{Q}) \doteq \{\alpha \in \ell^2 : q_\alpha = \sum_{j=1}^{\infty} \alpha_j g_j \in \mathcal{Q}\}$, y dado $\alpha^k \in \ell^2(\mathcal{Q})$ ($q^k \in \mathcal{Q}$) determinar $\alpha^{k+1} \in \ell^2(\mathcal{Q})$ resolviendo (2.4) para q .

Más precisamente, para cada α , denotemos con q_α a la expresión $\sum_{j=1}^{\infty} \alpha_j g_j$, y sea

$$D(\alpha)\gamma \doteq \sum_{i=1}^m M(t_i; q_\alpha)^* [M(t_i; q_\alpha)\gamma], \quad \gamma \in \ell^2, \quad (2.5)$$

donde para cada $q \in \mathcal{Q}$, $t \in [0, T]$, $M(t; q) : \ell^2 \rightarrow Y$ representa el operador “observación de la derivada de Fréchet” en la base $\{g_j\}$ de $\tilde{\mathcal{Q}}$ que se define por

$$M(t; q)\alpha \doteq \begin{bmatrix} \mathcal{C}z_q(t; q)g_1 & \mathcal{C}z_q(t; q)g_2 & \cdots \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \end{bmatrix} = \mathcal{C}z_q(t; q)q_\alpha$$

y $M(t; q)^* : Y \rightarrow \ell^2$ denota el operador adjunto de $M(t; q)$.

Con esta notación, se deduce fácilmente a partir de (2.4) que α^{k+1} puede calcularse como

$$\begin{aligned} \alpha^{k+1} &= \alpha^k - [D(\alpha^k)]^{-1} \sum_{i=1}^m M(t_i; q_{\alpha^k})^* [\mathcal{C}z(t_i; q_{\alpha^k}) - \hat{z}_i] \\ &\doteq E(\alpha^k) \end{aligned} \quad (2.6)$$

siempre que exista $[D(\alpha^k)]^{-1}$.

Claramente, el esquema iterativo definido por (2.6) no puede implementarse debido a que el espacio de parámetros es de dimensión infinita. Sin embargo, si sabemos que el minimizante q^* pertenece a algún subespacio de dimensión finita $\tilde{\mathcal{Q}}^s \subset \tilde{\mathcal{Q}}$, o si nuestro objetivo es determinar el minimizante q^{s*} de $J(q)$ en un subespacio de dimensión finita $\tilde{\mathcal{Q}}^s$ que de algún modo aproxime al espacio $\tilde{\mathcal{Q}}$, entonces es fácil modificar el algoritmo propuesto previamente. El resultado es un algoritmo en dimensión finita perfectamente implementable.

En efecto, supongamos que queremos encontrar un minimizante $q^{s*} \in \mathcal{Q} \subset \tilde{\mathcal{Q}}^s$, donde $\tilde{\mathcal{Q}}^s$ es un subespacio de dimensión finita de $\tilde{\mathcal{Q}}$, con base ortonormal $\{g_j : j = 1, 2, \dots, s\}$. Dado un estimador $q^{s,k} \in \tilde{\mathcal{Q}}^s$ del minimizante $q^{s*} \in \mathcal{Q}$, procedemos del mismo modo

en que lo hicimos con anterioridad, pero pedimos a $q^{k+1,s}$ que satisfaga (2.4) sólo para $j = 1, 2, \dots, s$.

Ahora bien, el parámetro $q \in \tilde{\mathcal{Q}}^s$ si y sólo si existe un único vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_s) \in \mathbb{R}^s$ tal que $q = \sum_{j=1}^s \alpha_j g_j$, y además $\|q\|_{\tilde{\mathcal{Q}}} = |\alpha| = (\sum_{j=1}^s \alpha_j^2)^{1/2}$. Por lo tanto, el algoritmo para la identificación de parámetros puede ser reformulado en términos de los vectores α de coeficientes de q como sigue. Definamos $\mathbb{R}^s(\mathcal{Q}) \doteq \{\alpha \in \mathbb{R}^s : q_\alpha = \sum_{j=1}^s \alpha_j g_j \in \mathcal{Q}\}$, y dado $\alpha^k \in \mathbb{R}^s(\mathcal{Q})$ ($q^k \in \mathcal{Q}$) determinar $\alpha^{k+1} \in \mathbb{R}^s(\mathcal{Q})$ resolviendo (2.4) en q .

Análogamente a como se hizo en el caso de dimensión infinita, denotando con q_α al elemento $\sum_{j=1}^s \alpha_j g_j$ de $\tilde{\mathcal{Q}}^s$, ($\alpha \in \mathbb{R}^s$) y definiendo el operador lineal $D(\alpha) : \mathbb{R}^s \rightarrow \mathbb{R}^s$ por

$$D(\alpha)\gamma \doteq \sum_{i=1}^m M(t_i; q_\alpha)^* [M(t_i; q_\alpha)\gamma], \quad \gamma \in \mathbb{R}^s,$$

donde ahora, para cada $q \in \mathcal{Q}$, $t \in [0, \tilde{T}]$, $M(t; q) : \mathbb{R}^s \rightarrow Y$ se define por

$$M(t; q)\alpha \doteq \begin{bmatrix} \mathcal{C}z_q(t; q)g_1 & \mathcal{C}z_q(t; q)g_2 & \cdots & \mathcal{C}z_q(t; q)g_s \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_s \end{bmatrix} = \mathcal{C}z_q(t; q)q_\alpha$$

y $M(t; q)^* : Y \rightarrow \mathbb{R}^s$ denota el operador adjunto de $M(t; q)$, obtenemos el esquema iterativo siguiente: Dado $\alpha^k \in \mathbb{R}^s$, calcular $\alpha^{k+1} \in \mathbb{R}^s$ como

$$\begin{aligned} \alpha^{k+1} &= \alpha^k - [D(\alpha^k)]^{-1} \sum_{i=1}^m M(t_i; q_{\alpha^k})^* [\mathcal{C}z(t_i; q_{\alpha^k}) - \hat{z}_i] \\ &\doteq E(\alpha^k) \end{aligned} \tag{2.7}$$

siempre que exista $[D(\alpha^k)]^{-1}$.

Observación: Hemos utilizado la misma notación para definir $M(t; q)$, $D(\alpha)$ y $E(\alpha)$, tanto para el algoritmo en dimensión infinita como para el algoritmo en dimensión finita.

La razón por la cual hemos hecho esto es porque los resultados de convergencia que daremos en la próxima sección permanecen válidos independientemente de la dimensión del espacio. En la próxima sección presentaremos los resultados de convergencia para el caso de dimensión infinita, pero las modificaciones para contemplar el caso de dimensión finita son solamente cuestión de notación (cambiar $\ell^2(\mathcal{Q})$ por $\mathbb{R}^s(\mathcal{Q})$), y no es necesario hacer distinciones entre estos dos casos.

2.2 Convergencia del Algoritmo de Cuasilinealización

En esta sección estudiaremos la convergencia del algoritmo presentado en la sección anterior. Presentaremos dos resultados que proveen condiciones suficientes para que el algoritmo converja. En el primero de ellos supondremos que existe un parámetro q^* que satisface $J(q^*) = 0$ (hipótesis de ajuste exacto), y en el segundo resultado reemplazaremos esta hipótesis por una más débil (datos con errores de medición).

Antes de presentar estos resultados serán necesarios dos lemas preliminares. En el primero de ellos, la hipótesis principal se refiere a la derivada de Fréchet del mapeo $q \rightarrow z(t; q)$ de \mathcal{Q} en Z_δ (para t fijo). Si esta derivada existe, entonces satisface $z_q(t; q) \in \mathcal{L}(\tilde{\mathcal{Q}}, Z_\delta)$ (ver Definición 2.2), pero como Z_δ está continuamente inmerso en Z (ver Teorema 1.26(b)), se tiene que $\mathcal{L}(\tilde{\mathcal{Q}}, Z_\delta) \subset \mathcal{L}(\tilde{\mathcal{Q}}, Z)$ y por lo tanto, también $z_q(t; q) \in \mathcal{L}(\tilde{\mathcal{Q}}, Z)$.

Lema 2.3 *Sea $t \in [0, \tilde{T}]$ fijo y sea $M(t; q)$ como se definió en la sección anterior. Si el mapeo $q \rightarrow z(t; q)$ de $\mathcal{Q} \rightarrow Z_\delta$ tiene derivada de Fréchet $z_q(t; q)$ localmente Lipschitz continua como función de $\mathcal{Q} \rightarrow \mathcal{L}(\tilde{\mathcal{Q}}, Z)$, entonces el mapeo $\alpha \rightarrow M(t; q_\alpha)$ es también localmente Lipschitz continuo de $\ell^2(\mathcal{Q}) \rightarrow \mathcal{L}(\ell^2, Y)$. Más precisamente, para todo $\alpha \in \ell^2(\mathcal{Q})$ y $t \in [0, \tilde{T}]$, existen constantes positivas y finitas η_α y \mathcal{L}_α dependientes de α y de t tales que*

$$\|M(t; q_\alpha) - M(t; q_{\tilde{\alpha}})\|_{\mathcal{L}(\ell^2, Y)} \leq \mathcal{L}_\alpha \|\alpha - \tilde{\alpha}\|_{\ell^2}, \quad \text{para todo } \tilde{\alpha} \in B(\alpha, \eta_\alpha),$$

donde $B(\alpha, \eta_\alpha) \doteq \{\beta \in \ell^2 : \|\alpha - \beta\|_{\ell^2} < \eta_\alpha\}$. El mismo resultado se cumple para el mapeo $\alpha \rightarrow M(t; q_\alpha)^*$.

Demostración. Sea $t \in [0, \tilde{T}]$ fijo. Por hipótesis, para todo $\alpha \in \ell^2(\mathcal{Q})$ existen $\eta_\alpha > 0$, y L_α tales que

$$\|z_q(t; q_\alpha) - z_q(t; q_{\tilde{\alpha}})\|_{\mathcal{L}(\tilde{\mathcal{Q}}, Z)} < L_\alpha \|q_\alpha - q_{\tilde{\alpha}}\|_{\tilde{\mathcal{Q}}}, \quad \text{para todo } \tilde{\alpha} \in B(\alpha, \eta_\alpha).$$

Se sigue entonces que

$$\begin{aligned}
\|M(t; q_\alpha) - M(t; q_{\tilde{\alpha}})\|_{\mathcal{L}(\ell^2, Y)} &= \sup_{\substack{\gamma \in \ell^2 \\ \|\gamma\|=1}} \|[M(t; q_\alpha) - M(t; q_{\tilde{\alpha}})] \gamma\|_Y \\
&= \sup_{\substack{\gamma \in \ell^2 \\ \|\gamma\|=1}} \|\mathcal{C}z_q(t; q_\alpha)q_\gamma - \mathcal{C}z_q(t; q_{\tilde{\alpha}})q_\gamma\|_Y \\
&\leq \|\mathcal{C}\|_{\mathcal{L}(Z, Y)} \sup_{\substack{\gamma \in \ell^2 \\ \|\gamma\|=1}} \left\{ \|z_q(t; q_\alpha) - z_q(t; q_{\tilde{\alpha}})\|_{\mathcal{L}(\tilde{\mathcal{Q}}, Z)} \|q_\gamma\|_{\tilde{\mathcal{Q}}} \right\} \\
&\leq \|\mathcal{C}\|_{\mathcal{L}(Z, Y)} L_\alpha \|q_\alpha - q_{\tilde{\alpha}}\|_{\tilde{\mathcal{Q}}} \\
&= \|\mathcal{C}\|_{\mathcal{L}(Z, Y)} L_\alpha \|\alpha - \tilde{\alpha}\|_{\ell^2} = \mathcal{L}_\alpha \|\alpha - \tilde{\alpha}\|_{\ell^2}
\end{aligned}$$

para todo $\tilde{\alpha} \in B(\alpha, \eta_\alpha)$. □

El siguiente lema utiliza el resultado del lema anterior para demostrar que bajo las mismas hipótesis, el mapeo $\alpha \rightarrow D(\alpha)$ es Lipschitz continuo, lo que será esencial para la demostración de los resultados de convergencia del algoritmo de cuasilinealización propuesto en la Sección 2.1.

Lema 2.4 *Bajo las mismas hipótesis del Lema 2.3, el mapeo $\alpha \rightarrow D(\alpha)$ es localmente Lipschitz continuo de $\ell^2(\mathcal{Q}) \rightarrow \mathcal{L}(\ell^2, \ell^2)$.*

Demostración. De la definición de $D(\alpha)$, se tiene que si $\alpha, \tilde{\alpha} \in \ell^2(\mathcal{Q})$ y $\gamma \in \ell^2$, entonces

$$\begin{aligned}
[D(\alpha) - D(\tilde{\alpha})] \gamma &= \sum_{i=1}^m M(t_i; q_\alpha)^* M(t_i; q_\alpha) \gamma - \sum_{i=1}^m M(t_i; q_{\tilde{\alpha}})^* M(t_i; q_{\tilde{\alpha}}) \gamma \\
&= \sum_{i=1}^m M(t_i; q_\alpha)^* [M(t_i; q_\alpha) - M(t_i; q_{\tilde{\alpha}})] \gamma \\
&\quad + \sum_{i=1}^m [M(t_i; q_\alpha)^* - M(t_i; q_{\tilde{\alpha}})^*] M(t_i; q_{\tilde{\alpha}}) \gamma,
\end{aligned}$$

y el resultado se sigue inmediatamente del Lema 2.3. □

Antes de enunciar los resultados principales con relación a la convergencia del algoritmo de cuasilinealización (AC), será necesario introducir el concepto de *punto de atracción*. A continuación se da su definición y una condición suficiente para que la iteración de un mapeo en un espacio de Banach tenga un punto de atracción.

Definición 2.5 Sea U un subconjunto abierto de un espacio de Banach X y $E : U \subset X \rightarrow X$. Se dice que x^* es un punto de atracción de la iteración $x^{k+1} = E(x^k)$ si existe un entorno S de x^* tal que $S \subset U$ y para todo $x^0 \in S$, las iteraciones $x^k \in U$ para todo $k \geq 1$ y $x^k \rightarrow x^*$ cuando $k \rightarrow \infty$.

Lema 2.6 (Teorema del Punto Fijo) Sea U un abierto de un espacio de Banach X , $E : U \subset X \rightarrow X$, $x^* \in U$ y supongamos que existen una bola $B = B(x^*, \eta) \subset U$ y una constante $\alpha \in (0, 1)$ tal que

$$\|E(x) - x^*\| \leq \alpha \|x - x^*\|, \quad \text{para todo } x \in B.$$

Entonces x^* es un punto de atracción de la iteración $x^{k+1} = E(x^k)$.

Demostración. Si $x^0 \in B$, se tiene que $\|x^1 - x^*\| = \|E(x^0) - x^*\| \leq \alpha \|x^0 - x^*\|$, por lo que $x^1 \in B$. Por inducción $\|x^{k+1} - x^*\| = \|E(x^k) - x^*\| \leq \alpha \|x^k - x^*\| \leq \alpha^{k+1} \|x^0 - x^*\|$ y $\alpha^{k+1} \|x^0 - x^*\| \rightarrow 0$ cuando $k \rightarrow \infty$. \square

A continuación presentamos el primer resultado de convergencia del algoritmo.

Teorema 2.7 (Convergencia del AC con la hipótesis de ajuste exacto) Supongamos que se cumple la hipótesis del Lema 2.3. Supongamos también que existen un abierto $U \subset \ell^2(\mathcal{Q})$ y $\alpha^* \in U$ tales que existe $D(\alpha^*)^{-1}$ y $J(q_{\alpha^*}) = 0$. Sea E el mapeo definido por (2.6). Entonces, para todo $\epsilon > 0$, existe una constante $\delta > 0$, dependiente de ϵ y α^* tal que $\|\alpha - \alpha^*\|_{\ell^2} < \delta$ implica

$$\|E(\alpha) - \alpha^*\|_{\ell^2} \leq K \|\alpha - \alpha^*\|_{\ell^2}^2 + \epsilon \|\alpha - \alpha^*\|_{\ell^2},$$

donde K es una constante que depende solamente de α^* (no de ϵ). En particular, α^* es un punto de atracción de la iteración $\alpha^{k+1} = E(\alpha^k)$.

Demostración. Por definición (ver (2.6))

$$E(\alpha) = \alpha - D(\alpha)^{-1} \sum_{i=1}^m M(t_i; q_\alpha)^* [\mathcal{C}z(t_i; q_\alpha) - \hat{z}_i]$$

siempre que $D(\alpha)^{-1}$ exista. Se tiene entonces que

$$\begin{aligned} E(\alpha) - \alpha^* &= \alpha - D(\alpha)^{-1} \left\{ \sum_{i=1}^m M(t_i; q_\alpha)^* [\mathcal{C}z(t_i; q_\alpha) - \hat{z}_i] \right\} - \alpha^* \\ &= D(\alpha)^{-1} \left\{ D(\alpha) (\alpha - \alpha^*) - \sum_{i=1}^m M(t_i; q_\alpha)^* [\mathcal{C}z(t_i; q_\alpha) - \hat{z}_i] \right\} \\ &= D(\alpha)^{-1} \left\{ \sum_{i=1}^m M(t_i; q_\alpha)^* [M(t_i; q_\alpha)(\alpha - \alpha^*) - \mathcal{C}z(t_i; q_\alpha) + \hat{z}_i] \right\} \end{aligned}$$

donde la última desigualdad se sigue de la definición de $D(\alpha)$ (ver (2.5)). Luego

$$\begin{aligned} E(\alpha) - \alpha^* &= D(\alpha)^{-1} \left\{ \sum_{i=1}^m M(t_i; q_\alpha)^* [M(t_i; q_\alpha) - M(t_i; q_{\alpha^*})] (\alpha - \alpha^*) \right\} \\ &\quad - D(\alpha)^{-1} \left\{ \sum_{i=1}^m M(t_i; q_\alpha)^* [\mathcal{C}z(t_i; q_\alpha) - \mathcal{C}z(t_i; q_{\alpha^*}) - M(t_i; q_{\alpha^*})(\alpha - \alpha^*)] \right\} \\ &\quad - D(\alpha)^{-1} \left\{ \sum_{i=1}^m M(t_i; q_\alpha)^* [\mathcal{C}z(t_i; q_{\alpha^*}) - \hat{z}_i] \right\}. \end{aligned}$$

Como $J(q_{\alpha^*}) = 0$, $\mathcal{C}z(t_i; q_{\alpha^*}) = \hat{z}_i$, para todo $i = 1, 2, \dots, m$ y el tercer término del lado derecho es igual a cero. Además, como $D(\alpha^*)^{-1}$ existe por hipótesis, por continuidad también existen constantes positivas δ_1 y D , dependientes de α^* tales que $D(\alpha)$ es invertible para todo $\alpha \in B(\alpha^*, \delta_1)$ y $\|D(\alpha)^{-1}\| \leq D$ para todo $\alpha \in B(\alpha^*, \delta_1)$. Del Lema 2.3 se sigue que existe M tal que $\|M(t_i; q_\alpha)^*\| \leq M$ para todo $i = 1, 2, \dots, m$ y $\|\alpha - \alpha^*\|_{\ell^2} < \delta_1$.

En consecuencia, para todo $\alpha \in B(\alpha^*, \delta_1)$

$$\begin{aligned} \|E(\alpha) - \alpha^*\|_{\ell^2} &\leq DM \sum_{i=1}^m \| [M(t_i; q_\alpha) - M(t_i; q_{\alpha^*})] (\alpha - \alpha^*) \| \\ &\quad + DM \sum_{i=1}^m \| \mathcal{C}z(t_i; q_\alpha) - \mathcal{C}z(t_i; q_{\alpha^*}) - M(t_i; q_{\alpha^*})(\alpha - \alpha^*) \| \\ &\doteq A + B. \end{aligned}$$

Por el Lema 2.3, existen constantes finitas $\eta_{\alpha^*} > 0$ y \mathcal{L}_{α^*} , tales que si $|\alpha - \alpha^*| < \eta_{\alpha^*}$, entonces $A \leq DMm\mathcal{L}_{\alpha^*} \|\alpha - \alpha^*\|_{\ell^2}^2$. Además, como

$$M(t_i; q_{\alpha^*})(\alpha - \alpha^*) = \mathcal{C}z_q(t_i; q_{\alpha^*})(q_\alpha - q_{\alpha^*}),$$

por la definición de la derivada de Fréchet $z_q(t; q)$, para todo $\epsilon > 0$, existe $\delta_2 = \delta_2(\epsilon, \alpha^*) > 0$ tal que $\|\alpha - \alpha^*\|_{\ell^2} < \delta_2$ ($\|q_\alpha - q_{\alpha^*}\|_{\mathcal{Q}} < \delta_2$) implica

$$\|\mathcal{C}z(t_i; q_\alpha) - \mathcal{C}z(t_i; q_{\alpha^*}) - M(t_i; q_{\alpha^*})(\alpha - \alpha^*)\|_Y \leq \epsilon \|q_\alpha - q_{\alpha^*}\|_{\mathcal{Q}} = \epsilon \|\alpha - \alpha^*\|_{\ell^2},$$

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

Resumiendo, se tiene que para todo $\epsilon > 0$ existen constantes finitas $\delta = \min \{\delta_1, \delta_2, \eta_{\alpha^*}\} > 0$ y \mathcal{L}_{α^*} tales que

$$\|E(\alpha) - \alpha^*\|_{\ell^2} \leq DMm [\mathcal{L}_{\alpha^*} \|\alpha - \alpha^*\|_{\ell^2}^2 + \epsilon \|\alpha - \alpha^*\|_{\ell^2}]$$

para todo $\alpha \in B(\alpha^*, \delta)$. Por el Lema 2.6, α^* es un punto de atracción de la iteración $\alpha^{k+1} = E(\alpha^k)$. \square

Es importante notar aquí que en el Teorema 2.7 asumimos que $J(q_{\alpha^*}) = 0$, es decir, supusimos que existe un parámetro α^* que ajusta *exactamente* los datos observados $\{\hat{z}_i\}_{i=1}^m$. En la práctica, cuando se trabaja en problemas reales de identificación de parámetros, ésta no es una hipótesis realista, debido a que pueden existir errores de observación, medición y modelado. En el próximo teorema se debilita esta hipótesis de ajuste exacto.

Teorema 2.8 (Convergencia del AC con datos no exactos) *Supongamos que se cumple la hipótesis del Lema 2.3. Supongamos también que existe un abierto $U \subset \ell^2(\mathcal{Q})$ y $\alpha^* \in U$ tal que $D(\alpha^*)$ es no singular y $\alpha^* = E(\alpha^*)$ (punto fijo). Sea $\delta_1 > 0$ y $D \doteq \sup \{\|D(\alpha)^{-1}\| : \|\alpha - \alpha^*\|_{\ell^2} < \delta_1\} < \infty$ como en el Teorema 2.7, y sea \mathcal{L} la menor constante que satisface*

$$\|M(t_i; q_\alpha)^* - M(t_i; q_{\alpha^*})^*\| \leq \mathcal{L} \|\alpha - \alpha^*\|_{\ell^2}, \quad \forall \|\alpha - \alpha^*\|_{\ell^2} < \delta_1, \quad i = 1, 2, \dots, m.$$

Supongamos que

$$\gamma \doteq D\mathcal{L} \sum_{i=1}^m \|\mathcal{C}z(t_i; q_{\alpha^*}) - \hat{z}_i\|_Y < 1, \tag{2.8}$$

entonces α^ es un punto de atracción de la iteración $\alpha^{k+1} = E(\alpha^k)$.*

Observación: En las hipótesis del Teorema 2.8 no se pide que $\mathcal{C}z(t_i; q_{\alpha^*}) = \hat{z}_i$ para todo $i = 1, 2, \dots, m$. Sin embargo, para poder asegurar la convergencia es necesario que $\mathcal{C}z(t_i; q_{\alpha^*})$ esté cerca de \hat{z}_i en algún sentido. Este sentido está dado por la desigualdad (2.8). En otras palabras, el enunciado del teorema dice que uno puede cometer errores al tomar las mediciones experimentales, pero que para poder realizar la identificación de los parámetros, esos errores deben ser pequeños.

Demostración. Siguiendo los mismos pasos que en la demostración del Teorema 2.7, se ve que dado $\epsilon > 0$ existen $\delta = \delta(\epsilon, \alpha^*) > 0$ y $M = M(\alpha^*)$ tales que

$$\begin{aligned} \|E(\alpha) - \alpha^*\|_{\ell^2} &\leq DMm [\mathcal{L} \|\alpha - \alpha^*\|_{\ell^2}^2 + \epsilon \|\alpha - \alpha^*\|_{\ell^2}] \\ &\quad + \left\| D(\alpha)^{-1} \sum_{i=1}^m M(t_i; q_{\alpha})^* [\mathcal{C}z(t_i; q_{\alpha^*}) - \hat{z}_i] \right\|, \end{aligned} \quad (2.9)$$

para todo $\alpha \in B(\alpha^*, \delta)$, donde el segundo sumando aparece porque ahora no estamos suponiendo $J(q_{\alpha^*}) = 0$.

Por otro lado,

$$\sum_{i=1}^m M(t_i; q_{\alpha^*})^* [\mathcal{C}z(t_i; q_{\alpha^*}) - \hat{z}_i] = 0, \quad (2.10)$$

pues, por hipótesis, $D(\alpha^*)$ es no singular y $\alpha^* = E(\alpha^*)$. Combinando (2.9) y (2.10) obtenemos

$$\begin{aligned} \|E(\alpha) - \alpha^*\|_{\ell^2} &\leq DMm [\mathcal{L} \|\alpha - \alpha^*\|_{\ell^2}^2 + \epsilon \|\alpha - \alpha^*\|_{\ell^2}] \\ &\quad + D \left\| \sum_{i=1}^m [M(t_i; q_{\alpha})^* - M(t_i; q_{\alpha^*})^*] [\mathcal{C}z(t_i; q_{\alpha^*}) - \hat{z}_i] \right\| \\ &\leq DMm [\mathcal{L} \|\alpha - \alpha^*\|_{\ell^2}^2 + \epsilon \|\alpha - \alpha^*\|_{\ell^2}] \\ &\quad + D\mathcal{L} \|\alpha - \alpha^*\|_{\ell^2} \sum_{i=1}^m \|\mathcal{C}z(t_i; q_{\alpha^*}) - \hat{z}_i\| \\ &= DMm [\mathcal{L} \|\alpha - \alpha^*\|_{\ell^2}^2 + \epsilon \|\alpha - \alpha^*\|_{\ell^2}] + \gamma \|\alpha - \alpha^*\|_{\ell^2} \end{aligned}$$

donde $\gamma < 1$ por hipótesis. Esto concluye la demostración del teorema. \square

2.3 Presentación Clásica del AC

En la literatura clásica [12, 28] el algoritmo de cuasilinealización se presenta de una manera muy diferente a la forma en que fue introducido aquí. Por razones de completitud,

presentamos brevemente esta otra forma de introducir el algoritmo. A pesar de que a primera vista los métodos pueden parecer completamente diferentes, mostraremos que ambos conducen al mismo proceso iterativo.

Supongamos por un momento que el término no lineal $F(q, t, z)$ de (1.11) es diferenciable Fréchet con respecto a q y a z . Dada una estimación $q^k \in \mathcal{Q}$ del parámetro óptimo $q^* \in \mathcal{Q}$, definimos $z^k(t) = z(t; q^k)$ y linealizamos el problema (1.11) alrededor de $(q^k, z^k(t))$. Esto nos da el siguiente PVI

$$\begin{cases} \frac{dz(t)}{dt} = Az^k(t) + F(q^k, t, z^k(t)) \\ \quad + F_q(q^k, t, z^k(t))(q - q^k) \\ \quad + A(z(t) - z^k(t)) + F_z(q^k, t, z^k(t))(z(t) - z^k(t)) \\ z(0) = z_0 \end{cases} \quad (2.11)$$

A continuación, definimos $z^{k+1}(t; q)$ como la solución del problema (2.11) y elegimos q^{k+1} como el minimizador del criterio de error modificado

$$J^k(q) \doteq \frac{1}{2} \sum_{i=1}^m \|\mathcal{C}z^{k+1}(t_i; q) - \hat{z}_i\|_Y^2.$$

Observando el problema (2.11) vemos que $v(t) = z^{k+1}(t; q) - z^k(t)$ es una solución del PVI lineal no homogéneo

$$\begin{cases} \frac{dv(t)}{dt} = Av(t) + F_q(q^k, t, z(t; q^k))(q - q^k) + F_z(q^k, t, z(t; q^k))v(t), \\ v(0) = 0. \end{cases} \quad (2.12)$$

La ecuación diferencial en (2.12) se conoce como *ecuación de sensibilidad* asociada al problema de identificación (ID). En la Sección 3.1 veremos que bajo ciertas condiciones sobre el operador A y la función F , resulta que la solución del PVI (2.12) $v(t)$ es la derivada de Fréchet de $z(t; q)$ con respecto a q evaluada en q^k aplicada a $(q - q^k)$, es decir,

$$v(t) = z_q(t; q^k)(q - q^k),$$

lo cual implica que

$$z^{k+1}(t; q) = z(t; q^k) + z_q(t; q^k)(q - q^k).$$

Luego, el criterio de error modificado resulta

$$J^k(q) = \frac{1}{2} \sum_{i=1}^m \left\| \mathcal{C} [z(t; q^k) + z_q(t; q^k)(q - q^k)] - \hat{z}_i \right\|_Y^2.$$

que es el mismo criterio de error modificado que se obtuvo en la Sección 2.1.

Como podemos ver, la presentación clásica al algoritmo de cuasilinealización se basa en la *linealización del PVI* inicial alrededor de la solución correspondiente al parámetro de la iteración actual y su derivación requiere del conocimiento previo de la ecuación de sensibilidad (2.12). Por otro lado, el método introducido en la Sección 2.1 se basa simplemente en la *linealización de la solución del PVI* alrededor del parámetro y para la derivación del algoritmo no fueron necesarias las ecuaciones de sensibilidad. Es importante señalar, sin embargo, que para la implementación computacional de ambos métodos se utilizan las derivadas de las soluciones con respecto al parámetro y en consecuencia, la solución del PVI (2.12).

Capítulo 3

Diferenciabilidad con Respecto a Parámetros

En el capítulo anterior vimos que una condición suficiente para que el algoritmo de cuasilinealización converja es que exista la derivada de Fréchet $z_q(t; q)$ del mapeo $q \rightarrow z(t; q)$ de \mathcal{Q} en Z_δ , y que además, el mapeo $q \rightarrow z_q(t; q)$ sea localmente Lipschitz continuo como función de \mathcal{Q} en $\mathcal{L}(\tilde{\mathcal{Q}}, Z)$. En este capítulo encontraremos condiciones suficientes sobre los operadores que definen el problema de valores iniciales (1.11) para que las soluciones cumplan con estas hipótesis. Específicamente, en la Sección 3.1 veremos condiciones suficientes para que la solución $z(t; q)$ sea diferenciable Fréchet y en la Sección 3.2 exponaremos condiciones suficientes para que la derivada $z_q(t; q)$ sea Lipschitz continua con respecto al parámetro q .

El problema de la diferenciabilidad con respecto a parámetros ha sido estudiado anteriormente pero en contextos diferentes del que nos interesa en este trabajo. En 1977, Clark y Gibson [20] analizaron este problema de diferenciabilidad en el contexto de problemas de Cauchy abstractos lineales del tipo

$$\frac{dz(t)}{dt} = A(q)z(t) + u(t), \quad z(0) = z_0$$

donde $A(q)$ es el generador infinitesimal de un semigrupo fuertemente continuo y puede escribirse como $A(q) = A + B(q)$, con $B(q)$ acotado, es decir, el parámetro aparece en la parte acotada del operador $A(q)$. Más adelante, en 1982, Brewer [11] estudió el mismo problema bajo hipótesis más débiles. Aquí el parámetro q no se restringió solamente a

aparecer en la parte acotada del operador $A(q)$. Hasta ahora nadie estudió el problema de la diferenciabilidad con respecto a parámetros de las soluciones de problemas no lineales del tipo (1.11), donde el parámetro aparece en la parte no lineal.

3.1 Existencia de la Derivada de Fréchet de $z(t; q)$

Antes de analizar y exponer las condiciones que aseguran la diferenciabilidad del mapeo $q \rightarrow z(t; q)$ veremos condiciones que nos permiten concluir que tal mapeo es Lipschitz. Este resultado será necesario para la demostración de aquellas. Supongamos de ahora en adelante que $\omega \in \mathbb{R}$ está fijo y $\omega > \sup \{\operatorname{Re} \lambda : \lambda \in \sigma(A)\}$, y llamemos, para $\delta \in (0, 1)$, Z_δ al espacio de Banach $D(-A_\omega)^\delta$ con la norma $\|z\|_\delta = \|(-A_\omega)^\delta z\|_Z$ tal como se definió en la Sección 1.5.

Teorema 3.1 *Supongamos que $F(q, t, z)$ satisface la hipótesis **(F)** presentada en la Sección 2.1, y que el mapeo $q \rightarrow F(q, \cdot, z)$ de \mathcal{Q} en $L^\infty(0, T : Z)$ es localmente Lipschitz continuo para todo $z \in Z_\delta$ con constante Lipschitz independiente de z en conjuntos Z_δ -acotados. Sean $z_0 \in Z_\delta$ y $q^* \in \mathcal{Q}$, y sea $t_1 = t_1(q^*, z_0) > 0$ tal que $[0, t_1)$ es el intervalo maximal de existencia de la solución $z(t; q^*)$ de (1.11). Entonces para todo $t'_1 < t_1$, existen constantes $\epsilon > 0$ y $C > 0$ tales que si $\|q - q^*\| \leq \epsilon$ entonces $z(t; q)$ existe en $[0, t'_1]$ y satisface*

$$\|z(\cdot; q) - z(\cdot; q^*)\|_{L^\infty(0, t'_1; Z_\delta)} \leq C \|q - q^*\|.$$

Es decir, el mapeo $q \rightarrow z(\cdot; q)$ es localmente Lipschitz continuo en el sentido de lo expresado en el enunciado de este teorema.

Demostración. Sea $\delta \in (0, 1)$ tal que se cumple la hipótesis **(F)**. Luego, para $z_0 \in Z_\delta$, y $q^* \in \mathcal{Q}$, existe $t_1 = t_1(q^*, z_0) > 0$ tal que $[0, t_1)$ es el intervalo maximal de existencia de $z(t; q^*)$. Fijemos $0 < t'_1 < t_1$. Para cada $q \in \mathcal{Q}$, sea $t_q > 0$ tal que $z(t; q)$ existe en $[0, t_q)$,

se tiene entonces que para $t \in [0, \min\{t_1, t_q\})$,

$$z(t; q^*) = T(t)z_0 + \int_0^t T(t-s)F(q^*, s, z(s; q^*)) ds,$$

y

$$z(t; q) = T(t)z_0 + \int_0^t T(t-s)F(q, s, z(s; q)) ds,$$

y por lo tanto, si $t \in [0, \min\{t_1, t_q\})$,

$$\begin{aligned} z(t; q^*) - z(t; q) &= \int_0^t T(t-s) [F(q^*, s, z(s; q^*)) - F(q, s, z(s; q))] ds \\ &= \int_0^t T(t-s) [F(q^*, s, z(s; q^*)) - F(q, s, z(s; q^*))] ds \\ &\quad + \int_0^t T(t-s) [F(q, s, z(s; q^*)) - F(q, s, z(s; q))] ds. \end{aligned}$$

Luego, se sigue por el Teorema 1.29(c) que para todo $q \in \mathcal{Q}$ y $t \in [0, \min\{t_1, t_q\})$

$$\begin{aligned} \|z(t; q^*) - z(t; q)\|_\delta &= \|(-A_\omega)^\delta (z(t; q^*) - z(t; q))\|_Z \\ &\leq \int_0^t \|(-A_\omega)^\delta T(t-s)\|_{\mathcal{L}(Z)} \|F(q^*, s, z(s; q^*)) - F(q, s, z(s; q^*))\|_Z ds \\ &\quad + \int_0^t \|(-A_\omega)^\delta T(t-s)\|_{\mathcal{L}(Z)} \|F(q, s, z(s; q^*)) - F(q, s, z(s; q))\|_Z ds \\ &\leq \int_0^t \frac{C}{(t-s)^\delta} \|F(q^*, s, z(s; q^*)) - F(q, s, z(s; q^*))\|_Z ds \\ &\quad + \int_0^t \frac{C}{(t-s)^\delta} \|F(q, s, z(s; q^*)) - F(q, s, z(s; q))\|_Z ds \\ &\doteq \rho_1(t) + \rho_2(t). \end{aligned}$$

Por la hipótesis de continuidad Lipschitz de $F(q, t, z)$ con respecto a q y como q^* está fijo, de modo que el conjunto $\{z(s; q^*) : 0 \leq s \leq t'_1\}$ es Z_δ -acotado, se tiene que

$$\|F(q^*, \cdot, z(\cdot; q^*)) - F(q, \cdot, z(\cdot; q^*))\|_{L^\infty(0, t'_1; Z)} \leq C_1 \|q - q^*\|$$

y por lo tanto existe una constante $C_F = \frac{CC_1(t'_1)^{1-\delta}}{1-\delta}$ tal que $\rho_1(t) \leq C_F \|q - q^*\|$, para todo $t \in [0, t'_1)$.

Resulta oportuno notar aquí que si $F(q, t, z)$ fuera globalmente Lipschitz continua con respecto a z , se tendría que $\rho_2(t) \leq \int_0^t \frac{C}{(t-s)^\delta} \|z(s; q^*) - z(s; q)\| ds$, y en virtud del

Lema 1.33 se tendría el resultado deseado. Sin embargo la hipótesis **(F)** es más débil: sólo dice que $F(q, t, z)$ es *localmente* Lipschitz continua con respecto a z . Si bien es cierto que la demostración del teorema, en este caso, se vuelve considerablemente más complicada, es importante destacar que este resultado permite su aplicabilidad a una gama mucho más amplia de problemas no lineales (por ejemplo, a los casos de no linealidades polinómicas en las variables de estado), algunos de los cuales estudiaremos en detalle en los capítulos siguientes.

Como $z(0; q) = z(0; q^*)$, se tiene que existe $\eta_q > 0$ tal que $\|z(t; q)\|_\delta \leq M + C_F + 2$ para todo $t \in [0, \eta_q]$, donde $M \doteq \sup_{t \in [0, t'_1]} \|z(t; q^*)\|_\delta$. Sea L la constante Lipschitz de $F(q, t, z)$ correspondiente al conjunto $U = [0, t'_1] \times \{\|z\|_\delta \leq M + C_F + 2\}$, entonces

$$\rho_2(t) \leq \int_0^t \frac{CL}{(t-s)^\delta} \|z(s; q^*) - z(s; q)\|_\delta ds \quad \text{para todo } t \in [0, \eta_q],$$

y si $K = K(\delta, CL, t'_1)$ como en el Lema 1.33, entonces

$$\|z(t; q^*) - z(t; q)\|_\delta \leq f_q(t), \quad \text{para todo } t \in [0, \eta_q], \quad (3.1)$$

donde $f_q(t)$ se define, para todo $t \in [0, t'_1]$, por

$$f_q(t) \doteq C_F \|q - q^*\| + KC_F \|q - q^*\| \int_0^t 1/(t-s)^\delta ds.$$

Observar que (3.1) nos da la cota buscada pero sólo en $[0, \eta_q]$. Veremos a continuación que tal cota vale también en todo el intervalo $[0, t'_1]$. Para ello, observemos que, a partir de la definición de $f_q(t)$, existe $\epsilon = \epsilon(K) > 0$ tal que $f_q(t) \leq C_F + 1$ en $[0, t'_1]$ si $\|q - q^*\| < \epsilon$ y entonces, para $\|q - q^*\| < \epsilon$,

$$\|z(t; q) - z(t; q^*)\|_\delta \leq C_F + 1, \quad \text{para todo } t \in [0, \eta_q].$$

Probaremos ahora que si $q \in \mathcal{Q}$ y $\|q - q^*\| < \epsilon$, entonces $z(t; q)$ existe en $[0, t'_1]$ y $\|z(t; q)\|_\delta < M + C_F + 2$ en $[0, t'_1]$. En efecto, supongamos por el contrario, que existe $t^* < t'_1$ tal que $\|z(t^*; q)\|_\delta = M + C_F + 2$ y $\|z(t; q)\|_\delta < M + C_F + 2$ en $[0, t^*)$. Entonces, por (3.1), $\|z(t; q) - z(t; q^*)\|_\delta \leq f_q(t) \leq C_F + 1$ en $[0, t^*]$ (dado que η_q puede elegirse

igual a t^*). Luego $\|z(t^*; q)\|_\delta \leq \|z(t; q^*)\|_\delta + C_F + 1 \leq M + C_1 + 1$, lo que contradice $\|z(t^*; q)\|_\delta = M + C_F + 2$.

Hemos probado entonces, que para todo $\|q - q^*\| < \epsilon$, se verifica que $\|z(t; q)\|_\delta < M + C_F + 2$ para todo $t \in [0, t'_1]$, lo cual implica que η_q puede elegirse igual a t'_1 y en consecuencia (3.1) se cumple en $[0, t'_1]$, es decir

$$\|z(t; q) - z(t; q^*)\|_\delta \leq \|q - q^*\| \left(C_F + K C_F \frac{t_1^{1-\delta}}{1-\delta} \right)$$

en $[0, t'_1]$ para todo $q \in \mathcal{Q}$ tal que $\|q - q^*\| < \epsilon$. \square

A continuación presentamos el resultado principal de esta sección, en el cual se dan condiciones suficientes sobre F para la diferenciabilidad del mapeo $q \rightarrow z(t; q)$.

Teorema 3.2 *Supongamos que la hipótesis (F) se cumple para algún $\delta \in (0, 1)$ y que el mapeo $(q, z(\cdot)) \rightarrow F(q, \cdot, z(\cdot))$ de $\mathcal{Q} \times L^\infty(0, T : Z_\delta)$ en $L^\infty(0, T : Z)$ es derivable Fréchet en ambas variables. Supongamos también que el mapeo $(q, z(\cdot)) \rightarrow F_q(q, \cdot, z(\cdot))$ de $\mathcal{Q} \times L^\infty(0, T : Z_\delta)$ en $L^\infty(0, T : \mathcal{L}(\tilde{\mathcal{Q}}, Z))$ es localmente Lipschitz continuo con respecto a q y a $z(\cdot)$, y que el mapeo $(t, z) \rightarrow F_z(q, t, z)$ es continuo de $[0, T] \times Z_\delta$ en Z .*

Sean $z_0 \in Z_\delta$ y $q \in \mathcal{Q}$, y sea $t_1 = t_1(q, z_0)$ tal que $[0, t_1)$ es el intervalo maximal de existencia de $z(t; q)$. Entonces el mapeo $q \rightarrow z(t; q)$ es diferenciable Fréchet de $\mathcal{Q} \rightarrow L^\infty(0, t'_1 : Z_\delta)$, $\forall t'_1 \in (0, t_1)$, y para todo $h \in \tilde{\mathcal{Q}}$, $z_q(t; q)h$ es la solución del PVI lineal

$$\begin{cases} \frac{dv_h(t)}{dt} = Av_h(t) + F_z(q, t, z(t; q))v_h(t) + F_q(q, t, z(t; q))h, & t \in (0, t'_1) \\ v_h(0) = 0. \end{cases}$$

Demostración. Sean $z_0 \in Z_\delta$, $q \in \mathcal{Q}$, $t_1 = t_1(q, z_0)$ tal que $[0, t_1)$ es el intervalo maximal de existencia de $z(t; q)$, $t'_1 \in (0, t_1)$, y $\epsilon > 0$ arbitrario. Mostraremos que existe $\gamma^* > 0$ tal que si $h \in \tilde{\mathcal{Q}}$ y $\|h\| < \gamma^*$ entonces $z(t; q+h)$ existe en $[0, t'_1]$ y $\|z(t; q+h) - z(t; q) - v_h(t)\|_\delta < \epsilon\|h\|$.

Por el Teorema 3.1 existen constantes positivas γ_1 y C_1 tales que si $h \in \tilde{\mathcal{Q}}$ y $\|h\| < \gamma_1$ entonces $z(t; q+h)$ existe en $[0, t'_1]$ y $\|z(t; q+h) - z(t; q)\|_\delta \leq C_1\|h\|$ para todo t en $[0, t'_1]$.

Además, para $h \in \tilde{\mathcal{Q}}$, con $\|h\| < \gamma_1$ se tiene que

$$\begin{aligned} z(t; q+h) &= T(t)z_0 + \int_0^t T(t-s)F(q+h, s, z(s; q+h)) ds, \\ z(t; q) &= T(t)z_0 + \int_0^t T(t-s)F(q, s, z(s; q)) ds \end{aligned}$$

en $[0, t'_1]$ y también

$$v_h(t) = \int_0^t T(t-s) [F_z(q, s, z(s; q))v_h(s) + F_q(q, s, z(s; q))h] ds$$

en $[0, t'_1]$. La existencia de $v_h(\cdot)$ en $[0, t_1)$ se sigue de la hipótesis de continuidad de $F_z(q, t, z)$ y del Teorema 1.31.

Observemos ahora que

$$\begin{aligned} & z(t; q+h) - z(t; q) - v_h(t) \\ &= \int_0^t T(t-s) [F(q+h, s, z(s; q+h)) - F(q, s, z(s; q))] ds \\ &\quad - \int_0^t T(t-s) [F_z(q, s, z(s; q))v_h(s) + F_q(q, s, z(s; q))h] ds \\ &= \int_0^t T(t-s) [F(q+h, s, z(s; q+h)) - F(q+h, s, z(s; q))] ds \\ &\quad + \int_0^t T(t-s) [F(q+h, s, z(s; q)) - F(q, s, z(s; q)) - F_q(q, s, z(s; q))h] ds \\ &\quad + \int_0^t T(t-s) [F(q, s, z(s; q+h)) - F(q, s, z(s; q)) - F_z(q, s, z(s; q))v_h(s)] ds \\ &\quad + \int_0^t T(t-s) [F(q, s, z(s; q)) - F(q, s, z(s; q+h))] ds \\ &\doteq I_1 + I_2 + I_3 + I_4, \end{aligned}$$

donde I_i , $i = 1, 2, 3, 4$ denota el i -ésimo término en el orden en que aparecen arriba.

Como F es diferenciable Fréchet con respecto a z de $L^\infty(0, T : Z_\delta)$ en $L^\infty(0, T : Z)$ por hipótesis, existe $\gamma_2 > 0$ tal que si $\|z(\cdot; q+h) - z(\cdot; q)\|_{L^\infty(0, T; Z_\delta)} < \gamma_2$, entonces

$$\begin{aligned} & \|F(q, s, z(s; q+h)) - F(q, s, z(s; q)) - F_z(q, s, z(s; q))(z(s; q+h) - z(s; q))\|_Z \\ & \leq \epsilon \|z(s; q+h) - z(s; q)\|_\delta \end{aligned}$$

para todo $s \in [0, t'_1]$, y por lo tanto

$$\begin{aligned}
 \|I_3\|_\delta &\leq \int_0^t \frac{C}{(t-s)^\delta} \|F(q, s, z(s; q+h)) - F(q, s, z(s; q)) \\
 &\quad - F_z(q, s, z(s; q)) (z(s; q+h) - z(s; q))\|_Z ds \\
 &\quad + \int_0^t \frac{C}{(t-s)^\delta} \|F_z(q, s, z(s; q)) (z(s; q+h) - z(s; q) - v_h(s))\|_Z ds \\
 &\leq \int_0^t \frac{C \epsilon}{(t-s)^\delta} \|z(s; q+h) - z(s; q)\|_\delta ds \\
 &\quad + \int_0^t \frac{C C_2}{(t-s)^\delta} \|z(s; q+h) - z(s; q) - v_h(s)\|_\delta ds,
 \end{aligned}$$

donde $C_2 = \|F_z(q, \cdot, z(\cdot; q))\|_{L^\infty(0, T; \mathcal{L}(Z_\delta, Z))}$. Por lo tanto, si $\|h\| < \min\{\gamma_2/C_1, \gamma_1\}$, obtenemos

$$\|I_3\|_\delta \leq \int_0^t \frac{C C_1 \epsilon}{(t-s)^\delta} \|h\| ds + \int_0^t \frac{C C_2}{(t-s)^\delta} \|z(s; q+h) - z(s; q) - v_h(s)\|_\delta ds.$$

para todo $t \in [0, t'_1]$.

Además, como F es diferenciable Fréchet con respecto a q , existe $\gamma_3 > 0$ tal que si $\|h\| < \gamma_3$, entonces

$$\|I_2\|_\delta \leq \int_0^t \frac{C \epsilon}{(t-s)^\delta} \|h\| ds.$$

Por otro lado, observemos que $I_1 + I_4$ puede escribirse como

$$\begin{aligned}
 I_1 + I_4 &= \int_0^t T(t-s) [F(q+h, s, z(s; q+h)) - F(q, s, z(s; q+h))] ds \\
 &\quad - \int_0^t T(t-s) [F(q+h, s, z(s; q)) - F(q, s, z(s; q))] ds \\
 &= \int_0^t T(t-s) F_q(q + \alpha_1(h)h; s, z(s; q+h)) h ds \\
 &\quad - \int_0^t T(t-s) F_q(q + \alpha_2(h)h; s, z(s; q)) h ds
 \end{aligned}$$

donde $0 < \alpha_1(h), \alpha_2(h) < 1$. En consecuencia, por la continuidad Lipschitz de F_q con

respecto a q y a $z(\cdot)$ tenemos que

$$\begin{aligned}
& \|I_1 + I_4\|_\delta \\
& \leq \int_0^t \frac{C}{(t-s)^\delta} \|F_q(q + \alpha_1(h)h, s, z(s; q+h)) - F_q(q + \alpha_1(h)h, s, z(s; q))\|_Z \|h\| ds \\
& \quad + \int_0^t \frac{C}{(t-s)^\delta} \|F_q(q + \alpha_1(h)h, s, z(s; q)) - F_q(q + \alpha_2(h)h, s, z(s; q))\|_Z \|h\| ds \\
& \leq \|h\| \int_0^t \frac{C}{(t-s)^\delta} (L_1 \|z(s; q+h) - z(s; q)\|_\delta + L_2 |\alpha_1(h) - \alpha_2(h)| \|h\|) ds \\
& \leq C_3 \|h\|^2,
\end{aligned}$$

donde la última desigualdad se verifica siempre que $\|h\| < \gamma_1$.

Resumiendo, existen constantes $\gamma^* = \min\{\gamma_1, \gamma_2/C_1, \gamma_3, \epsilon\}$, $K_1, K_2 > 0$, tales que, si $h \in \tilde{\mathcal{Q}}$ y $\|h\| < \gamma^*$, entonces para todo $t \in [0, t'_1]$

$$\|z(t; q+h) - z(t; q) - v_h(t)\|_\delta \leq K_1 \|h\| \epsilon + K_2 \int_0^t \frac{\|z(s; q+h) - z(s; q) - v_h(s)\|_\delta}{(t-s)^\delta} ds.$$

Esta última desigualdad, junto con el Lema 1.33 implica la existencia de una constante \tilde{K} , independiente de ϵ , tal que para todo $t \in [0, t'_1]$

$$\|z(t; q+h) - z(t; q) - v_h(t)\|_\delta \leq \tilde{K} \|h\| \epsilon, \quad \text{si } \|h\| < \gamma^*,$$

y el teorema queda probado. □

3.2 Continuidad Lipschitz de la Derivada de Fréchet

Como vimos en la Sección 2.2, una condición suficiente para que el algoritmo de cuasilocalización converja es que exista la derivada de Fréchet del mapeo $q \rightarrow z(t; q)$, pero que además esta derivada sea Lipschitz continua. El siguiente teorema provee condiciones suficientes para que se verifique esta última hipótesis.

Teorema 3.3 *Supongamos que se cumplen las hipótesis del Teorema 3.2. Supongamos también que el mapeo $(q, z(\cdot)) \rightarrow F_z(q, \cdot, z(\cdot))$ de $\mathcal{Q} \times L^\infty(0, T : Z_\delta)$ en $L^\infty(0, T : \mathcal{L}(Z_\delta, Z))$ es localmente Lipschitz continuo con respecto a ambas variables q y $z(\cdot)$. Entonces, el mapeo $q \rightarrow z_q(\cdot; q)$ de $\mathcal{Q} \rightarrow L^\infty(0, T : \mathcal{L}(\tilde{\mathcal{Q}}, Z))$ es localmente Lipschitz continuo.*

Demostración. Por el Teorema 3.2 $z_q(t; q)h$ coincide con la solución del problema a valores iniciales

$$\begin{cases} \frac{dv_{q,h}(t)}{dt} = Av_{q,h}(t) + G(q, t, v_{q,h}, h) \\ v_{q,h}(0) = 0, \end{cases}$$

donde $G(q, t, v, h) \doteq F_z(q, t, z(t; q))v + F_q(q, t, z(t; q))h$. Ahora, sea \mathcal{Q}_C un subconjunto compacto de \mathcal{Q} and $q_1, q_2 \in \mathcal{Q}_C$. Entonces, para $v \in Z_\delta$ y $t \in [0, T]$, se sigue que

$$\begin{aligned} \|G(q_1, t, v, h) - G(q_2, t, v, h)\|_Z &\leq \|F_z(q_1, t, z(t; q_1)) - F_z(q_2, t, z(t; q_2))\|_{\mathcal{L}(Z_\delta, Z)} \|v\|_{Z_\delta} \\ &\quad + \|F_q(q_1, t, z(t; q_1)) - F_q(q_2, t, z(t; q_2))\|_{\mathcal{L}(\tilde{\mathcal{Q}}, Z)} \|h\|_{\tilde{\mathcal{Q}}} \\ &\leq L_1 \|z(t; q_1) - z(t; q_2)\|_{Z_\delta} \|v\|_{Z_\delta} + L_2 \|q_1 - q_2\|_{\tilde{\mathcal{Q}}} \|h\|_{\tilde{\mathcal{Q}}} \\ &\leq (\tilde{L}_1 \|v\|_{Z_\delta} + L_2 \|h\|_{\tilde{\mathcal{Q}}}) \|q_1 - q_2\|_{\tilde{\mathcal{Q}}}, \end{aligned}$$

donde, en la tercera desigualdad se utilizó el hecho que $z(t; q)$ es localmente Lipschitz continua con respecto a q (Teorema 3.1). Por lo tanto, el mapeo $q \rightarrow G(q, \cdot, v, h)$ es localmente Lipschitz continuo y la constante Lipschitz puede elegirse independiente de v y de h en subconjuntos compactos de Z_δ y $\tilde{\mathcal{Q}}$, respectivamente. En consecuencia, $G(q, t, v, h)$ satisface la hipótesis del Teorema 3.1 y el mapeo $q \rightarrow z_q(\cdot; q)h$ es localmente Lipschitz continuo. Más aún, como la constante Lipschitz de G es independiente de h en conjuntos $\tilde{\mathcal{Q}}$ -acotados, se sigue inmediatamente de la demostración del Teorema 3.1 que la constante Lipschitz del mapeo $q \rightarrow z_q(\cdot, q)h$ también puede elegirse independiente de h en conjuntos $\tilde{\mathcal{Q}}$ -acotados. Esto culmina la demostración del teorema. \square

Capítulo 4

Materiales con Memoria de Forma

En este capítulo presentaremos un sistema de ecuaciones en derivadas parciales que se origina en el modelado de la dinámica de un sólido unidimensional con memoria de forma. Este sistema de ecuaciones posee parámetros no-físicos que desean identificarse a partir de datos experimentales, y puede escribirse formalmente como un problema de Cauchy abstracto no lineal sobre un espacio de Hilbert determinado, donde los parámetros mencionados aparecen en la parte no lineal del mismo. Esta formulación abstracta, que presentaremos en la Sección 4.3, brinda un marco matemático adecuado para el tratamiento de diferentes problemas relacionados con el sistema mencionado, como ser la existencia, unicidad y regularidad de soluciones, dependencia de las soluciones con respecto a los parámetros, identificación de parámetros, control, etc.

El problema de identificación recientemente mencionado generó el interés en el estudio de los problemas de identificación presentados en la Sección 1.7 y fue el que motivó, a fin de cuentas, todo el presente trabajo de tesis.

4.1 Introducción

Los metales se caracterizan por sus propiedades físicas tales como conductividad térmica, maleabilidad, resistencia a la tracción, etc. Para una nueva familia de aleaciones podemos agregar las antropomórficas propiedades de *memoria* y *capacidad de aprendizaje*. Estas aleaciones exhiben lo que llamamos efecto de memoria de forma: en un cierto

rango de temperaturas pueden deformarse en forma plástica, permanecer deformados y recuperar luego su forma original sin la ayuda de ninguna fuerza externa, solamente mediante la aplicación de calor. De allí el nombre de materiales con memoria de forma o “shape memory alloys” (SMA de ahora en adelante). Durante el proceso de recuperación de su forma original, estas aleaciones pueden producir una fuerza y un desplazamiento, en función de la temperatura. También pueden ser “entrenados” para comportarse de una forma determinada ante ciertos cambios térmicos. Estas propiedades han dado origen a una amplia variedad de aplicaciones tecnológicas entre las cuales podemos citar: aplicaciones en aviación, en la construcción de ortodoncias y otros dispositivos dentales, antenas autodesplegables para satélites artificiales, cuplas para tuberías de alta presión, robótica, grampas, válvulas y retenes accionados por temperatura, interruptores térmicos, motores de calor, ventanas automáticas para invernaderos y además una amplia variedad de aplicaciones en medicina, en el tratamiento de la escoliosis, en la construcción de prótesis y dispositivos de tracción para fracturas óseas, en filtros para coágulos sanguíneos y hasta existen ya prototipos de corazones artificiales construidos de SMA (ver [4, 14, 18, 22, 34, 43, 44, 45]).

El comportamiento de los SMA depende fuertemente de la temperatura. A bajas temperaturas el comportamiento es elasto-plástico. Las curvas de fuerza deformación en este rango de temperaturas muestran una rama elástica para fuerzas pequeñas y un límite elástico o punto de deformación plástica. Sin embargo, a diferencia de lo que ocurre con un material plástico que se rompe cuando la fuerza excede ese límite, existe una segunda rama elástica, que le permite al material soportar fuerzas mayores que el límite elástico (Figura 4.1a). En este rango de temperaturas también se dice que el comportamiento es ferroelástico pues las curvas de fuerza-deformación muestran mucha similitud con las relaciones campo-magnetización observadas en ferromagnetismo. Ambas presentan, por ejemplo, el fenómeno de histéresis.

A temperaturas intermedias, el tipo de comportamiento que se observa se denomina pseudoelástico o superelástico. Aquí, un rango relativamente amplio de deformaciones

puede corresponder a una misma fuerza aplicada (Figura 4.1b). Este fenómeno de superelasticidad ha sido utilizado exitosamente en la construcción de ortodoncias y otros dispositivos dentales, ya que una determinada fuerza correctora se puede conseguir para un amplio rango de deformaciones, sin necesidad de tener que readaptar el material.

A temperaturas altas, el material se comporta en forma casi linealmente elástica, con mayor módulo de elasticidad (pendiente de la curva fuerza deformación) cuanto mayor es la temperatura (Figura 4.1c).

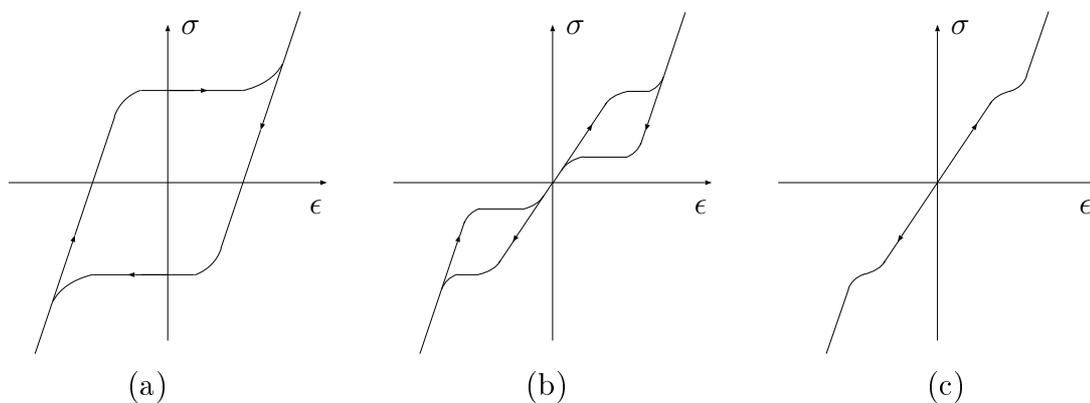


Figura 4.1: Esquemas de las relaciones fuerza–deformación (σ – ϵ) para materiales con memoria a diferentes temperaturas: (a) bajas; (b) intermedias; (c) altas.

Desde el punto de vista microscópico, todos los fenómenos peculiares que caracterizan a los SMA están asociados a ciertos cambios en su estructura cristalina que originan las llamadas “transiciones de fase” de tipo sólido–sólido o transformaciones martensíticas. En el artículo [4] de M. Ahlers se encuentra un detallado y muy claro resumen de las propiedades cristalográficas de los SMA. Algunas de las aleaciones que exhiben las características que acabamos de mencionar son, entre otras, AgCd, AuCd, CuAlNi, CuAuZn, CuSn, NiAl, NiTi.

4.2 Modelos Matemáticos

Los primeros intentos por modelar matemáticamente la dinámica de los SMA se remontan al final de la década del 60 con los trabajos de M. E. Gurtin [27]. Desde entonces y hasta

el año 1986 se propusieron y estudiaron varios modelos matemáticos [1, 2, 3, 24, 25, 33, 35, 36, 37, 52].

La mayoría de estos modelos, sin embargo, eran estáticos y no tenían en cuenta el fuerte acoplamiento entre las propiedades térmicas y mecánicas, que caracterizan a los SMA, y algunos de ellos trataron simplemente el problema de encontrar funciones que ajustaran de manera apropiada las extrañas relaciones de fuerza–deformación, y que al mismo tiempo fuesen capaces de describir y explicar algunos de los fenómenos observados experimentalmente. Sin embargo, todos estos modelos suponían que la temperatura, o la fuerza aplicada, eran constantes, los procesos se trataban en forma puntual en el espacio y no permitían la inclusión de acciones externas variables. Estas limitaciones hicieron que estos modelos resultaran de muy poco interés práctico y pronto fueron abandonados.

En 1988 M. Niezgodka y J. Sprekels [38] propusieron un modelo general basado en las leyes de conservación de energía y momento lineal para estos materiales. Esto constituyó un gran paso hacia adelante puesto que este modelo no solamente subsanaba todas las deficiencias de los modelos anteriores, sino que, además, podía incorporar gran parte de la información que ellos brindaban. A continuación describiremos este modelo desde sus orígenes.

Consideraremos sólo el caso unidimensional. Supongamos que tenemos una barra o alambre de longitud unitaria $\Omega = (0, 1)$ construida con un material con memoria de forma. Denotemos con $u(x, t)$ al desplazamiento transversal y con $\theta(x, t)$ la temperatura absoluta del punto x en el instante t . Permitiremos la inclusión de fuentes distribuidas de calor $g(x, t)$ y también fuerzas distribuidas $f(x, t)$, las que pueden ser utilizadas como variables de control. Con ρ , k y β denotamos constantes positivas (ρ es la densidad de masa, k es el coeficiente de conductividad térmica y β es la viscosidad). Teniendo en cuenta relaciones constitutivas y algunas ecuaciones termomecánicas bien conocidas, las leyes de conservación de momento lineal y energía para Ω dan origen al siguiente sistema

de ecuaciones (ver [46] para una deducción detallada del mismo):

$$\begin{cases} \rho u_{tt}(x, t) - \beta \rho u_{xxt}(x, t) - [\Psi_\epsilon(\epsilon, \epsilon_x, \theta)]_x + [\Psi_{\epsilon_x}(\epsilon, \epsilon_x, \theta)]_{xx} = f(x, t) \\ -\theta [\Psi_\theta(\epsilon, \epsilon_x, \theta)]_t - k\theta_{xx}(x, t) - \beta \rho u_{xt}(x, t)^2 = g(x, t) \\ \epsilon = u_x(x, t) \end{cases} \quad (4.1)$$

para $x \in \Omega$ y $t > 0$, donde los subíndices x y t denotan derivadas. Aquí $\epsilon = u_x$ es lo que se llama strain tangencial linealizado y la función $\Psi(\epsilon, \epsilon_x, \theta)$ es el potencial de energía libre de Helmholtz.

Supongamos por el momento que las constantes ρ , k y β son conocidas y que también lo son las funciones $f(x, t)$ y $g(x, t)$. ¿Es posible encontrar un par de funciones $u(x, t)$, $\theta(x, t)$ que satisfagan el sistema de ecuaciones (4.1)? Resulta claro que antes de poder contestar a esta pregunta necesitaremos, al menos

(i) conocer la forma explícita del funcional $\Psi(\epsilon, \epsilon_x, \theta)$;

(ii) e imponer condiciones iniciales y de frontera relevantes

para poder recién comenzar a intentar resolver el sistema de ecuaciones diferenciales en derivadas parciales resultante. El problema (i) está íntimamente relacionado con las relaciones fuerza–deformación (o stress–strain) que analizamos al principio y con una teoría de transiciones de fase llamada teoría de Landau [23] que dice, entre otras cosas, qué condiciones debe satisfacer el funcional Ψ . Es aquí donde los modelos isotérmicos mencionados anteriormente juegan un rol importante. Dado que el stress σ puede definirse como la derivada parcial del potencial de energía libre con respecto al strain ϵ , el potencial Ψ debe ser tal que, para los diferentes rangos de temperatura, las curvas ϵ – Ψ_ϵ muestren un comportamiento similar al que exhiben las curvas fuerza–deformación obtenidas experimentalmente. Para el caso de materiales perfectamente elásticos se tiene que el stress es proporcional al strain: $\sigma = c\epsilon$ (Ley de Hooke) donde c es una constante llamada módulo de elasticidad o módulo de Young. En el caso de materiales linealmente termoelásticos, el módulo de elasticidad depende de la temperatura: $\sigma = c(\theta)\epsilon$. Hablamos de elasticidad

(termoelasticidad) no lineal cuando σ es una función no lineal de ϵ (y de θ), monótona creciente con respecto a ϵ para cada θ fijo. En los casos de elasticidad lineal, los potenciales de energía asociados son de tipo cuadrático y, en general, convexos para la gran mayoría de los metales. Para el caso de los materiales con memoria de forma, la histéresis y la termo-elasto-plasticidad que caracterizan las relaciones fuerza–deformación a temperaturas bajas e intermedias se traduce en potenciales de energía no convexos (“two well potentials”). En este aspecto han sido muy importantes los trabajos de F. Falk ([24], [25]), quien descubrió que un potencial de energía simple que satisface todas las hipótesis de la teoría de Landau y al mismo tiempo es capaz de reproducir muchos de los fenómenos observados experimentalmente, está dado por el llamado potencial de Landau-Ginzburg:

$$\Psi(\epsilon, \epsilon_x, \theta) = -C_v \theta \log\left(\frac{\theta}{\theta_2}\right) + C_v \theta + C + \alpha_2(\theta - \theta_1)\epsilon^2 - \alpha_4 \epsilon^4 + \alpha_6 \epsilon^6 + \frac{\gamma}{2} \epsilon_x^2, \quad (4.2)$$

donde C_v , θ_1 , θ_2 , α_2 , α_4 , α_6 , γ y C son constantes positivas dependientes del material considerado θ_1 y θ_2 son dos temperaturas críticas y C_v denota el calor específico; las constantes α_2 , α_4 , α_6 , son constantes **no-físicas**, la constante C es un nivel de energía de referencia fijo y γ denota la rigidez a la torsión (bending stiffness). Observar que para valores de θ cerca de θ_1 , Ψ es una función no convexa de ϵ y que si en particular $\alpha_4 = \alpha_6 = 0$ estamos en un caso de termoelasticidad lineal. En el mismo trabajo, Falk propone —para una aleación particular— ciertos valores para los parámetros para los cuales las curvas $\epsilon - \Psi_\epsilon$ muestran gran similitud con las de la Figura 4.1 (ver Figura 4.2).

Con el potencial Ψ dado por (4.2), el sistema (4.1) toma la forma

$$\begin{cases} \rho u_{tt} - \beta \rho u_{xxt} + \gamma u_{xxxx} = f(x, t) + [2\alpha_2(\theta - \theta_1)u_x - 4\alpha_4 u_x^3 + 6\alpha_6 u_x^6]_x \\ C_v \theta_t - k \theta_{xx} = g(x, t) + 2\alpha_2 \theta u_x u_{xt} + \beta \rho u_{xt}^2, \end{cases} \quad (4.3)$$

para $x \in \Omega$, $t > 0$. A este sistema se lo conoce frecuentemente como ecuaciones de termo-visco-elasto-plasticidad unidimensional. La primera ecuación en (4.3) es no lineal e hiperbólica en $u(x, t)$ y puede entenderse como una ecuación de viga no lineal. La segunda es una ecuación de difusión del calor con un término fuente y dos términos no lineales (uno de acople); esta ecuación es parabólica en $\theta(x, t)$. Las ecuaciones están acopladas debido a

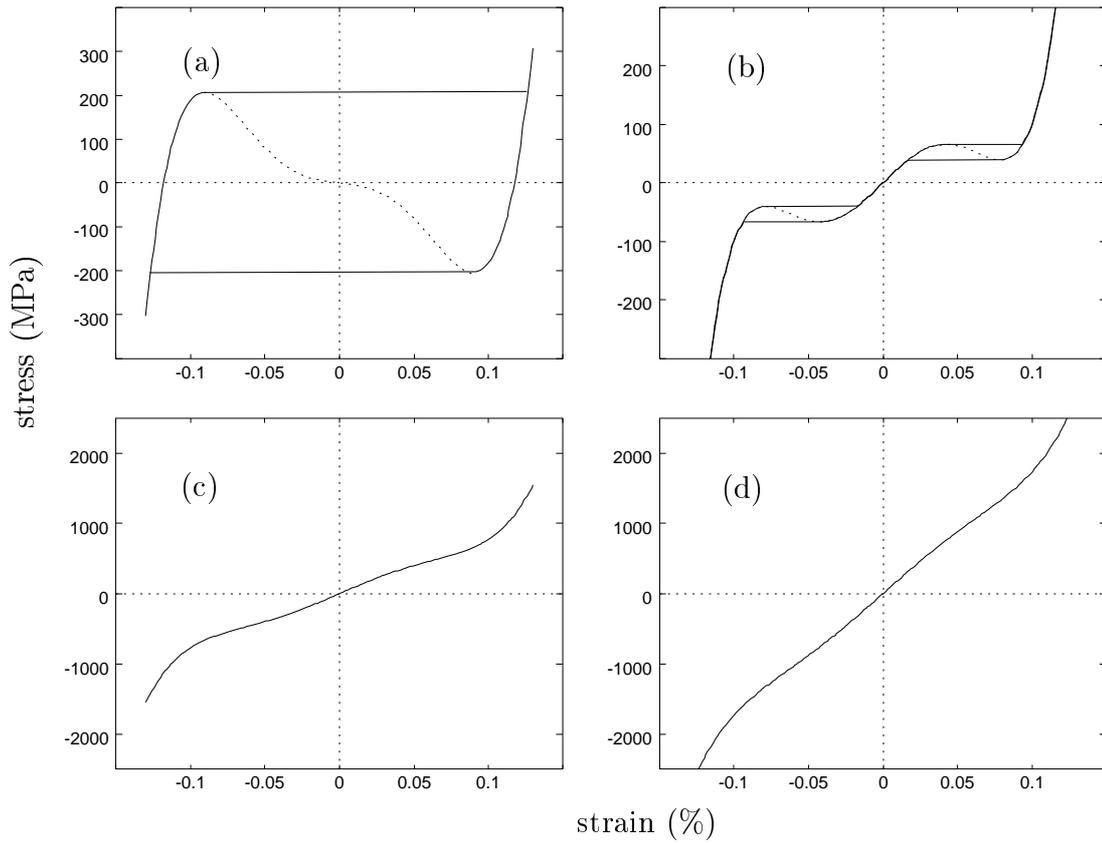


Figura 4.2: Gráfico de las curvas $\epsilon-\Psi_\epsilon$ para diferentes temperaturas: (a) $\theta = 200^0\text{K}$; (b) $\theta = 260^0\text{K}$; (c) $\theta = 400^0\text{K}$; (d) $\theta = 600^0\text{K}$. Obtenidas con Ψ como en (4.2) con los valores de α_2 , α_4 , α_6 y θ_1 tomados de [24]. Las líneas punteadas representan partes inestables de las curvas. Las líneas horizontales representan posibles lazos de histéresis.

la presencia de los términos que contienen simultáneamente $\theta(x, t)$ y derivadas parciales de $u(x, t)$. El sistema es, por lo tanto, un sistema no lineal, acoplado, de tipo mixto, hiperbólico-parabólico de ecuaciones en derivadas parciales.

Los primeros trabajos sobre existencia y unicidad de soluciones para sistemas del tipo (4.3) se deben a M. Niezgodka, J. Sprekels, K. H. Hoffmann y S. Zheng [38, 30, 31, 40]. En estos trabajos, sin embargo, se supone que el flujo de calor está dado por la ecuación $q = -k\theta_x - \alpha k\theta_{xt}$, en lugar de la clásica ley de Fourier (la misma ecuación con $\alpha = 0$), donde k es el coeficiente de conductividad térmica y $\alpha > 0$ es una constante. Esta ecuación caracteriza la conducción de calor con una pequeña memoria térmica y tiene su origen en algunos artículos de P. J. Chen, M. E. Gurtin y J. Nunziato [19, 27, 41]. Al considerar el flujo de calor en esta forma aparece el término adicional $-\alpha k\theta_{xxt}$ en el lado izquierdo de la segunda ecuación en (4.3), lo que trae ciertas ventajas en lo que se refiere a la demostración de la existencia y unicidad de soluciones del problema (ver [38, 40]). Sin embargo, lo cierto es que, sin mayores dificultades, puede demostrarse que si el coeficiente α es estrictamente positivo, no se verifica la segunda ley de la termodinámica para la producción de entropía, dada por la desigualdad de Clausius-Duhem. De allí que los resultados obtenidos en los trabajos arriba mencionados, no obstante matemáticamente interesantes, sean, al menos, físicamente cuestionables.

Los primeros trabajos sobre la existencia de soluciones para el caso $\alpha = 0$ se deben a J. Sprekels [49] y S. Zheng [50, 53]. Una breve síntesis sobre la evolución de los resultados en este área puede encontrarse en [46]. Nosotros seguiremos, para este problema, la línea de trabajo introducida en [46] donde se escriben las ecuaciones como un problema de Cauchy abstracto en un espacio de Hilbert apropiado.

Dijimos anteriormente que para poder analizar la existencia de soluciones del problema debíamos, además de conocer la forma del funcional $\Psi(\epsilon, \epsilon_x, \theta)$, imponer condiciones

iniciales y de frontera relevantes. Consideraremos las siguientes:

$$\begin{cases} u(x, 0) = u_0(x), & u_t(x, 0) = v_0(x), & \theta(x, 0) = \theta_0(x); \\ u(0, t) = u(1, t) = 0 = u_{xx}(0, t) = u_{xx}(1, t); \\ \theta_x(0, t) = \theta_x(1, t) = 0, \end{cases} \quad (4.4)$$

de manera que la barra se encuentra fija (“simply supported”) y térmicamente aislada en ambos extremos. Las condiciones iniciales indican que el desplazamiento, la velocidad y la temperatura de la barra se suponen conocidas en $t = 0$.

De todas las constantes intervinientes en el sistema (4.3), los parámetros α_2 , α_4 , α_6 que definen el potencial Ψ son no físicos y por lo tanto no pueden determinarse a partir de experimentos de laboratorio. Desde un punto de vista práctico y de las aplicaciones, sería muy interesante encontrar los valores de estos parámetros, como así también de la temperatura crítica θ_1 para un cierto material. Más adelante en el Capítulo 6 utilizaremos el algoritmo de cuasilinealización presentado en el Capítulo 2 para la identificación de estos cuatro parámetros.

4.3 Formulación Abstracta del Problema

En esta sección, siguiendo el enfoque introducido en [46], presentaremos una formulación abstracta del problema (4.3)–(4.4) que ya ha sido utilizada para la demostración de existencia y unicidad de soluciones locales del problema y que además brinda un marco matemático adecuado para atacar el problema de la identificación de parámetros, dado que la forma del problema abstracto coincide con la ecuación (1.11) para la cual se estudió el problema de identificación en el Capítulo 2.

Teniendo en cuenta que uno de nuestros objetivos es el de identificar los parámetros α_2 , α_4 , α_6 y θ_1 , y sabiendo de antemano que son positivos, definimos el *conjunto de parámetros admisibles* $\mathcal{Q} = \{q = (\alpha_2, \alpha_4, \alpha_6, \theta_1) \mid q \in \mathbb{R}_{>0}^4\}$, y definimos el espacio de estados Z como

el espacio de Hilbert $H_0^1(0,1) \cap H^2(0,1) \times L^2(0,1) \times L^2(0,1)$ con el producto interno

$$\left\langle \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}, \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\theta} \end{pmatrix} \right\rangle = \gamma \int_0^1 u''(x) \overline{\tilde{u}''(x)} dx + \rho \int_0^1 v(x) \overline{\tilde{v}(x)} dx + \frac{C_v}{k} \int_0^1 \theta(x) \overline{\tilde{\theta}(x)} dx.$$

Denotaremos con $\|\cdot\|$ a la norma inducida por $\langle \cdot, \cdot \rangle$ en Z .

Con esta notación, el problema de valores iniciales y de frontera (4.3)–(4.4) puede escribirse formalmente como el siguiente problema de Cauchy abstracto en Z

$$\begin{cases} \frac{d}{dt} z(t) = Az(t) + F(q, t, z(t)), & 0 \leq t \leq T \\ z(0) = z_0, \end{cases} \quad (4.5)$$

donde

$$z(t)(x) = \begin{pmatrix} u(x, t) \\ u_t(x, t) \\ \theta(x, t) \end{pmatrix}, \quad z_0(x) = \begin{pmatrix} u_0(x) \\ v_0(x) \\ \theta_0(x) \end{pmatrix},$$

$$D(A) = \left\{ \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} \in Z \mid \begin{array}{l} u \in H^4(0,1), u(0) = u(1) = 0 = u''(0) = u''(1), \\ v \in H_0^1(0,1) \cap H^2(0,1), \\ \theta \in H^2(0,1), \theta'(0) = \theta'(1) = 0 \end{array} \right\} \quad (4.6)$$

y para $\begin{pmatrix} u \\ v \\ \theta \end{pmatrix} \in D(A)$

$$A \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} = \begin{pmatrix} v \\ \beta v'' - \frac{\gamma}{\rho} u'''' \\ \frac{k}{C_v} \theta'' \end{pmatrix} = \begin{pmatrix} 0 & I & 0 \\ -\frac{\gamma}{\rho} \frac{\partial^4}{\partial x^4} & \beta \frac{\partial^2}{\partial x^2} & 0 \\ 0 & 0 & \frac{k}{C_v} \frac{\partial^2}{\partial x^2} \end{pmatrix} \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}. \quad (4.7)$$

El término no lineal $F : \mathcal{Q} \times [0, T] \times D(F) \rightarrow Z$ se define por $D(F) = H^2(0,1) \times H^1(0,1) \times H^1(0,1)$ y

$$F(q, t, z) = F\left(q, t, \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}\right) = \begin{pmatrix} 0 \\ f_2(q, t, z) \\ f_3(q, t, z) \end{pmatrix} \quad (4.8)$$

donde

$$\rho f_2(q, t, \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}) = f(x, t) + \frac{d}{dx} [2\alpha_2 (\theta(x) - \theta_1) u'(x) - 4\alpha_4 u'(x)^3 + 6\alpha_6 u'(x)^5], \quad (4.9)$$

$$C_v f_3(q, t, \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}) = g(x, t) + 2\alpha_2 \theta(x) u'(x) v'(x) + \beta \rho v'(x)^2. \quad (4.10)$$

4.4 Resultados Previos

El siguiente resultado puede derivarse fácilmente a partir de los Teoremas 3.7 y 3.11 de [46] con mínimas modificaciones debidas a las condiciones de frontera diferentes que se consideran aquí.

Teorema 4.1 *Sea A el operador lineal definido por (4.6)–(4.7). Entonces*

(i) *El operador A es disipativo;*

(ii) *El operador adjunto A^* de A es también disipativo, $D(A^*) = D(A)$, y para $\begin{pmatrix} u \\ v \\ \theta \end{pmatrix} \in D(A^*)$,*

$$A^* \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} = \begin{pmatrix} -v \\ \beta v'' + \frac{\gamma}{\rho} u'''' \\ \frac{k}{C_v} \theta'' \end{pmatrix} = \begin{pmatrix} 0 & -I & 0 \\ \frac{\gamma}{\rho} \frac{\partial^4}{\partial x^4} & \beta \frac{\partial^2}{\partial x^2} & 0 \\ 0 & 0 & \frac{k}{C_v} \frac{\partial^2}{\partial x^2} \end{pmatrix} \begin{pmatrix} u \\ v \\ \theta \end{pmatrix};$$

(iii) *El operador A tiene espectro puramente puntual $\sigma_p(A)$ dado por*

$$\sigma_p(A) = \{\lambda_n^+\}_{n=1}^\infty \cup \{\lambda_n^-\}_{n=1}^\infty \cup \{\alpha_n\}_{n=0}^\infty,$$

donde

$$\lambda_n^{+,-} = \sqrt{\mu_n} \left(-r \pm \sqrt{r^2 - 1} \right), \quad \alpha_n = -\frac{k}{C_v} n^2 \pi^2$$

y $\mu_n = \gamma n^2 \pi^2 / \rho$, $r = \beta \sqrt{\rho} / 2 \sqrt{\gamma}$. Los correspondientes autovectores normalizados en Z son, respectivamente,

$$\begin{pmatrix} e_n(x) \\ \lambda_n^+ e_n(x) \\ 0 \end{pmatrix}_{n=1,2,\dots}, \quad \begin{pmatrix} k_n e_n(x) \\ k_n \lambda_n^- e_n(x) \\ 0 \end{pmatrix}_{n=1,2,\dots}, \quad \begin{pmatrix} 0 \\ 0 \\ \chi_n(x) \end{pmatrix}_{n=0,1,\dots},$$

donde

$$k_n^2 = \frac{\mu_n + |\lambda_n^+|^2}{\mu_n + |\lambda_n^-|^2}, \quad e_n(x) = \left[\frac{2}{\rho(\mu_n + |\lambda_n^+|^2)} \right]^{1/2} \text{sen}(\pi n x), \quad n = 1, 2, \dots,$$

$$\chi_0(x) = \left(\frac{k}{C_v} \right)^{1/2}, \quad \chi_n(x) = \left(\frac{2k}{C_v} \right)^{1/2} \cos(\pi n x), \quad n = 1, 2, \dots;$$

(iv) *El operador A genera un semigrupo analítico de contracciones $T(t)$ sobre Z .*

Del Teorema 4.1 se sigue que $\{\lambda \in \mathbb{C} : \operatorname{Re}(\lambda) > 0\} \subset \rho(A)$. Luego, las potencias fraccionarias $(I - A)^\delta$ de $I - A$ son operadores lineales cerrados e inversibles con inversa acotada para cualquier $\delta \in [0, 1]$ (ver Sección 1.5). Por lo tanto, $D(I - A)^\delta$ con la norma del gráfico de $(I - A)^\delta$ es un espacio de Banach y como $(I - A)^\delta$ es inversible con inversa acotada, la norma $\|z\|_\delta \doteq \|(I - A)^\delta z\|$ es equivalente a la norma del gráfico. Llamaremos Z_δ al espacio $(D(I - A)^\delta, \|\cdot\|_\delta)$.

Sobre las entradas del sistema (fuerzas y fuentes de calor distribuidas) consideraremos la siguiente hipótesis de regularidad:

(H1) *Existen funciones nonegativas $K_f, K_g \in L^2(0, 1)$, tales que*

$$|f(x, t_1) - f(x, t_2)| \leq K_f(x) |t_1 - t_2| \quad y \quad |g(x, t_1) - g(x, t_2)| \leq K_g(x) |t_1 - t_2|$$

para casi todo $x \in (0, 1)$ y todo $t_1, t_2 \in [0, T]$.

Con esta hipótesis, se tiene el siguiente resultado de regularidad sobre la parte no lineal del PVI (4.5) [47, Teorema 3.1].

Teorema 4.2 *Supongamos que se cumple la hipótesis (H1). Sean $q \in \mathcal{Q}$, $\epsilon > 0$ y U un conjunto acotado de $[0, T] \times Z_{\frac{3}{4}+\epsilon}$. Entonces existe una constante $L = L(q, U, f, g)$ tal que*

$$\|F(q, t_1, z_1) - F(q, t_2, z_2)\| \leq L(|t_1 - t_2| + \|z_1 - z_2\|_{\frac{3}{4}+\epsilon})$$

para todo $(t_1, z_1), (t_2, z_2) \in U$, es decir, la función $F : \mathcal{Q} \times [0, T] \times Z_{\frac{3}{4}+\delta} \rightarrow Z$ es localmente Lipschitz continua en t y en z . Más aún, la constante L puede elegirse independiente de q en cualquier subconjunto compacto de \mathcal{Q} .

En la demostración de este teorema [47, Teorema 3.1] se demuestra, utilizando resultados clásicos de teoría de interpolación (ver por ejemplo [32] o [51]) la siguiente caracterización de los espacios Z_δ que aquí presentamos en el siguiente

Corolario 4.3 Si $\frac{3}{4} < \delta \leq 1$, entonces

$$Z_\delta = \left\{ \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} \in H^{2+2\delta} \times H^{2\delta} \times H^{2\delta} \left| \begin{array}{l} u(0) = u(1) = u''(0) = u''(1) = 0 \\ v(0) = v(1) = 0 \\ \theta'(0) = \theta'(1) = 0 \end{array} \right. \right\}$$

y la norma $\|\cdot\|_\delta$ es equivalente a la norma del espacio producto $H^{2+2\delta} \times H^{2\delta} \times H^{2\delta}$.

El Teorema 4.2 provee condiciones suficientes para que, dados $\delta > \frac{3}{4}$ y $z_0 \in D(I - A)^\delta$, exista $t_1 > 0$ tal que el mapeo $\mathcal{G} : C([0, t_1] : Z) \rightarrow C([0, t_1] : Z)$, dado por

$$\mathcal{G}y(t) = T(t)(I - A)^\delta z_0 + \int_0^t (I - A)^\delta T(t - s)F(q, s, (I - A)^{-\delta}y(s)) ds$$

tenga un punto fijo $y^*(t) = y^*(t; q, z_0)$. Resulta luego que $(I - A)^{-\delta}y^*(t)$ es una solución de (4.5) en $[0, t_1]$. Más precisamente, como consecuencia del Teorema 6.3.1 y del Corolario 6.3.2 en [42], se tiene el siguiente resultado de existencia y regularidad.

Teorema 4.4 Sea $q \in \mathcal{Q}$, $\delta > \frac{3}{4}$ y supongamos que se cumple **(H1)**. Entonces, para todo dato inicial $z_0 \in D(I - A)^\delta$ existe $t_1 = t_1(q, z_0)$ tal que el problema de valores iniciales (4.5) tiene una única solución $z(t; q, z_0) \in C([0, t_1] : Z_\delta) \cap C^1((0, t_1) : Z)$. Más aún, $\frac{d}{dt}z(t; q, z_0) \in C_{loc}^{1-\delta}((0, t_1) : Z)$, es decir, $\frac{d}{dt}z(t; q, z_0)$ es localmente Hölder continua en $(0, t_1)$ con exponente $1 - \delta$. Además, si \mathcal{Q}_C es un subconjunto compacto del conjunto de parámetros admisibles \mathcal{Q} , entonces $t_1^* = \inf \{t_1(q, z_0) : q \in \mathcal{Q}_C\} > 0$, es decir, existe un tiempo final $t_1^* > 0$ tal que la solución $z(t; q)$ existe en $[0, t_1^*)$ para todo $q \in \mathcal{Q}_C$.

Capítulo 5

Aproximaciones Numéricas

En este capítulo definiremos aproximaciones de dimensión finita al problema de Cauchy abstracto (4.5), demostraremos la convergencia de las soluciones de los problemas discretos a la solución del problema (4.5) y utilizaremos las mismas para encontrar aproximaciones numéricas a las soluciones de (4.3)–(4.4) para distintas condiciones iniciales y acciones externas. Las aproximaciones se basarán en la utilización de los autovalores y autofunciones de la parte lineal de la ecuación (4.5).

A lo largo de este capítulo, el parámetro $q \in \mathcal{Q}$ estará fijo, así que, siempre que quede claro en el contexto, será suprimido de la notación.

5.1 Aproximaciones Modales

Definimos, para $N \in \mathbb{N}$ fijo

$$\beta_n^N(x) \doteq \begin{pmatrix} \sin \pi n x \\ \lambda_n^+ \sin \pi n x \\ 0 \end{pmatrix}, \quad \beta_{N+n}^N(x) \doteq \begin{pmatrix} \sin \pi n x \\ \lambda_n^- \sin \pi n x \\ 0 \end{pmatrix}, \quad \beta_{2N+n}^N(x) \doteq \begin{pmatrix} 0 \\ 0 \\ \cos \pi(n-1)x \end{pmatrix},$$

para $n = 1, 2, \dots, N$, donde $\lambda_n^{+,-}$ son como en el Teorema 4.1, y definimos Z^N como el espacio generado por $\hat{\beta}_N \doteq \{\beta_n^N(x)\}_{n=1}^{3N}$ con la norma heredada de Z . Luego, $\bigcup_{N=1}^{\infty} Z^N$ es denso en Z y, como las β_n^N son autofunciones de A , se sigue que Z^N es invariante bajo A . Notemos también que con el producto interno heredado de Z , Z^N es un espacio de Hilbert.

A continuación, definimos el problema en dimensión finita aproximante como sigue

$$\begin{cases} \frac{d}{dt} z^N(t) = A^N z^N(t) + F^N(t, z^N(t)), & 0 \leq t \leq T \\ z^N(0) = P^N z_0 \end{cases} \quad (5.1)$$

donde $P^N : Z \rightarrow Z^N$ es la proyección ortogonal de Z sobre Z^N , A^N es la restricción de A a Z^N y $F^N(t, z) \doteq P^N F(t, z)$. La densidad de $\bigcup_{N=1}^{\infty} Z^N$ en Z implica la convergencia fuerte de P^N a la identidad. Más aún, un cálculo sencillo utilizando la descomposición espectral de A muestra que $\bigcup_{N=1}^{\infty} Z^N$ es también denso en Z_δ y $\|P^N z - z\|_\delta \rightarrow 0$ para todo $z \in Z_\delta$ y $0 < \delta \leq 1$.

Como Z^N es de dimensión finita, el operador A^N es lineal y acotado sobre Z^N , y por ende genera un semigrupo uniformemente continuo de operadores lineales y acotados $T^N(t)$ sobre Z^N .

El siguiente resultado trata sobre la existencia local de soluciones del problema aproximante (5.1).

Teorema 5.1 *Sea $z_0 \in Z$. Entonces, para todo entero positivo N , existe $t_1^N > 0$ tal que (5.1) tiene una única solución en $[0, t_1^N)$.*

Demostración. Sea $\delta \in (\frac{3}{4}, 1)$, $z_0 \in Z$ y $N \in \mathbb{N}$ fijos. Por el Teorema 4.2, para todo $r > 0$, $t' > 0$, existe una constante $L(r, t')$ tal que

$$\|F(t, q) - F(s, w)\|_Z \leq L(r, t')(|t - s| + \|z - w\|_\delta)$$

para todo $t, s \in [0, t']$ y $z, w \in Z_\delta$ con $\|z\|_\delta \leq r$, $\|w\|_\delta \leq r$. Luego, para todo $t, s \in [0, t']$ y $z, w \in Z_\delta$ con $\|z\|_\delta \leq r$, $\|w\|_\delta \leq r$ se tiene que

$$\begin{aligned} \|F^N(t, z) - F^N(s, w)\|_{Z^N} &= \|P^N(F(t, z) - F(s, w))\|_Z \\ &\leq \|F(t, z) - F(s, w)\|_Z \\ &\leq L(r, t')(|t - s| + \|z - w\|_\delta) \\ &\leq L(r, t')C(\delta, N)(|t - s| + \|z - w\|_{Z^N}) \end{aligned}$$

donde la constante $C(\delta, N)$ aparece como consecuencia de la equivalencia de normas en Z^N .

Concluyendo, el mapeo $(t, z) \rightarrow A^N z + F^N(t, z)$ es localmente Lipschitz continuo de $[0, T] \times Z^N$ en Z^N y por lo tanto existe $t_1^N > 0$ tal que el problema (5.1) tiene solución única en $[0, t_1^N)$. \square

Como es de esperarse, los semigrupos $T(t)$ y $T^N(t)$ están fuertemente relacionados, esta relación se describe en el siguiente

Lema 5.2 Sean $T(t)$, $T^N(t)$, A , A^N , Z , y Z^N como se definieron anteriormente y denotemos con $R(\lambda, A) \doteq (\lambda I - A)^{-1}$ a la resolvente de A en λ . Entonces

- (i) para todo $\lambda \in \rho(A)$, el espacio Z^N es invariante bajo $R(\lambda, A)$;
- (ii) la restricción de $T(t)$ a Z^N coincide con $T^N(t)$ para todo $t \geq 0$, es decir

$$T(t)|_{Z^N} = T^N(t) \quad \text{para todo } t \geq 0.$$

Demostración.

- (i) Sea $\lambda \in \rho(A)$ y ξ un elemento de la base $\hat{\beta}_N$ de Z^N y definamos $z \doteq R(\lambda; A)\xi$. Luego z es un autovector de A correspondiente al mismo autovalor σ de ξ . En efecto, $(\lambda I - A)Az = A\xi = \sigma\xi$, lo que implica $Az = R(\lambda; A)\sigma\xi = \sigma z$. Como todos los autovalores de A son simples, z debe ser un múltiplo de ξ y por lo tanto $z \in Z^N$. Como $R(\lambda; A)$ es lineal (i) está probado.

- (ii) Como Z^N es invariante bajo A , el operador \tilde{A}^N , la parte de A en Z^N , definido por

$$D(\tilde{A}^N) \doteq \{z \in D(A) \cap Z^N : Az \in Z^N\}$$

$$A^N z \doteq Az, \quad z \in D(\tilde{A}^N)$$

coincide con A^N , la restricción de A a Z^N . Luego, \tilde{A}^N genera un semigrupo uniformemente continuo en Z^N y por la parte (i), Z^N es invariante bajo $R(\lambda; A)$ para

todo $\lambda \in \rho(A)$. Por el Teorema 4.5.5 en [42] se sigue que $\tilde{A}^N = A^N$ es el generador infinitesimal de la restricción de $T(t)$, el semigrupo generado por A , a Z^N . \square

El siguiente teorema es el resultado principal de esta sección y trata sobre la convergencia de las soluciones de (5.1) a la solución de (4.5).

Teorema 5.3 *Sea $\delta \in (\frac{3}{4}, 1)$, $z_0 \in Z_\delta$ y supongamos que $z^N(t)$, $z(t)$ son soluciones de (5.1) y (4.5), respectivamente, y sea $[0, t_1)$ el intervalo maximal de existencia de $z(t)$. Entonces, para todo $t'_1 < t_1$ existe una constante N_0 tal que $z^N(t)$ existe en $[0, t'_1]$ para todo $N \geq N_0$ y $z^N(t)$ converge a $z(t)$ en la norma de Z para todo $t \in [0, t'_1]$. Más aún, la convergencia se logra en la norma de Z_δ .*

Demostración. Sea $\delta \in (\frac{3}{4}, 1)$, $z_0 \in Z_\delta$, $t'_1 < t_1$ y para cada $N \in \mathbb{N}$ sea $t_1^N > 0$ tal que $z^N(t)$ existe en $[0, t_1^N)$. Entonces, para $t \in [0, \min\{t'_1, t_1^N\})$ y $N \in \mathbb{N}$

$$\begin{aligned} z^N(t) &= T^N(t)P^N z_0 + \int_0^t T^N(t-s)P^N F(s, z^N(s)) ds \\ z(t) &= T(t)z_0 + \int_0^t T(t-s)F(s, z(s)) ds. \end{aligned}$$

Luego,

$$\begin{aligned} \|z^N(t) - z(t)\|_\delta &= \|(I - A)^\delta (z^N(t) - z(t))\|_Z \\ &\leq \|(I - A)^\delta (T^N(t)P^N z_0 - T(t)z_0)\|_Z \\ &\quad + \int_0^t \|(I - A)^\delta [T^N(t-s)P^N F(s, z^N(s)) - T(t-s)F(s, z(s))]\|_Z ds \\ &\doteq \rho_1^N(t) + \rho_2^N(t). \end{aligned}$$

Como $T(t)$ conmuta con $(I - A)^\delta$ y como $T^N(t)$ es la restricción de $T(t)$ a Z^N

$$\begin{aligned} \rho_1^N(t) &= \|(I - A)^\delta T(t) (P^N z_0 - z_0)\|_Z \\ &\leq \|T(t)\|_{\mathcal{L}(Z)} \|(I - A)^\delta (P^N z_0 - z_0)\|_Z. \end{aligned}$$

Del mismo modo, para el integrando de $\rho_2^N(t)$ tenemos que

$$\begin{aligned}
& \|(I - A)^\delta [T^N(t - s)P^N F(s, z^N(s)) - T(t - s)F(s, z(s))]\|_Z \\
&= \|(I - A)^\delta T(t - s) [P^N F(s, z^N(s)) - F(s, z(s))]\|_Z \\
&\leq \|(I - A)^\delta T(t - s)\|_{\mathcal{L}(Z)} \|P^N F(s, z^N(s)) - F(s, z(s))\|_Z \\
&\leq \frac{C_\delta e^{t_1}}{(t - s)^\delta} \|P^N F(s, z^N(s)) - F(s, z(s))\|_Z.
\end{aligned} \tag{5.2}$$

Pero

$$\begin{aligned}
\|P^N F(s, z^N(s)) - F(s, z(s))\|_Z &\leq \|P^N [F(s, z^N(s)) - F(s, z(s))]\|_Z \\
&\quad + \|P^N F(s, z(s)) - F(s, z(s))\|_Z \\
&\leq \|F(s, z^N(s)) - F(s, z(s))\|_Z \\
&\quad + \|(P^N - I)F(s, z(s))\|_Z.
\end{aligned} \tag{5.3}$$

Luego, de (5.2) y de (5.3) se sigue que

$$\begin{aligned}
\rho_2^N(t) &\leq C_\delta e^{t_1} \int_0^t \frac{1}{(t - s)^\delta} \|F(s, z^N(s)) - F(s, z(s))\|_Z ds \\
&\quad + C_\delta e^{t_1} \int_0^t \frac{1}{(t - s)^\delta} \|(P^N - I)F(s, z(s))\|_Z ds.
\end{aligned} \tag{5.4}$$

El integrando del segundo término del lado derecho de (5.4) está acotado uniformemente en N por $\frac{2}{(t-s)^\delta} \|F(s, z(s))\|_Z \in L^1(0, t)$, y converge a cero cuando N tiende a infinito. Por el Teorema de la Convergencia Dominada de Lebesgue, el segundo término de la última desigualdad tiende a cero cuando N crece.

Resumiendo, tenemos que

$$\|z^N(t) - z(t)\|_\delta \leq \epsilon^N(t) + \int_0^t \frac{\tilde{C}}{(t - s)^\delta} \|F(s, z^N(s)) - F(s, z(s))\|_Z ds \tag{5.5}$$

donde, para $t \in [0, t'_1]$, $\epsilon^N(t) \leq C$ para todo $N \in \mathbb{N}$ y $\epsilon^N(t) \rightarrow 0$ cuando $N \rightarrow \infty$. En particular $\int_0^{t'_1} \epsilon^N(t) dt \rightarrow 0$ cuando $N \rightarrow \infty$.

Como $z^N(0) = P^N z_0$ existe $\delta^N > 0$ tal que $\|z^N(t)\|_\delta \leq M + C + 2$ para todo $t \in [0, \delta^N]$, donde $M \doteq \sup_{0 \leq t \leq t'_1} \|z(t)\|_\delta$. Sea L la constante Lipschitz para F correspondiente al conjunto $U \doteq [0, t'_1] \times \{\|z\|_\delta \leq M + C + 2\}$. Entonces, de (5.5)

$$\|z^N(t) - z(t)\|_\delta \leq \epsilon^N(t) + \tilde{C}L \int_0^t \frac{1}{(t - s)^\delta} \|z^N(s) - z(s)\|_\delta ds$$

para todo $t \in [0, \delta^N]$, y si $K = K(\delta, \tilde{C}L, t'_1)$ como en el Lema 1.33, entonces

$$\|z^N(t) - z(t)\|_\delta \leq f^N(t), \quad \text{para todo } t \in [0, \delta^N], \quad (5.6)$$

donde, para $t \in [0, t'_1]$, se define $f^N(t) \doteq \epsilon^N(t) + K \int_0^t \epsilon^N(t)/(t-s)^\delta ds$.

A continuación se mostrará que existe $N_0 \in \mathbb{N}$ tal que $f^N(t) \leq C + 1$ para todo $t \in [0, t'_1]$ y para todo $N \geq N_0$. Como se verá más adelante, ésto no solo implicará la existencia de $z^N(t)$ en todo el intervalo $[0, t'_1]$ para todo $N \geq N_0$, sino también que $\|z^N(t)\|_\delta \leq M + C + 2$, para $t \in [0, t'_1]$ y $N \geq N_0$.

En efecto, observemos que

$$\int_0^t \frac{\epsilon^N(t)}{(t-s)^\delta} ds \leq \int_0^t \frac{C}{(t-s)^\delta} ds \quad (5.7)$$

$$= \int_0^t \frac{C}{s^\delta} ds \quad (5.8)$$

$$= \frac{C}{1-\delta} t^{1-\delta}. \quad (5.9)$$

Elijiendo $\eta = \eta(C, K) = \left(\frac{1-\delta}{2CK}\right)^{\frac{1}{1-\delta}}$ se tiene que $t^{1-\delta} \leq \frac{1-\delta}{2CK}$ para todo $t \in [0, \eta]$, luego

$$\int_0^t \frac{\epsilon^N(t)}{(t-s)^\delta} ds \leq \frac{1}{2K} \quad \text{para todo } t \in [0, \eta]. \quad (5.10)$$

Por otro lado, si $\eta < t \leq t'_1$

$$\begin{aligned} \int_0^t \frac{\epsilon^N(t)}{(t-s)^\delta} ds &= \int_0^t \frac{\epsilon^N(t-s)}{s^\delta} ds \\ &= \int_0^\eta \frac{\epsilon^N(t-s)}{s^\delta} ds + \int_\eta^t \frac{\epsilon^N(t-s)}{s^\delta} ds \\ &\leq \frac{C}{1-\delta} \eta^{1-\delta} + \int_\eta^t \frac{\epsilon^N(t-s)}{\eta^\delta} ds \\ &\leq \frac{1}{2K} + \frac{1}{\eta^\delta} \int_\eta^t \epsilon^N(t-s) ds \\ &\leq \frac{1}{2K} + \frac{1}{\eta^\delta} \int_\eta^{t'_1} \epsilon^N(s) ds. \end{aligned}$$

Luego, como $\int_\eta^{t'_1} \epsilon^N(s) ds \rightarrow 0$, existe N_0 tal que

$$\int_0^t \frac{\epsilon^N(t)}{(t-s)^\delta} ds \leq \frac{1}{K} \quad \text{para todo } t \in [\eta, t'_1] \text{ y } N \geq N_0. \quad (5.11)$$

De (5.10) y (5.11) y la definición de f^N se sigue que

$$f^N(t) \leq C + \frac{K}{K} = C + 1 \quad \text{para todo } t \in [0, t'_1] \text{ y } N \geq N_0, \quad (5.12)$$

como queríamos.

Consecuentemente, de (5.6) y (5.12) se tiene

$$\|z^N(t) - z(t)\|_\delta \leq C + 1 \quad \text{para todo } N \geq N_0 \text{ y } t \in [0, \delta^N],$$

lo que implica que

$$\|z^N(t)\|_\delta \leq M + C + 1 \quad \text{para todo } N \geq N_0 \text{ y } t \in [0, \delta^N].$$

Sea ahora $N \geq N_0$ fijo. Entonces $z^N(t)$ existe en $[0, t'_1]$ y para $t \in [0, t'_1]$, $\|z^N(t)\|_\delta < M + C + 2$. En efecto, supongamos, por el contrario, que existe $t^* \leq t'_1$ tal que $\|z^N(t^*)\|_\delta = M + C + 2$ y $\|z^N(t)\|_\delta < M + C + 2$ para $0 \leq t < t^*$. Entonces, por (5.6) $\|z^N(t) - z(t)\|_\delta \leq f^N(t) \leq C + 1$ en $[0, t^*]$, y por lo tanto $\|z^N(t)\|_\delta \leq M + C + 1$ en $[0, t^*]$, luego $\|z^N(t^*)\|_\delta \leq M + C + 1$, que contradice $\|z^N(t^*)\|_\delta = M + C + 2$.

Hasta el momento se ha probado que para todo $N \geq N_0$, $\|z^N(t)\|_\delta < M + C + 2$ para todo $t \in [0, t'_1]$. Por lo tanto, para $N \geq N_0$, δ^N puede elegirse estrictamente mayor a t'_1 y en consecuencia (5.6) se cumple en $[0, t'_1]$, es decir

$$\|z^N(t) - z(t)\|_\delta \leq f^N(t), \quad \text{para todo } t \in [0, t'_1].$$

Finalmente, por el Teorema de la Convergencia Dominada $f^N(t) \rightarrow 0$ para todo $t \in [0, t'_1]$ cuando $N \rightarrow \infty$ y el teorema está probado. \square

Hasta aquí hemos visto que las soluciones del problema en dimensión finita (5.1) convergen a la solución del problema (4.5). Aquél es un problema más tratable a nivel computacional, pero aún debemos hacer una discretización en la variable t , y ver que ésta converge. La siguiente sección trata este problema.

5.2 Discretización Temporal

En esta sección primero encontraremos la representación del problema aproximante (5.1) en términos de los coeficientes de la solución en la base $\hat{\beta}_N$ de Z^N . Luego definiremos

la discretización temporal que utilizaremos y a continuación se verá un resultado de convergencia y estabilidad del método resultante.

Llamemos w^N al vector cuyas componentes son los coeficientes de la solución $z^N(t) = \begin{pmatrix} u^N(t) \\ v^N(t) \\ \theta^N(t) \end{pmatrix}$ del problema (5.1) en la base $\hat{\beta}_N$. Luego w^N es solución del PVI

$$\begin{cases} \dot{w}^N(t) = \tilde{A}^N w^N(t) + \tilde{F}(t, w^N(t)) \\ w^N(0) = \gamma^N \end{cases} \quad (5.13)$$

con

$$\tilde{A}^N = (Q^N)^{-1} K^N, \quad \tilde{F}^N(t, w) = (Q^N)^{-1} R^N F(t, Q^N w), \quad \gamma^N = (Q^N)^{-1} R^N \begin{pmatrix} u_0 \\ v_0 \\ \theta_0 \end{pmatrix},$$

donde las matrices Q^N , K^N , y el mapeo $R^N : Z^N \rightarrow \mathbb{R}^{3N}$ se definen por

$$(Q^N)_{i,j} = \langle \beta_i^N, \beta_j^N \rangle, \quad (K^N)_{i,j} = \langle \beta_i^N, A^N \beta_j^N \rangle, \quad (R^N z)_i = \langle \beta_i^N, z \rangle,$$

$i, j = 1, 2, \dots, 3N$.

Para la discretización temporal definimos el siguiente esquema semi-implícito de Euler:

$$\begin{aligned} w_0^N &= \gamma^N \\ \frac{1}{\Delta t} (w_{k+1}^N - w_k^N) &= \tilde{A}^N w_{k+1}^N + \tilde{F}^N(k\Delta t, w_k^N), \quad k = 0, 1, \dots \end{aligned}$$

La convergencia de las soluciones del sistema completamente discretizado a la solución de (5.13) cuando $\Delta t \rightarrow 0$ se sigue inmediatamente del siguiente teorema.

Teorema 5.4 *Sea n un entero positivo, $B \in \mathcal{L}(\mathbb{R}^n)$, $G : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función continua tal que $G(t, \cdot)$ es Lipschitz continua en \mathbb{R}^n para todo $t \in [0, T]$, con constante K independiente de t . Supongamos que el PVI*

$$\begin{cases} W'(t) = BW(t) + G(t, W(t)), & W(t) \in \mathbb{R}^n, \\ W(0) = W_0 \end{cases}$$

tiene una única solución $W(\cdot)$ en $[0, T]$ tal que $W''(t)$ es acotada en $[0, T]$. Dado $0 < h < 1/\|B\|$ y $w_0 \in \mathbb{R}^n$ definamos $N(h) \doteq [T/h]$ (la parte entera de N/h) y dado w_0 ,

definamos w_j , $j = 1, 2, \dots, N(h)$ por

$$w_{j+1} = w_j + hBw_{j+1} + hG(jh, w_j), \quad j = 0, 1, \dots, N(h) - 1.$$

Entonces, existe una constante \tilde{C} tal que

$$\max_{j=0,1,\dots,N(h)} |W(jh) - w_j| \leq e^{T\left(\frac{K+\|B\|}{1-h\|B\|}\right)} |W_0 - w_0| + h\tilde{C} \left[e^{T\left(\frac{K+\|B\|}{1-h\|B\|}\right)} - 1 \right].$$

Más aún, si $h < 1/(2\|B\|)$, entonces

$$\max_{j=0,1,\dots,N(h)} |W(jh) - w_j| \leq e^{2T(K+\|B\|)} |W_0 - w_0| + h\tilde{C} \left[e^{2T(K+\|B\|)} - 1 \right].$$

Comentario. No es difícil demostrar que las conclusiones de este teorema permanecen válidas bajo las condiciones más débiles de G localmente Lipschitz y W' Lipschitz continua.

Demostración. Sea $0 < h < 1/\|B\|$ y definamos $W_j = W(jh)$. Usando el Teorema de Aproximación de Taylor

$$\begin{aligned} W_{j+1} &= W_j + hW'(jh) + \frac{h^2}{2}W''(\xi_{j,h}) \\ &= W_j + hBW_j + hG(jh, W_j) + \frac{h^2}{2}W''(\xi_{j,h}), \end{aligned}$$

donde $\xi_{i,j} \in (jh, (j+1)h)$. Definiendo $e_j = W_j - w_j$ se tiene que

$$\begin{aligned} e_{j+1} &= e_j + hB(W_j - w_{j+1}) + h[G(jh, W_j) - G(jh, w_j)] + \frac{h^2}{2}W''(\xi_{j,h}) \\ &= e_j + hBe_{j+1} + h[G(jh, W_j) - G(jh, w_j)] + hB(W_j - W_{j+1}) + \frac{h^2}{2}W''(\xi_{j,h}), \end{aligned}$$

o, equivalentemente

$$e_{j+1} - hBe_{j+1} = e_j + h[G(jh, W_j) - G(jh, w_j)] + hB(W_j - W_{j+1}) + \frac{h^2}{2}W''(\xi_{j,h}).$$

Por lo tanto,

$$\begin{aligned} (1 - h\|B\|)|e_{j+1}| &\leq |e_j| + h|G(jh, W_j) - G(jh, w_j)| + h\|B\||W_j - W_{j+1}| + \frac{h^2}{2}|W''(\xi_{j,h})| \\ &\leq |e_j| + hK|e_j| + h^2\|B\| \left| W'(\tilde{\xi}_{j,h}) \right| + \frac{h^2}{2}|W''(\xi_{j,h})| \end{aligned}$$

donde $\tilde{\xi}_{j,h} \in (jh, (j+1)h)$.

Si C es una cota superior para $|W'(t)| + |W''(t)|$ en $[0, T]$, se sigue que

$$|e_{j+1}| \leq \frac{1+hK}{1-h\|B\|} |e_j| + \frac{h^2 C(1+\|B\|)}{1-h\|B\|}.$$

Por inducción se tiene que

$$|e_j| \leq \left[\frac{1+hK}{1-h\|B\|} \right]^j |e_0| + \frac{h^2 C(1+\|B\|)}{1-h\|B\|} \frac{\left[\frac{1+hK}{1-h\|B\|} \right]^j - 1}{\left[\frac{1+hK}{1-h\|B\|} \right] - 1}.$$

Ahora bien, como $(1+x)^j \leq e^{jx}$ para todo $x > -1$ y $j \in \mathbb{N}$, obtenemos

$$\left[\frac{1+hK}{1-h\|B\|} \right]^j = \left[1 + h \frac{K+\|B\|}{1-h\|B\|} \right]^j \leq e^{jh \left(\frac{K+\|B\|}{1-h\|B\|} \right)} \leq e^{T \left(\frac{K+\|B\|}{1-h\|B\|} \right)}$$

y

$$\frac{\left[\frac{1+hK}{1-h\|B\|} \right]^j - 1}{\left[\frac{1+hK}{1-h\|B\|} \right] - 1} = \frac{\left[\frac{1+hK}{1-h\|B\|} \right]^j - 1}{h \left[\frac{K+\|B\|}{1-h\|B\|} \right] - 1}.$$

Por lo tanto

$$|W_j - w_j| \leq e^{T \left(\frac{K+\|B\|}{1-h\|B\|} \right)} |W_0 - w_0| + hC \frac{1+\|B\|}{K+\|B\|} \left[e^{T \left(\frac{K+\|B\|}{1-h\|B\|} \right)} - 1 \right]$$

para todo $j = 0, 1, \dots, N(h)$. Esto prueba la primera parte del teorema. La afirmación final se sigue del hecho que $1 - h\|B\| > 1/2$ si $h < 1/2\|B\|$. \square

Comentario. El teorema anterior junto con el resultado del Teorema 5.3 asegura la convergencia del sistema discretizado en espacio y tiempo a las soluciones de (4.5) cuando $N \rightarrow \infty$ y $\Delta t \rightarrow 0$ de algún modo apropiado. Más aún, para N fijo, el Teorema 5.4 dice que el orden de convergencia para las ecuaciones discretizadas en tiempo a la solución de (5.13) es $O(\Delta t)$.

5.3 Resultados Numéricos

En esta sección utilizaremos el algoritmo resultante de las aproximaciones modales introducidas en la Sección 5.1 y de la discretización temporal presentada en la Sección 5.2

para obtener aproximaciones de la solución al problema de valores iniciales (4.5) que es equivalente al problema de valores iniciales y de frontera (4.3)–(4.4). La intención es ver experimentalmente si la solución simula, al menos cualitativamente, el comportamiento de un sólido unidimensional de material con memoria de forma bajo diferentes condiciones iniciales, acciones de fuerza y aplicación de calor externas.

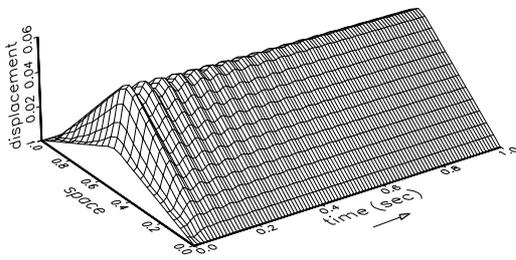
Para los experimentos numéricos que presentamos a continuación hemos utilizado el método expuesto en la sección anterior con $N = 32$ y $\Delta t = 10^{-5}$ y los valores de los parámetros informados por Falk en [24] para la aleación $\text{Au}_{23}\text{Cu}_{30}\text{Zn}_{47}$: $\alpha_2 = 24 \text{ J cm}^{-3} \text{ K}^{-1}$, $\alpha_4 = 1.5 \times 10^5 \text{ J cm}^{-3}$, $\alpha_6 = 7.5 \times 10^6 \text{ J cm}^{-3} \text{ K}^{-1}$, $\theta_1 = 208 \text{ K}$, $C_v = 2.9 \text{ J cm}^{-3} \text{ K}^{-1}$, $k = 1.9 \text{ W cm}^{-1} \text{ K}^{-1}$, $\rho = 11.1 \text{ g cm}^{-3}$. También tomamos $\gamma = 10^{-12} \text{ J cm}^{-1}$ como fue reportado en [26]. Para β tomamos $\beta = 1$. Esta elección no tiene ningún significado físico particular. A nuestro conocimiento, no se han reportado aún valores de β para materiales reales, aunque parece haber evidencia que para algunos materiales con memoria de forma β es muy pequeño o cero.

Experimento 1: *Estado Estacionario a Baja Temperatura*

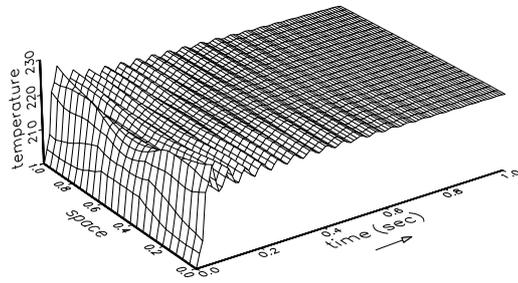
En este experimento tomamos $f = g = 0$, $\theta_0(x) \equiv 200$ K, $u_0(x) = P^N h(x)$, donde

$$h(x) = \begin{cases} 0.05x, & \text{si } 0 \leq x \leq 0.5, \\ 0.05(1-x), & \text{si } 0.5 \leq x \leq 1, \end{cases}$$

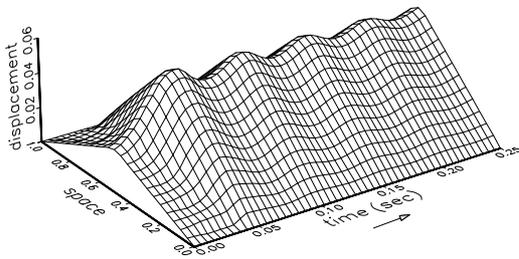
y $v_0 \equiv 0$. Es decir, la viga se encuentra inicialmente en el rango de bajas temperaturas compuesto de dos segmentos de martensita, a saber, martensita M_+ en $0 \leq x < \frac{1}{2}$ y martensita M_- en $\frac{1}{2} < x \leq 1$ (con una deformación inicial del 5%). La evolución del desplazamiento y la temperatura puede verse en las Figuras 5.1a y 5.1b, respectivamente. Esta evolución se debe a que la condición inicial $z_0(x) = \begin{pmatrix} u_0(x) \\ v_0(x) \\ \theta_0(x) \end{pmatrix}$ no es estado estacionario del sistema (4.5). El sistema evoluciona hasta que se llega a un estado estacionario compuesto de dos segmentos simétricos de martensita M_+ y M_- (deformación $\cong 11.25\%$) y temperatura constante $\cong 222$ K. La Figura 5.1c muestra en mayor detalle el desplazamiento durante los primeros 250 milisegundos.



(a)



(b)



(c)

Figura 5.1: Estado estacionario a baja temperatura. Evolución del desplazamiento (a,c) y temperatura (b) a partir de una condición inicial no estacionaria.

Experimento 2: *Estado Estacionario a Alta Temperatura*

Aquí tomamos $\theta_0(x) \equiv 600$ K y u_0, v_0, f , y g como en el Experimento 1. La evolución del desplazamiento y de la temperatura se muestra en las Figuras 5.2a y 5.2b, respectivamente. La viga oscila hasta que se llega al estado estacionario consistente de deformación nula y temperatura constante $\theta \cong 505.6$ K. Este comportamiento está de acuerdo con el hecho que por encima de la temperatura *austenite-finish* (fin de la fase austenítica) $\theta = A_f$ (en este caso $A_f \cong 283$ K) los estados estacionarios satisfacen $u \equiv 0$ y $\theta = const.$ Debido al estado inicial no estacionario a alta temperatura, la viga inmediatamente se desdobra aproximándose al estado $u \equiv 0$ mientras que la temperatura decrece levemente, originando las oscilaciones amortiguadas que se observan en las Figuras 5.2a y 5.2b. En la Figura 5.2c, pueden apreciarse las oscilaciones del punto medio de la viga.

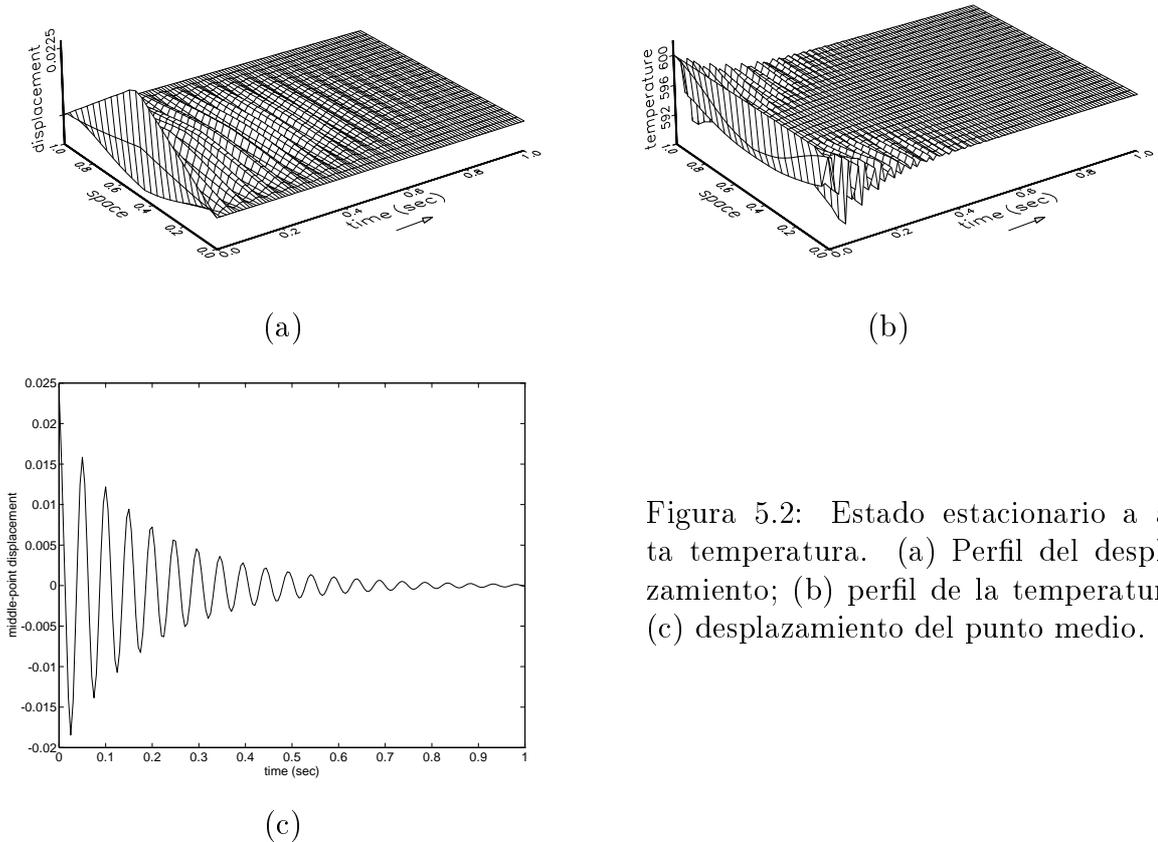


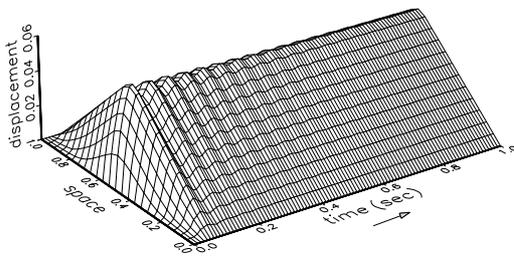
Figura 5.2: Estado estacionario a alta temperatura. (a) Perfil del desplazamiento; (b) perfil de la temperatura; (c) desplazamiento del punto medio.

Experimento 3: Pulso a Baja Temperatura

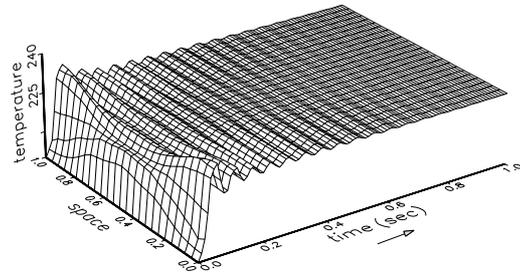
En este experimento pretendimos estudiar los efectos de una fuerza distribuida consistente de un pulso de fuerza alrededor del punto medio de la viga cuando la temperatura inicial está debajo de la temperatura *martensite-finish* (fin de la fase martensítica) $\theta = M_f \cong 208$ K. Se tomó $u_0(x) = v_0(x) \equiv 0$, $\theta(x) \equiv 200$ K, $g(x, t) \equiv 0$, y

$$f(x, t) = \begin{cases} 5 \times 10^4, & \text{si } 0.4 \leq x \leq 0.6 \text{ y } 0 < t < 0.5 \times 10^{-3} \\ 0, & \text{en otro caso.} \end{cases}$$

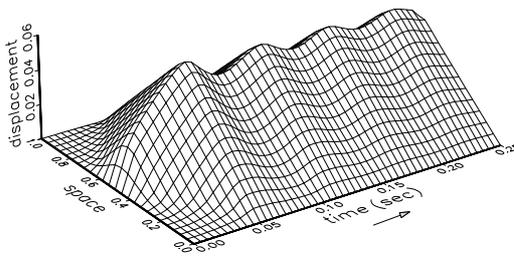
Inicialmente, los puntos alrededor del centro se desplazan hacia arriba mientras el efecto del pulso se propaga hacia los extremos de la viga (Figuras 5.3a y 5.3c). En el mismo instante en que este efecto alcanza los extremos, el punto medio llega a un máximo y pequeñas oscilaciones amortiguadas comienzan a ocurrir (Figura 5.3c) alrededor del estado final de equilibrio consistente de dos segmentos simétricos de martensita M_+ , M_- (deformación $\cong 11.05\%$) y temperatura constante 226 K (Figura 5.3b).



(a)



(b)



(c)

Figura 5.3: Pulso a baja temperatura. (a), (c) perfil del desplazamiento; (b) perfil de la temperatura.

Experimento 4: Pulso a Alta Temperatura

En este caso investigamos los efectos de un pulso alrededor del punto medio de la viga, que estaba inicialmente a temperatura constante mayor a A_f . Se tomó $\theta_0(x) \equiv 600$ K y u_0, v_0, f , y g como en el experimento 3. Al principio, la viga se dobla hacia arriba hasta que se elimina la aplicación de la fuerza (Figura 5.4c). Inmediatamente después, comienzan a ocurrir oscilaciones amortiguadas alrededor del estado final de equilibrio $u \equiv 0$ y temperatura constante $\theta \cong 602$ K (Figuras 5.4a y 5.4b). Recordemos que sobre la temperatura de finalización de la fase austenítica el único estado estacionario es $u \equiv 0$ y $\theta \equiv const$.

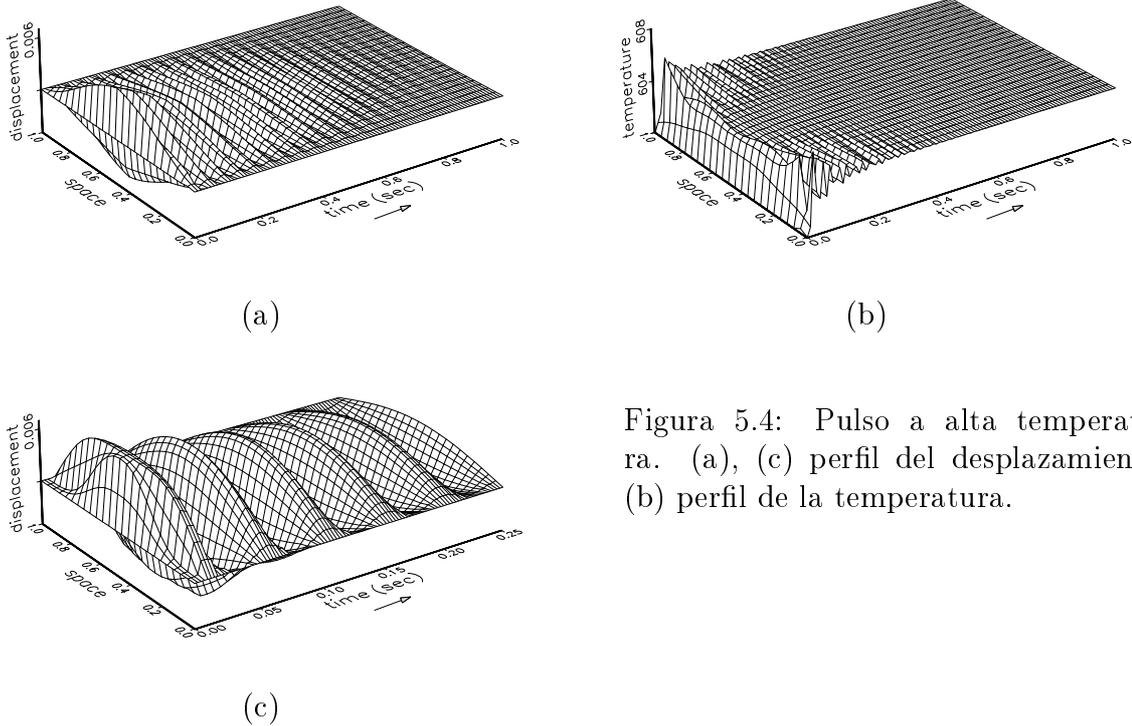


Figura 5.4: Pulso a alta temperatura. (a), (c) perfil del desplazamiento; (b) perfil de la temperatura.

Experimento 5: Esperar y Calentar

En este experimento observamos los efectos de calentar la viga cuando inicialmente está en un estado de equilibrio consistente de dos segmentos simétricos de martensita M_+ y M_- . Para esto tomamos como estado inicial el estado estacionario final del Experimento 1 (deformación inicial 11.25% y $\theta_0(x) \equiv 222$ K), $f(x,t) \equiv 0$ y la fuente de calor $g(x,t)$ consistente de un pulso de calor distribuido uniformemente como sigue

$$g(x,t) = \begin{cases} 5 \times 10^4, & \text{si } 0.2 < t < 0.25, \\ 0, & \text{en otro caso.} \end{cases}$$

El sistema se mantiene en el estado inicial hasta que se aplica el pulso de calor. En este momento la temperatura comienza a crecer (Figura 5.5b), los cristales de martensita se convierten en austenita y la viga se desdobra mostrando pequeñas oscilaciones amortiguadas alrededor del estado de deformación nula (Figura 5.5a). Estas oscilaciones desaparecen rápido y la viga alcanza el estado estacionario $u \equiv 0$, $\theta \cong 336$ K.

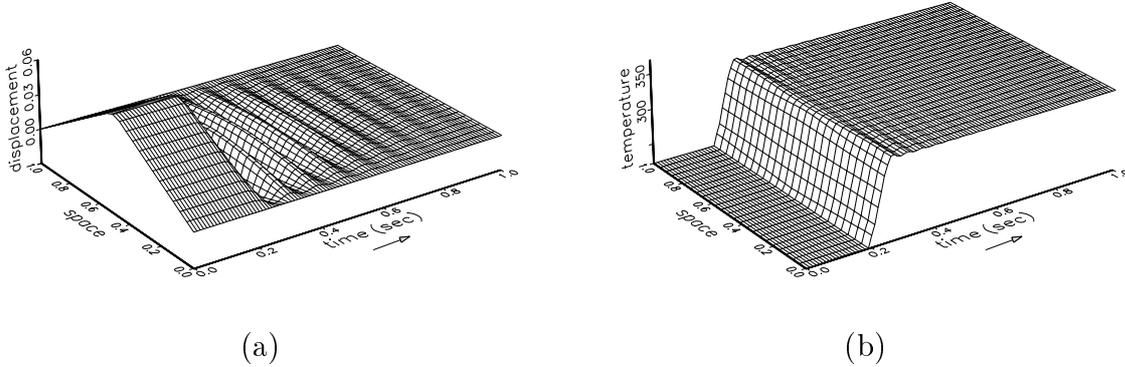


Figura 5.5: Esperar y Calentar. (a), (c) perfil del desplazamiento; (b) perfil de la temperatura.

Experimento 6: *Calentar Esperar y Enfriar*

Para este experimento tomamos nuevamente como estado inicial el estado final del Experimento 1. También se tomó $f(x, t) \equiv 0$ y las fuentes de calor distribuidas $g(x, t)$ consistentes de un pulso inicial uniformemente distribuido que se suspende después de $t = 0.05$ seg. y cuando $t = 1.45$ seg. se aplica un pulso opuesto hasta $t = 1.50$ seg. cuando se suspende nuevamente. Más precisamente,

$$g(x, t) = \begin{cases} 8 \times 10^3, & \text{si } t < 0.05, \\ -8 \times 10^3, & \text{si } 1.45 < t < 1.50, \\ 0 & \text{en otro caso.} \end{cases}$$

La temperatura crece uniformemente hasta aproximadamente 336 K mientras la viga se aproxima al estado de deformación nula. Después que se suspende el pulso de calor, la temperatura permanece a 336 K mientras que el desplazamiento muestra pequeñas oscilaciones amortiguadas alrededor de $u \equiv 0$. La muestra está ahora completamente en la fase austenítica. Cuando $t = 1.45$ seg., al aplicar el pulso opuesto, la temperatura decrece uniformemente y finalmente permanece aproximadamente a 222 K, mientras que la viga sufre el proceso inverso que la lleva al estado inicial original mostrando el efecto de memoria de forma de dos vías (*two-way shape memory effect*). Figuras 5.6a-d.

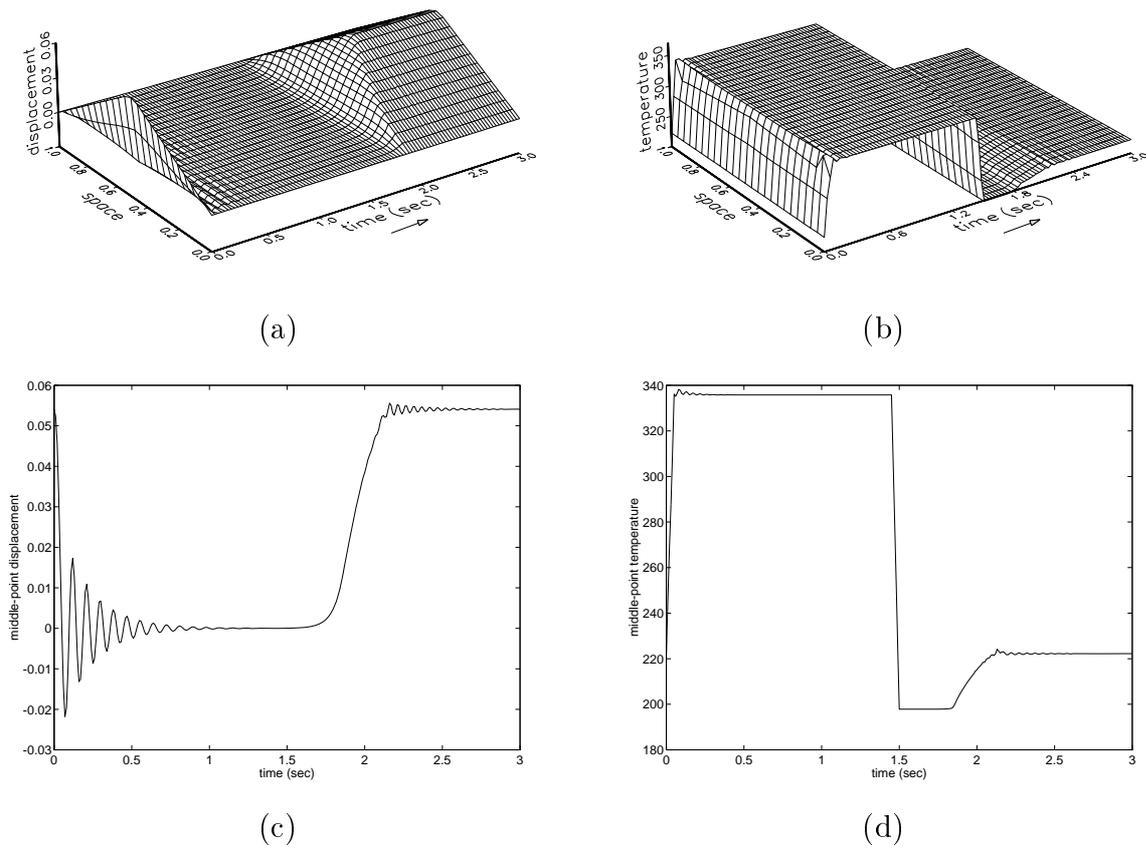


Figura 5.6: Calentar Esperar Enfriar. (a) perfil del desplazamiento; (b) perfil de la temperatura; (c) desplazamiento del punto medio; (d) temperatura del punto medio.

Capítulo 6

Una Aplicación del Algoritmo de Cuasilinealización

En este capítulo utilizaremos el algoritmo de Cuasilinealización que se expuso en la Sección 2.1 para identificar los parámetros α_2 , α_4 , α_6 y θ_1 del problema de valores iniciales y de frontera (4.3)–(4.4).

En la Sección 6.1 daremos un resultado sobre la identificabilidad de los parámetros α_2 , α_4 , α_6 y θ_1 del sistema (4.3)–(4.4). En la Sección 6.2 mostraremos que hay algunas situaciones en que los parámetros mencionados arriba no pueden ser identificados debido a que existen diversos juegos de parámetros que dan lugar a la misma solución. En la Sección 6.3 tomaremos como observaciones datos sintéticos generados con la computadora resolviendo numéricamente las ecuaciones y aplicaremos el algoritmo de cuasilinealización con esos datos, para comprobar que el algoritmo efectivamente funciona en la práctica y si tiene sentido intentar realizar la identificación con datos experimentales.

6.1 Identificabilidad de Parámetros

A “grosso modo” la *identificabilidad de parámetros* puede definirse como la continuidad e inversibilidad local del mapeo $q \rightarrow z(t; q)$ del conjunto \mathcal{Q} de parámetros admisibles en el espacio de soluciones. El siguiente teorema muestra un resultado concerniente a la inversibilidad del mapeo $q \rightarrow z(t; q)$ donde $z(t; q)$ es solución de (4.5), con A y $F(q, t, z)$ dados por (4.7) y (4.8), respectivamente. En la Sección 6.3 veremos un resultado

relacionado con la continuidad, donde además veremos que existe la derivada de Fréchet de $z(t; q)$ con respecto a q y es q -Lipschitz continua.

Teorema 6.1 Sean $\delta \in (\frac{3}{4}, 1)$, $q = (\alpha_2, \alpha_4, \alpha_6, \theta_1) \in \mathcal{Q}$, y $\tilde{q} = (\tilde{\alpha}_2, \tilde{\alpha}_4, \tilde{\alpha}_6, \tilde{\theta}_1) \in \mathcal{Q}$. Sea

$$z_0(x) = \begin{pmatrix} u_0(x) \\ v_0(x) \\ \theta_0(x) \end{pmatrix} \in Z_\delta$$

y supongamos que las soluciones $z(t; q)$ y $z(t; \tilde{q})$ de (4.5) coinciden en $T_1 \leq t \leq T_2$ para algún par de valores $0 \leq T_1 < T_2$. Supongamos además que

$$z(t; q) = \begin{pmatrix} u(\cdot, t) \\ v(\cdot, t) \\ \theta(\cdot, t) \end{pmatrix}$$

es tal que $v(\cdot, t^*) \not\equiv 0$ para algún $t^* \in (T_1, T_2)$. Entonces $q = \tilde{q}$.

Comentario. Notemos que las hipótesis de este teorema se satisfacen si $T_1 = 0$ y la condición inicial

$$z_0(x) = \begin{pmatrix} u_0(x) \\ v_0(x) \\ \theta_0(x) \end{pmatrix}$$

se elige de tal modo que $v_0 \not\equiv 0$. O bien, si se elige una función $f(x, t)$ de fuerzas distribuidas que sea capaz de lograr un desplazamiento fuera de la posición de equilibrio.

Demostración. Como $v(\cdot, t^*) \not\equiv 0$ y $v(0, t^*) = v(1, t^*) = 0$ por las condiciones de frontera, existe $\hat{x} \in (0, 1)$ tal que $v_x(\hat{x}, t^*) \neq 0$. Por continuidad $v_x(\hat{x}, t) \neq 0$ en $t^* - \epsilon < t < t^* + \epsilon$ para algún $\epsilon > 0$. Dado que $v_x = u_{xt}$ debe existir $\hat{t} \in (t^* - \epsilon, t^* + \epsilon)$ tal que

$$u_x(\hat{x}, \hat{t}) \neq 0 \tag{6.1}$$

Por otro lado,

$$\begin{aligned} \dot{z}(t; q) &= Az(t; q) + F(q, t, z(t; q)), & y \\ \dot{z}(t; \tilde{q}) &= Az(t; \tilde{q}) + F(\tilde{q}, t, z(t; \tilde{q})), & \forall t \in [T_1, T_2]. \end{aligned} \tag{6.2}$$

Luego, como $z(t; q) = z(t; \tilde{q})$ para todo $t \in [T_1, T_2]$, (6.2) implica

$$F(q, t, z(t; q)) = F(\tilde{q}, t, z(t; \tilde{q})) \quad \forall t \in [T_1, T_2], \tag{6.3}$$

y usando la definición de $F(q, t, z)$,

$$\begin{aligned} \frac{\partial}{\partial x} [2\alpha_2(\theta(x, t) - \theta_1)u_x(x, t) - 4\alpha_4u_x(x, t)^3 + 6\alpha_6u_x(x, t)^5] \\ = \frac{\partial}{\partial x} [2\tilde{\alpha}_2(\theta(x, t) - \tilde{\theta}_1)u_x(x, t) - 4\tilde{\alpha}_4u_x(x, t)^3 + 6\tilde{\alpha}_6u_x(x, t)^5] \end{aligned} \quad (6.4)$$

y

$$\alpha_2\theta(x, t)u_x(x, t)v_x(x, t) = \tilde{\alpha}_2\theta(x, t)u_x(x, t)v_x(x, t) \quad (6.5)$$

para todo $x \in (0, 1)$, $t \in [T_1, T_2]$.

De (6.5) se sigue que

$$(\alpha_2 - \tilde{\alpha}_2)\theta(x, t)u_x(x, t)v_x(x, t) = 0 \quad \forall x \in (0, 1), t \in [T_1, T_2].$$

Como $u_x(\hat{x}, \hat{t})v_x(\hat{x}, \hat{t}) \neq 0$ y $\theta(x, t) > 0$, concluimos que $\alpha_2 = \tilde{\alpha}_2$. Luego, (6.4) luce

$$u_{xx}(x, t) \left(2\alpha_2(\theta_1 - \tilde{\theta}_1) + 12(\alpha_4 - \tilde{\alpha}_4)u_x(x, t)^2 - 30(\alpha_6 - \tilde{\alpha}_6)u_x(x, t)^4 \right) = 0 \quad (6.6)$$

para todo $x \in (0, 1)$, $t \in [T_1, T_2]$.

Ahora bien, si $u_{xx}(x, \hat{t})$ fuera idénticamente cero en $(0, 1)$, entonces, por las condiciones de frontera, tendríamos que $u(\cdot, \hat{t}) \equiv 0$ y $u_x(\cdot, \hat{t}) \equiv 0$, lo que obviamente contradice (6.1). Por lo tanto, deben existir constantes a y b , $0 < a, b < 1$ tales que $u_{xx}(x, \hat{t}) \neq 0$ para $a < x < b$, lo que implica que $u_x(x, \hat{t})$ no puede ser constante en (a, b) . Entonces las funciones 1 , $u_x(x, \hat{t})^2$, y $u_x(x, \hat{t})^4$ son linealmente independientes como funciones de x en (a, b) . Por lo tanto, (6.6) implica que $\alpha_4 = \tilde{\alpha}_4$, $\alpha_6 = \tilde{\alpha}_6$, y $\theta_1 = \tilde{\theta}_1$. \square

6.2 No Identificabilidad de Parámetros

En esta sección veremos que si se toma un estado estacionario del sistema, entonces existen diferentes juegos de parámetros para los cuales el estado estacionario es solución del problema.

Teorema 6.2 Sean A y $F(q, t, z)$ definidos como en (4.7) y (4.8), respectivamente, con $f(x, t) \equiv g(x, t) \equiv 0$, y sea $q \in \mathcal{Q}$. Sea $z = \begin{pmatrix} u(x) \\ v(x) \\ \theta(x) \end{pmatrix} \in D(A)$ tal que $Az + F(q, t, z) = 0$, entonces, existen infinitos $\tilde{q} \in \mathcal{Q}$ tal que $Az + F(\tilde{q}, t, z) = 0$.

Demostración. Por la definición de A y $F(q, t, z)$, si $z = \begin{pmatrix} u(x) \\ v(x) \\ \theta(x) \end{pmatrix}$ es tal que $Az + F(q, t, z) = 0$ se tiene que

$$v(x) \equiv 0, \quad x \in [0, 1], \quad (6.7)$$

$$\begin{aligned} \rho\beta v''(x) - \gamma u^{(4)}(x) + 2\alpha_2\theta'(x)u'(x) + 2\alpha_2(\theta(x) - \theta_1)u''(x) \\ - 12\alpha_4u''(x)u'(x)^2 + 30\alpha_6u''(x)u'(x)^4 = 0, \end{aligned} \quad (6.8)$$

$$k\theta''(x) + 2\alpha_2\theta(x)u'(x)v'(x) + \beta\rho v'(x)^2 = 0. \quad (6.9)$$

De (6.7) se sigue que $v'(x) = 0$ en $(0, 1)$ y entonces de (6.9), y teniendo en cuenta las condiciones de frontera $\theta'(0) = \theta'(1) = 0$, se tiene que $\theta(x) \equiv \text{const}$ y por ende $\theta'(x) = 0$ en $(0, 1)$. Luego, (6.8) se transforma en

$$-\gamma u^{(4)}(x) + 2\alpha_2(\text{const} - \theta_1)u''(x) - 12\alpha_4u''(x)u'(x)^2 + 30\alpha_6u''(x)u'(x)^4 = 0,$$

ecuación que se satisface para cualquier par de constantes positivas $\tilde{\alpha}_2$ y $\tilde{\theta}_1$ que cumplan

$$\tilde{\alpha}_2(\text{const} - \tilde{\theta}_1) = \alpha_2(\text{const} - \theta_1)$$

□

6.3 Estimación de los Parámetros

En esta sección utilizaremos el algoritmo de Cuasilinealización introducido en la Sección 2.1 para identificar los parámetros α_2 , α_4 , α_6 , θ_1 que definen el funcional de energía libre Ψ del sistema, pero primero verificaremos que se cumplen las condiciones sobre A y F que aseguran la convergencia del algoritmo de cuasilinealización.

El siguiente teorema muestra que el operador A y la función F definidos en (4.7) y (4.8), respectivamente, satisfacen ciertas condiciones de regularidad, que, teniendo en cuenta los Teoremas 3.2 y 3.3 aseguran la existencia y continuidad Lipschitz local de la derivada de Fréchet del mapeo $q \rightarrow z(t; q)$.

Teorema 6.3 Sean Z , A y $F(q, t, z)$ los definidos en la Sección 4.3 con las hipótesis de regularidad sobre $f(x, t)$ y $g(x, t)$ allí mencionadas. Entonces el mapeo $(q, z(\cdot)) \rightarrow F(q, t, z(\cdot))$ de $\mathcal{Q} \times L^\infty(0, T : Z_\delta)$ en $L^\infty(0, T : Z)$ es diferenciable Fréchet en ambas variables. Además, los mapeos $(q, z(\cdot)) \rightarrow F_q(q, \cdot, z(\cdot))$ y $(q, z(\cdot)) \rightarrow F_z(q, \cdot, z(\cdot))$ son localmente Lipschitz continuos de $\mathcal{Q} \times L^\infty(0, T : Z_\delta)$ en $L^\infty(0, T : \mathcal{L}(\tilde{\mathcal{Q}}, Z))$ y de $\mathcal{Q} \times L^\infty(0, T : Z_\delta)$ en $L^\infty(0, T : \mathcal{L}(Z_\delta, Z))$, respectivamente.

Demostración. Este resultado se obtiene inmediatamente observando que $f_2(q, t, z)$ y $f_3(q, t, z)$, como han sido definidas en (4.9) son diferenciables Fréchet con respecto a q y z . Más aún, estas derivadas pueden calcularse explícitamente y están dadas por

$$\begin{aligned} D_z f_2(q, t, \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}) \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\theta} \end{pmatrix} &= f_{2,u} \tilde{u} + f_{2,v} \tilde{v} + f_{2,\theta} \tilde{\theta}, \\ D_z f_3(q, t, \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}) \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\theta} \end{pmatrix} &= f_{3,u} \tilde{u} + f_{3,v} \tilde{v} + f_{3,\theta} \tilde{\theta}, \\ D_q f_2(q, t, \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}) &= \frac{1}{\rho} [2\theta' u' + 2(\theta - \theta_1) u'', -12(u')^2 u'', 30(u')^4 u'', -2\alpha_2 u''], \\ D_q f_3(q, t, \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}) &= \frac{1}{C_v} [2\theta u' v', 0, 0, 0], \end{aligned}$$

Donde los operadores lineales $f_{i,u}$, $f_{i,v}$, y $f_{i,\theta}$, $i = 2, 3$ están dados por

$$\begin{aligned} f_{2,u} &= \frac{1}{\rho} \{ 2\alpha_2 \theta' D + 2\alpha_2 (\theta - \theta_2) D^2 - 24\alpha_4 u' u'' D - 12\alpha_4 (u')^2 D^2 \\ &\quad + 120\alpha_6 (u')^3 u'' D + 30\alpha_6 (u')^4 D^2 \}, \\ f_{2,v} &= 0, \\ f_{2,\theta} &= \frac{1}{\rho} \{ 2\alpha_2 u' D + 2\alpha_2 u'' \}, \\ f_{3,u} &= \frac{1}{C_v} \{ 2\alpha_2 \theta v' D \}, \\ f_{3,v} &= \frac{1}{C_v} \{ 2\alpha_2 \theta u' D + 2\beta \rho v' D \}, \\ f_{3,\theta} &= \frac{1}{C_v} \{ 2\alpha_2 u' v' \}, \end{aligned}$$

donde D denota al operador derivada con respecto a x . \square

Por el teorema anterior vemos que se cumplen las hipótesis de los Teoremas 3.2 y 3.3. Por lo tanto se tiene la convergencia del algoritmo de cuasilinealización definido en la Sección 2.1. A continuación utilizaremos este algoritmo para identificar el parámetro $q = (\alpha_2, \alpha_4, \alpha_6, \theta_1)$ en el sistema de ecuaciones en derivadas parciales (4.3) que modela, como ya hemos mencionado, el comportamiento de un sólido unidimensional con memoria de forma.

En los ejemplos que siguen haremos uso de los mismos valores de los parámetros que se utilizaron para la resolución numérica de las ecuaciones en el Capítulo 5. Estos valores son: $\alpha_2 = 24 \text{ J cm}^{-3} \text{ K}^{-1}$, $\alpha_4 = 1.5 \times 10^5 \text{ J cm}^{-3}$, $\alpha_6 = 7.5 \times 10^6 \text{ J cm}^{-3} \text{ K}^{-1}$, $\theta_1 = 208 \text{ K}$, $C_v = 2.9 \text{ J cm}^{-3} \text{ K}^{-1}$, $k = 1.9 \text{ w cm}^{-1} \text{ K}^{-1}$, $\rho = 11.1 \text{ g cm}^3$, $\beta = 1$, y $\gamma = 10^{-12} \text{ J cm}^{-1}$. Queremos estimar $q^* = (\alpha_2, \alpha_4, \alpha_6, \theta_1) = (24, 1.5 \times 10^5, 7.5 \times 10^6, 208)$.

Ejemplo 1: Datos Exactos

En este primer ejemplo investigamos el comportamiento del algoritmo cuando suponemos la existencia de un q^* que ajusta exactamente los datos. Para ello tomamos $u_0 \equiv 0$, $v_0 \equiv 0$ y $\theta_0 \equiv 200$ K, $g(x, t) \equiv 0$,

$$f(x, t) = \begin{cases} 1 \times 10^5, & \text{si } 0.4 \leq x \leq 0.6, \\ 0, & \text{en otro caso,} \end{cases}$$

y $T = 0.01$. Primero, obtenemos $u(x, t, q^*)$ y $\theta(x, t, q^*)$ resolviendo numéricamente el problema, utilizando el algoritmo propuesto en el Capítulo 5. Tomamos las observaciones como $\hat{z}_i = \left\{ \begin{pmatrix} u(x_j, t_i; q^*) \\ \theta(x_j, t_i; q^*) \end{pmatrix} \right\}_{j=1}^9$, donde $t_i = 0.001 i$, $i = 1, 2, \dots, 10$. Comenzamos con una estimación inicial $q^0 = (50, 3 \times 10^5, 15 \times 10^6, 420)$ aproximadamente igual a dos veces q^* . Los resultados de las iteraciones producidas por el algoritmo de cuasilinealización se muestran en la Tabla 6.1 y en la Figura 6.1. En la Figura 6.2a se muestra una comparación entre $u(x, T; q^*)$ y $u(x, T; q^k)$ mientras que en la Figura 6.2b se muestran los valores de $\theta(x, T; q^*)$ y $\theta(x, T; q^k)$ para diferentes valores de k .

k	α_2	α_4	α_6	θ_1	$J(q^k)$
0	50.0000	300000	1.50000e+07	420.000	1994.6900
1	16.1807	228111	1.40769e+07	459.904	611.1950
2	26.1790	222964	8.71784e+06	33.096	280.8220
3	25.3531	246241	8.83171e+06	126.468	15.3156
4	24.2770	178223	7.87660e+06	181.091	7.1313
5	24.0166	151184	7.51451e+06	206.550	0.6210
6	24.0012	150073	7.50096e+06	207.927	0.0122
7	24.0001	150006	7.50008e+06	207.994	0.0030
8	24.0001	150002	7.50003e+06	207.998	0.0029
9	24.0000	150002	7.50002e+06	207.998	0.0029
10	24.0000	150002	7.50002e+06	207.998	0.0029
11	24.0000	150002	7.50002e+06	207.998	0.0029
12	24.0000	150002	7.50002e+06	207.998	0.0029

Tabla 6.1: Valores de los parámetros y del criterio de error en las iteraciones del algoritmo de cuasilinealización en el ejemplo 1.

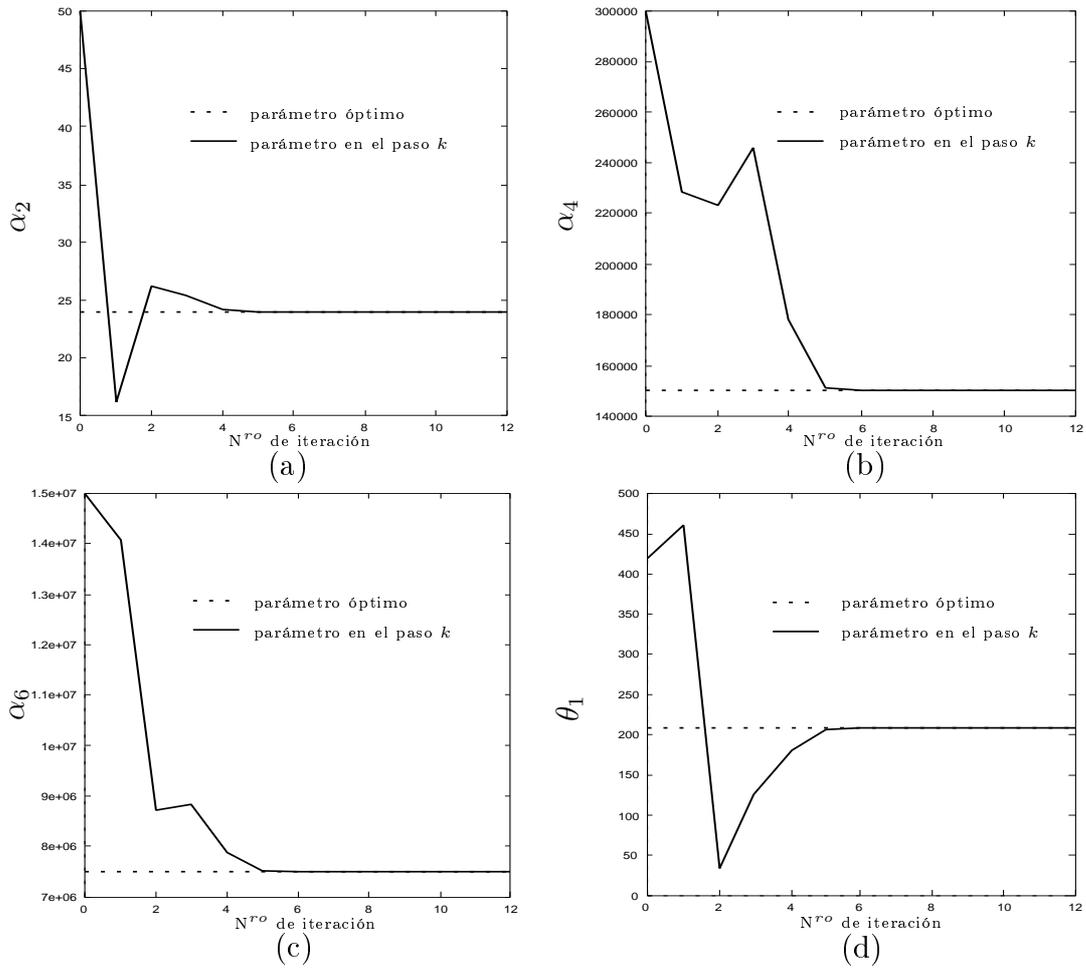


Figura 6.1: Evolución de las iteraciones para el Ejemplo 1.

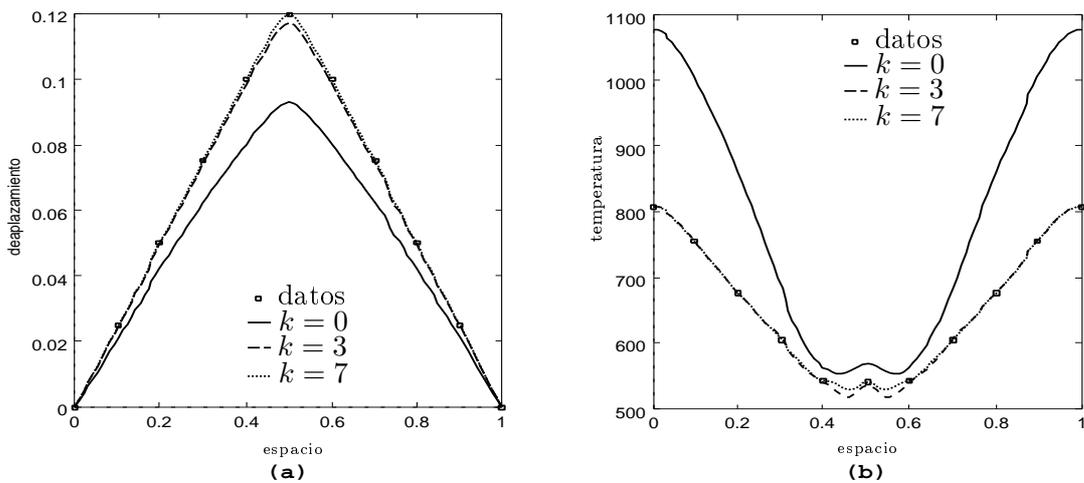


Figura 6.2: Desplazamiento (a) y Temperatura (b) para $T = 0.01$ y $q = q^k$, $k = 0, 3, 7$.

Ejemplo 2: *Datos Ruidosos*

Este ejemplo es análogo al Ejemplo 1, con la diferencia que ahora agregamos errores aleatorios a los datos observados con el objeto de simular errores de medición. Más precisamente, las observaciones se tomaron como $\hat{z}_i = \left\{ \begin{pmatrix} u(x_j, t_i; q^*) + r_{i,j} \\ \theta(x_j, t_i; q^*) + \tilde{r}_{i,j} \end{pmatrix} \right\}_{j=1}^9$, donde $r_{i,j}$ y $\tilde{r}_{i,j}$ son números al azar uniformemente distribuidos en $(-0.05\bar{u}, 0.05\bar{u})$ y $(-0.05\bar{\theta}, 0.05\bar{\theta})$, respectivamente, con

$$\bar{u} = \frac{1}{90} \sum_{i=1}^{10} \sum_{j=1}^9 |u(x_i, t_j; q^*)|, \quad \text{y} \quad \bar{\theta} = \frac{1}{90} \sum_{i=1}^{10} \sum_{j=1}^9 |\theta(x_i, t_j; q^*)|.$$

El estimador inicial es nuevamente $q^0 = (50, 3 \times 10^5, 15 \times 10^6, 420)$. Los resultados de las iteraciones se muestran en la Tabla 6.2 y en la Figura 6.3. La Figura 6.4a muestra una comparación entre $u(x, T; q^*)$ y $u(x, T; q^k)$ mientras que en la Figura 6.4b se muestra $\theta(x, T; q^*)$ y $\theta(x, T; q^k)$ para diferentes valores de k .

k	α_2	α_4	α_6	θ_1	$J(q^k)$
0	50.0000	300000	1.50000e+07	420.000	1987.240
1	16.5263	251533	1.43413e+07	450.975	604.570
2	26.7351	173032	7.92651e+06	77.3584	261.591
3	25.1282	223785	8.54573e+06	148.386	111.619
4	24.2875	176007	7.84280e+06	189.479	111.030
5	24.4436	183683	7.95702e+06	183.663	110.985
6	24.4070	180771	7.91592e+06	186.193	110.977
7	24.4184	181677	7.92857e+06	185.411	110.979
8	24.4151	181408	7.92483e+06	185.645	110.978
9	24.4161	181487	7.92593e+06	185.576	110.979
10	24.4158	181464	7.92560e+06	185.596	110.978
11	24.4159	181471	7.92570e+06	185.590	110.978
12	24.4159	181469	7.92567e+06	185.592	110.978
13	24.4159	181469	7.92568e+06	185.592	110.978

Tabla 6.2: Valores de los parámetros y del criterio de error en las iteraciones del algoritmo de cuasilinealización en el ejemplo 2.

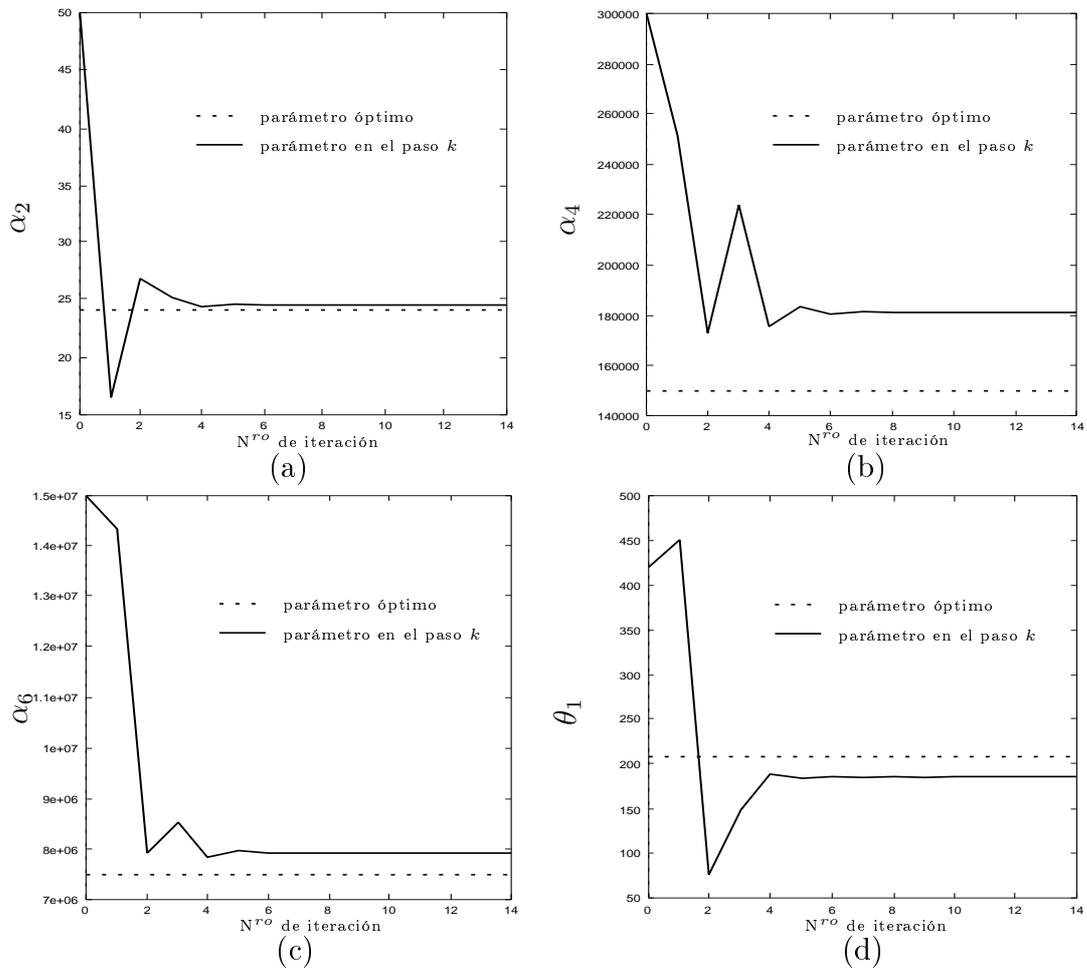


Figura 6.3: Evolución de las iteraciones para el Ejemplo 2.

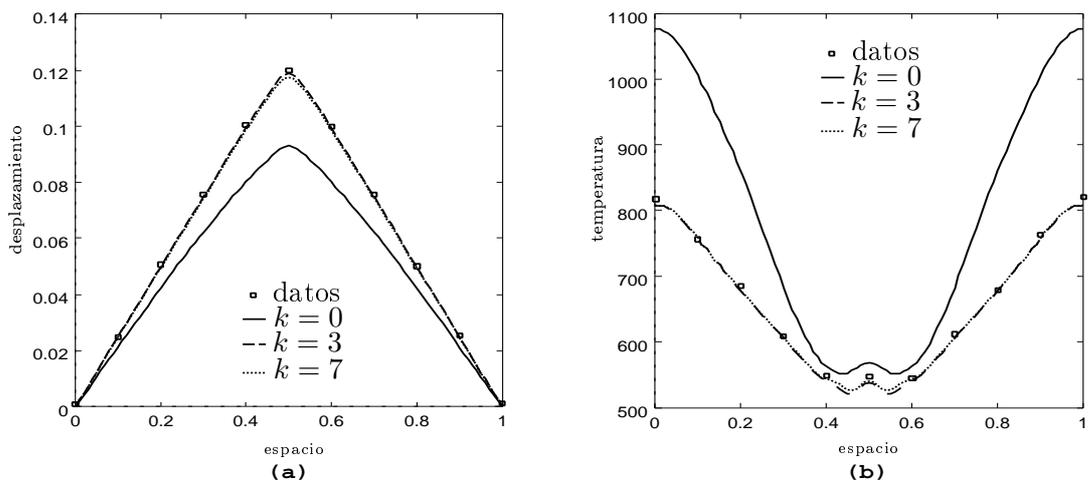


Figura 6.4: Desplazamiento (a) y Temperatura (b) para $T = 0.01$ y $q = q^k$, $k = 0, 3, 7$.

Ejemplo 3: *Comparación entre Métodos Directos e Indirectos*

En este ejemplo hemos resuelto el problema de identificación de dos maneras, primero usando un método indirecto y luego utilizando el algoritmo propuesto en la Sección 2.1. El propósito es ilustrar las distintas velocidades de convergencia de los dos métodos. Tomamos u_0 , v_0 , θ , f y g como en los ejemplos anteriores. El método indirecto consistió en aproximar la solución del sistema dinámico utilizando el algoritmo propuesto en el Capítulo 5 y aplicar el algoritmo de optimización de Hooke y Jeeves [8] para resolver el problema de optimización resultante. Se obtuvo \hat{z}_i como en el Ejemplo 1 y se comenzó con la estimación inicial $q^0 = (25, 2 \times 10^5, 9 \times 10^6, 220)$. Los resultados de las iteraciones se muestran en la Tabla 6.3.

k	α_2		α_4		α_6		θ_1	
	D	I	D	I	D	I	D	I
0	25	25	200000	200000	9e+06	9e+06	215	215
12	24.004	24.1250	149991	176000	7500020	8865000	207.999	202.1
40	24.004	24.9375	149991	161500	7500020	7537500	207.999	202.1
100	24.004	25.7344	149991	154000	7500020	6907500	207.999	202.1
500	24.004	24.8140	149991	149738	7500020	7333770	207.999	206.564
1000	24.004	24.4651	149991	149967	7500020	7462530	207.999	207.355
2000	24.004	24.1638	149991	150040	7500020	7499950	207.999	207.801
3000	24.004	24.0651	149991	150011	7500020	7500490	207.999	207.924

Tabla 6.3: Comparación de las velocidades de convergencia entre un método directo y uno indirecto.

Capítulo 7

Resolución Numérica del Problema de Control

En este capítulo utilizaremos el algoritmo de Cuasilinealización que se expuso en la Sección 2.1 para resolver numéricamente el problema de control terminal (C) planteado en la Sección 1.8, para el caso particular en el cual el operador lineal A y la función $F(t, z)$ están dados por (4.6)–(4.9).

No haremos un estudio completo ni exhaustivo del problema de control, simplemente ilustraremos otra utilización del método directo para la identificación de parámetros propuesto en el Capítulo 2.

Como notáramos en la Sección 1.8, si consideramos la función de control $u(t)$ como un parámetro distribuido temporalmente, además de espacialmente, el problema de control planteado en esa sección es un caso particular de un problema de identificación de parámetros. Tomaremos ventaja de esta similaridad entre los problemas para realizar lo que nos hemos propuesto en este capítulo: Hallar aproximaciones numéricas de los controles que logren cierto estado final a partir de un estado inicial.

Si bien es cierto que deberíamos hacer un estudio teórico de qué estados son alcanzables desde un estado inicial dado, con una cierta familia de funciones de control, también es cierto que los experimentos numéricos nos pueden ayudar a dilucidar cuál será la respuesta a este interrogante, o al menos a imaginarnos cuáles no podrán ser respuestas válidas, para después afrontar el problema de la demostración teórica del o de los

resultados imaginados.

El estudio teórico de la controlabilidad y de los estados alcanzables mediante una familia dada de controles se dejará para trabajos posteriores, en los cuales intentaremos también resolver problemas de control óptimo.

7.1 Hipótesis para la Convergencia del Algoritmo

En esta sección definiremos los operadores lineales A y B , y la función $F(t, z)$ de tal modo que la ecuación (1.12) sea equivalente al sistema (4.3)–(4.4). A tal efecto, mantenemos las definiciones del espacio de estados Z y del operador A de la Sección 4.3, y definimos el operador $F : [0, T] \times D(F) \rightarrow Z$ por

$$F(t, z) = \begin{pmatrix} 0 \\ f_2(t, z) \\ f_3(t, z) \end{pmatrix} \quad (7.1)$$

donde

$$\rho f_2(t, \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}) = \frac{d}{dx} [2\alpha_2 (\theta(x) - \theta_1) u'(x) - 4\alpha_4 u'(x)^3 + 6\alpha_6 u'(x)^5], \quad (7.2)$$

$$C_v f_3(t, \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}) = 2\alpha_2 \theta(x) u'(x) v'(x) + \beta \rho v'(x)^2. \quad (7.3)$$

y $D(F) = H^2(0, 1) \times H^1(0, 1) \times H^1(0, 1)$. Llamaremos \mathcal{U} al espacio de controles admisibles

$$\mathcal{U} \doteq H^1(0, T : H^2(0, 1) \times H^2(0, 1))$$

y definimos el operador lineal $B : \mathcal{U} \rightarrow L^\infty(0, T : Z_\delta)$ por

$$BU(t) = B \begin{pmatrix} f \\ g \end{pmatrix} (t) = \begin{pmatrix} 0 \\ \frac{1}{\rho} f(t) \\ \frac{1}{C_v} g(t) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \frac{1}{\rho} & 0 \\ 0 & \frac{1}{C_v} \end{pmatrix} \begin{pmatrix} f(t) \\ g(t) \end{pmatrix}, \quad t \in [0, T].$$

De ahora en adelante dejamos fijo $\delta \in (\frac{3}{4}, 1)$. Sabemos por los resultados de la Sección 4.4 que si f y g satisfacen la hipótesis **(H1)** (ver página 62), entonces la ecuación (1.12) tiene solución para todo $z_0 \in D(I - A)^\delta$. En particular para todo $U \in \mathcal{U}$ la ecuación (1.12) tiene solución única a la que llamaremos $z(t; U) = (u(t; U), v(t; U), \theta(t; U))$.

Según los resultados de la Sección 2.2 una condición suficiente para la convergencia del algoritmo de identificación, es que la derivada de Fréchet $D_U z(t; U)$ sea localmente Lipschitz continua con respecto a la variable U . Teniendo en cuenta los resultados del Capítulo 3, una condición suficiente para que esto ocurra es que el mapeo $\tilde{F} : \mathcal{U} \times [0, T] \times D(F) \rightarrow L^\infty(0, T : Z_\delta)$ definido por $\tilde{F}(U, t, z) = F(t, z) + BU(t)$ sea derivable Fréchet con respecto a U y a z y que las derivadas $D_U \tilde{F}(U, t, z)$ y $D_z \tilde{F}(U, t, z)$ sean funciones localmente Lipschitz continuas de $\mathcal{U} \times L^\infty(0, T : Z_\delta)$ en $L^\infty(0, T : \mathcal{L}(\mathcal{U}, Z))$ y en $L^\infty(0, T : \mathcal{L}(Z_\delta, Z))$, respectivamente, con respecto a ambas variables.

Estas hipótesis se cumplen como consecuencia del Teorema 6.3 y del hecho que B es un operador lineal y acotado de \mathcal{U} en $L^\infty(0, T : Z_\delta)$.

7.2 Experimentos Numéricos

En esta sección utilizaremos el algoritmo de identificación basado en cuasilinealización para obtener aproximaciones numéricas de las funciones de control f y g que en el problema de valores iniciales y de frontera (4.3)–(4.4) logran un desplazamiento \hat{u} y una temperatura $\hat{\theta}$ deseados, en tiempo final T dado. En todos los experimentos partiremos del estado inicial consistente de desplazamiento y velocidad nulos y temperatura constante de 200 K, e intentaremos hallar las funciones que mejor aproximen el estado final

$$\hat{u}(x) = \begin{cases} 0.1x & \text{si } 0 \leq x \leq 0.5 \\ 0.1(1-x) & \text{si } 0.5 \leq x \leq 1 \end{cases}, \quad \hat{\theta}(x) = 200 \text{ K}, \quad 0 \leq x \leq 1$$

en $T = 0.01$. Para ello definimos $J(U) = C_d J_d(U) + C_\theta J_\theta(U)$, donde

$$J_d(U) = \sum_{i=0}^{100} (u(0.01 i, T; U) - \hat{u}(0.01 i))^2$$

y

$$J_\theta(U) = \sum_{i=0}^{100} (\theta(0.01 i, T; U) - \hat{\theta}(0.01 i))^2,$$

y elegimos $C_d = 1000$ y $C_\theta = 1$ para equilibrar o balancear el funcional de modo que se le dé la misma importancia al desplazamiento y a la temperatura. En todos los casos, iniciaremos las iteraciones con las funciones de control $f \equiv g \equiv 0$.

Experimento 1.

En este experimento, deseamos ver qué resultados podemos obtener suponiendo que f y g son constantes a trozos en los rectángulos $R_1 = [0, 1/3) \times [0, T]$, $R_2 = [1/3, 2/3) \times [0, T]$ y $R_3 = [2/3, 1] \times [0, T]$ (ver Figura 7.1). Los valores de f y g en cada uno de estos

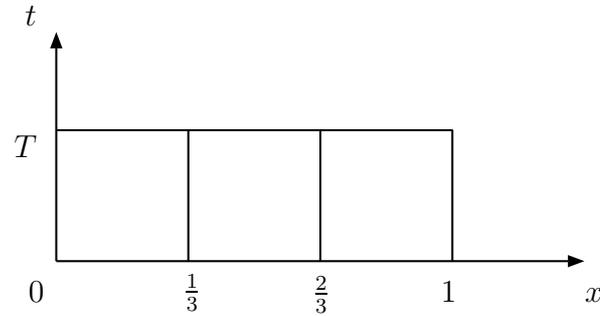


Figura 7.1: Rectángulos donde f y g son constantes en el Experimento 1.

rectángulos para cada iteración pueden observarse en la Tabla 7.1, mientras que en la Figura 7.2 se ve el desplazamiento y la temperatura finales para algunas iteraciones.

k	f			g		
	R_1	R_2	R_3	R_1	R_2	R_3
0	0	0	0	0	0	0
1	3.84051	10.6160	3.84051	-1.38e-14	1.08e-14	2.07e-14
2	3.78732	11.0925	4.12042	-8.63138	-7.15525	-11.67290
3	3.74838	11.3075	4.00717	-8.79441	-8.12705	-9.98224
4	3.74878	11.4273	3.89327	-8.91942	-8.61023	-9.52561
5	3.75231	11.4838	3.79721	-8.97732	-8.94461	-9.07119
6	3.75572	11.5135	3.72982	-9.00843	-9.15649	-8.76404
7	3.75829	11.5306	3.68533	-9.02664	-9.29116	-8.56287
8	3.76007	11.5411	3.65646	-9.03790	-9.37726	-8.43289
9	3.76127	11.5477	3.63774	-9.04504	-9.43274	-8.34878
10	3.76207	11.5519	3.62557	-9.04963	-9.46871	-8.29412

Tabla 7.1: Valores de las funciones de control f y g para el **Experimento 1** en las primeras 10 iteraciones.

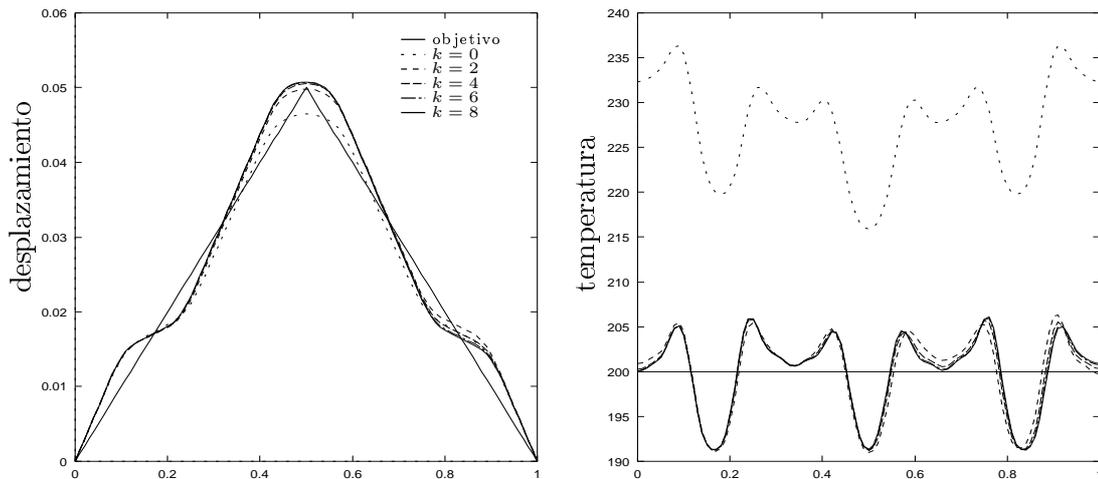


Figura 7.2: Desplazamiento y temperatura en $T = 0.01$ para $U = U^k$, con $k = 0, 2, 4, 6, 8$ en el Experimento 1.

Experimento 2.

En este experimento, en un intento por mejorar la aproximación a \hat{u} y a $\hat{\theta}$ aumentamos a 5 el número de rectángulos donde las funciones de control son constantes, dividiendo nuevamente sólo el intervalo espacial $[0, 1]$ en 5 subintervalos de longitud $1/5$ (ver Figura 7.3). En la Figura 7.4 podemos observar cuál es el desplazamiento y la temperatura

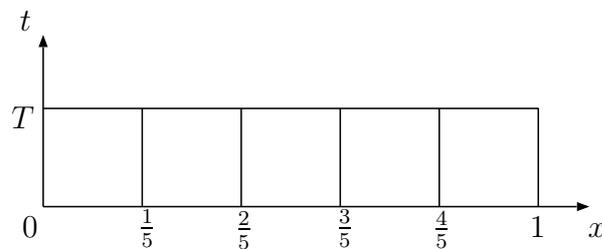


Figura 7.3: Rectángulos donde f y g son constantes en el Experimento 2.

final con los controles obtenidos en algunas de las iteraciones. Como podemos observar en la figura la aproximación al desplazamiento deseado es mejor, pero la temperatura sigue presentando oscilaciones pronunciadas.

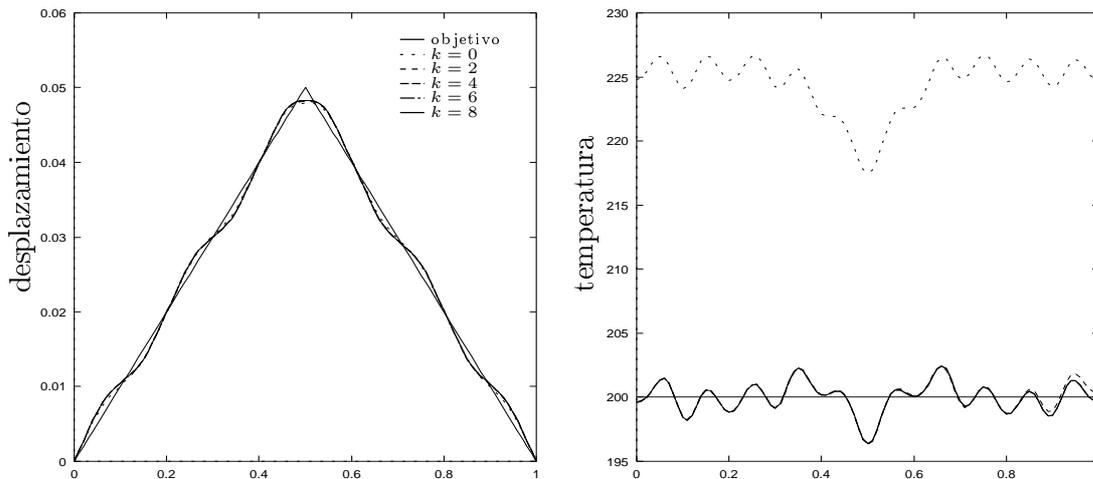


Figura 7.4: Desplazamiento y temperatura en $T = 0.01$ para $U = U^k$, con $k = 0, 2, 4, 6, 8$, en el Experimento 2.

Experimento 3.

En este experimento, en un intento por mejorar la aproximación que se había obtenido en el Experimento 1, y evitar a la vez las oscilaciones en la temperatura que se obtuvieron en el Experimento 2, aumentamos a 6 el número de rectángulos donde las funciones de control son constantes, pero esta vez dividiendo el intervalo espacial en 3 subintervalos y el temporal en 2 (ver Figura 7.5). En la Figura 7.6 podemos observar cuál es el despla-

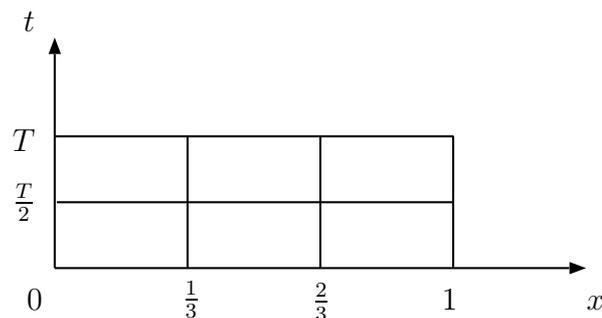


Figura 7.5: Rectángulos donde f y g son constantes en el Experimento 3.

zamiento y la temperatura final con los controles obtenidos en algunas de las iteraciones. Como podemos observar en la Figura 7.6, tanto la aproximación al desplazamiento como la aproximación a la temperatura deseados son similares a las del Experimento 2, pero

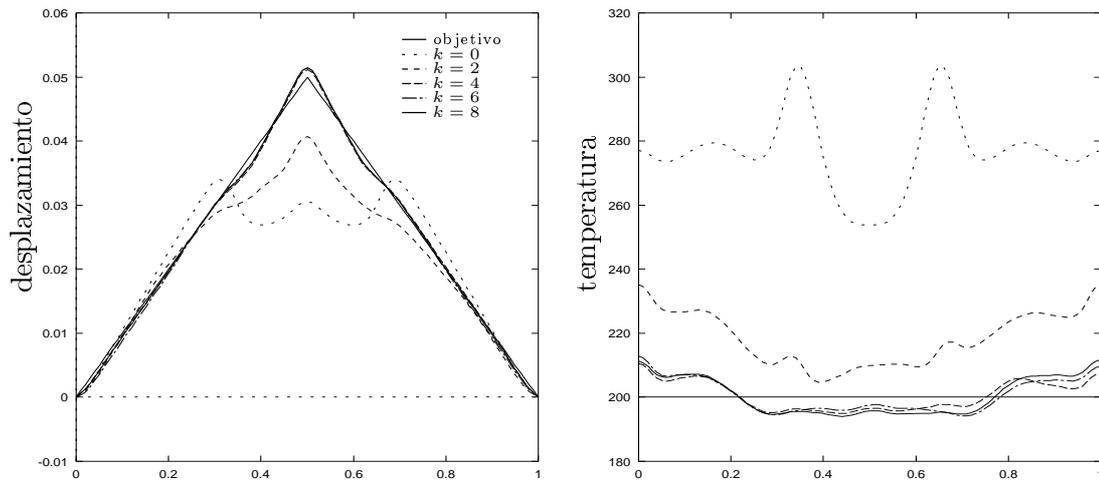


Figura 7.6: Desplazamiento y temperatura en $T = 0.01$ para $U = U^k$, con $k = 0, 2, 4, 6, 8$, en el Experimento 3.

las oscilaciones de la temperatura disminuyen considerablemente. Esto nos induce a creer que una estrategia de refinamiento de la malla debe tener en cuenta un refinamiento en ambas variables y no sólo en la variable espacial. Esto guarda una fuerte relación con el hecho que el problema es no lineal, pues si el problema fuera lineal, la discretización temporal no tendría ningún efecto.

Capítulo 8

Conclusiones y Planes para el Futuro

En este trabajo se ha desarrollado un algoritmo para la identificación de parámetros en un problema de Cauchy abstracto no lineal en un espacio de Banach Z , de la forma

$$\begin{cases} \frac{dz(t)}{dt} = Az(t) + F(q, t, z(t)), & t \in (0, T) \\ z(0) = z_0. \end{cases}$$

donde A es el generador infinitesimal de un semigrupo analítico en Z y la parte no lineal F depende del parámetro q . El mismo algoritmo y los mismos teoremas de convergencia permanecen válidos para cualquier problema de Cauchy abstracto que dependa de un parámetro q , en particular para el caso en que también el operador lineal A depende del parámetro q . Sin embargo, en el caso en que A y F dependen ambos del parámetro q , es más difícil de probar que se cumplen las hipótesis para la convergencia del algoritmo. Uno de los problemas que quedan abiertos para futuras investigaciones es el de encontrar condiciones suficientes sobre $A(q)$ y $F(q, t, z)$ para que exista la derivada de Fréchet de la solución $z(t; q)$ con respecto al parámetro q y que ésta sea Lipschitz continua también con respecto al parámetro q .

Por otro lado, el algoritmo para la identificación de parámetros fue desarrollado para el caso en que q se encuentra en un espacio de dimensión infinita, y los teoremas de convergencia del mismo permanecen válidos en este caso. Sin embargo, para que el algoritmo pueda ser implementado computacionalmente es necesario aproximar el espacio de parámetros \tilde{Q} por un espacio de dimensión finita \tilde{Q}^s . Pero al hacer esto, los teoremas

de convergencia dejan de ser válidos, pues sólo hemos demostrado la convergencia en el caso en que se sabe a priori que el parámetro óptimo $q^* \in \tilde{Q}^s$ para algún s . Un problema que nos proponemos para el futuro es el de obtener resultados de convergencia cuando el espacio de estados \tilde{Q} se aproxima por espacios de dimensión finita \tilde{Q}^s , y no se supone a priori que q^* pertenezca a algún \tilde{Q}^s , como así también obtener cotas para el error.

El algoritmo de identificación ha sido utilizado para identificar los parámetros α_2 , α_4 , α_6 y θ_1 del problema de valores iniciales y de frontera (4.3)–(4.4). Si bien los experimentos numéricos se han realizado utilizando datos artificiales, se están realizando esfuerzos por conseguir mediciones que permitirán probar el algoritmo contra datos experimentales.

También se ha visto que el algoritmo basado en cuasilinealización para la identificación de parámetros aquí desarrollado es de utilidad para la resolución numérica de un problema de control terminal, asociado al sistema de ecuaciones que modela el comportamiento de un sólido unidimensional con memoria de forma. Esta aplicación se realizó solamente a modo de ejemplo, pero permite imaginar respuestas al problema de la controlabilidad. El problema de la controlabilidad, como también el problema de la existencia de controles óptimos asociados a las ecuaciones diferenciales mencionadas en este párrafo es un problema que nos proponemos estudiar en breve. Otro problema que nos proponemos estudiar es el desarrollo y aplicación de algún algoritmo basado en cuasilinealización para la resolución numérica de problemas de control óptimo.

Bibliografía

- [1] M. ACHENBACH and I. MULLER, A Model for Shape Memory, *J. Physique C4* **43**, No. 12, Suppl. (1982), pp. 163–167.
- [2] M. ACHENBACH and I. MULLER, Creep and Yield in Martensitic Transformations, *Ingenieur-Archiv.* **53** (1983), pp. 73–83.
- [3] M. ACHENBACH and I. MULLER, Shape Memory as a Thermally Activated Process, in “Plasticity Today: Modeling Methods and Applications”, pp. 515–531, Sawczuk and Bianchi.
- [4] M. AHLERS, Materiales con Memoria de Forma, *Rev. Ciencia Hoy*, Vol. 2, No. 2, marzo–abril 1991, pp. 24–33.
- [5] G. F. ANDREASEN, and R. E. MORROW, Laboratory and Clinical Analyses of NiTiNol Wire, *Am. J. Orthod.* **73**, No. 2, pp. 142–151.
- [6] H. T. BANKS, and G. M. GROOME, JR., Convergence Theorems for Parameter Estimation by Quasilinearization, *J. Math. Anal. Appl.* **42** (1973), pp. 91–109.
- [7] H. T. BANKS, and K. KUNISCH, “Estimation Techniques for Distributed Parameter Systems”, *Systems and Control: Foundations and Applications*, No. 1, Birkhäuser, Boston, 1989.
- [8] M. S. BAZARAA, and C. M. SHETTY, “Nonlinear Programming. Theory and Algorithms”, John Wiley and Sons, 1979.

- [9] R. E. BELLMAN, and R. E. KALABA, “Quasilinearization and Nonlinear Boundary Value Problems”, American Elsevier, New York, 1965.
- [10] A. BENSOUSSAN, G. DA PRATO, M. DELFOUR and S. MITTER, “Representation and Control of Infinite Dimensional Systems”, Vol. I y II, Birkäuser, 1992.
- [11] D. W. BREWER, The Differentiability with respect to a Parameter of the Solution of a Linear Abstract Cauchy Problem, *SIAM J. Math. Anal.* **13** (1982), pp. 607–620.
- [12] D. W. BREWER, J. A. BURNS, and E. M. CLIFF, Parameter Identification for an Abstract Cauchy Problem by Quasilinearization, *Quart. Math. Anal.* **51** (1993), pp. 1–22. Aparecido en octubre de 1989 como NASA Report No. 89-75.
- [13] M. BROKATE and J. SPREKELS, “Hysteresis and Phase Transitions”, Applied Mathematical Sciences, Vol. 121, Springer, New York 1996.
- [14] A. BROWN, Materials Get Smarter, *Aerospace America*, March 1990, pp. 30–36.
- [15] J. BUELL, and R. E. KALABA, Quasilinearization and the Fitting of Nonlinear Models of Drug Metabolism to Experimental Kinetic Data, *Math. Biosciences* **5** (1969), pp. 121–132.
- [16] J. A. BURNS, and E. M. CLIFF, An Abstract Quasilinearization Algorithm for Estimating Parameters in Hereditary Systems, *IEEE Trans. Automat. Contr.* **25** (1980), pp.126–129.
- [17] J. A. BURNS and S. KANG, A Control Problem for Burgers’ Equation with Bounded Input/Output, *Nonlinear Dynamics* **2** (1991), pp. 235–262.
- [18] L. S. CASTELMAN, S. M. MOTZKIN, F. P. ALICANDRI and V. L. BONAWIT, Biocompatibility of Nitinol Alloy as an Implant Material, *J. Biomedical Materials Res.* **10** (1976), pp. 614–627.

-
- [19] P. J. CHEN, and M. E. GURTIN, On a Theory of Heat Conduction Involving Two Temperatures, *Z. Angew. Math. Phys.* **19** (1968), pp. 614–627.
- [20] L. G. CLARK, and J. S. GIBSON, Sensitivity Analysis for a Class of Evolution Equations, *J. Math. Anal. Appl.* **58** (1977), pp. 22–31.
- [21] R. F. CURTAIN, and A. J. PRITCHARD, “Functional Analysis in Modern Applied Mathematics”, Academic Press, London, 1977.
- [22] H. FUNAKUBO (Ed.), “Shape Memory Alloys”, Precision Machinery and Robotics, Vol. 1, Translated from the Japanese by J. B. Kennedy, Gordon and Breach Science Publishers, 1987.
- [23] F. FALK, Landau Theory and Martensitic Phase Transitions, *Journal de Physique*, Colloque C4, Supplément au No. 12, Tome 43, décembre 1982, pp. C4 3–15.
- [24] F. FALK, Model Free Energy, Mechanics and Thermodynamics of Shape Memory Alloys, *Acta Metallurgica* **28** (1980), pp. 1773–1780.
- [25] F. FALK, One Dimensional Model of Shape Memory Alloys, *Arch. Mech.* **35** (1983), pp. 63–84.
- [26] A. FRIEDMAN, and J. SPREKELS, Steady-States of Austenitic-Martensitic Domains in the Ginzburg-Landau Theory of Shape Memory Alloys, *Continuum Mech. Thermodyn.*, Vol. 2 (1990), pp. 199–213.
- [27] M. E. GURTIN, On the Thermodynamics of Materials with Memory, *Arch. Rational Mech. Anal.* **28** (1968), pp. 40–50.
- [28] P. W. HAMMER, Parameter Identification in Parabolic Partial Differential Equations using Quasilinearization, Ph. D. Thesis, ICAM Report 90-07-01, Interdisciplinary Center for Applied Mathematics, Virginia Polytechnic Institute and State University.

-
- [29] D. HENRY, “Geometric Theory of Semilinear Parabolic Equations”, Lecture Notes in Mathematics No. 840, Springer-Verlag, 1989.
- [30] K. H. HOFFMANN and S. ZHENG, Uniqueness for Nonlinear Coupled Equations Arising from Alloy Mechanism, Preprint 118, Institut Für Mathematik, Ausburg, 1986.
- [31] K. H. HOFFMANN and S. ZHENG, Uniqueness for Structural Phase Transitions in Shape Memory Alloys, *Math. Methods Appl. Sci.* **10** (1988), pp. 145–151.
- [32] J. L. LIONS and E. MAGENES, “Non-Homogeneous Boundary Value Problems and Applications”, vol. 1, Springer, Berlin, 1972.
- [33] R. LOHMAN AND I. MULLER, A Model for the Qualitative Description of Martensitic Transformations in Memory Alloys, in “Phase Transformations” (E. Aifantis and J. Gittus, Eds.), pp. 55–75.
- [34] E. C. MALLOY, Nitinol Provides Shape Memory Capabilities, *On the Surf. Mag.* (22 June 1990), pp. 1–3.
- [35] I. MULLER, A Model for a Body with Shape Memory, *Arch. Rational Mechanics Anal.*, **70** (1979), pp. 61–67.
- [36] I. MULLER AND P. VILLAGGIO, A Model for an Elastic-Plastic Body, *Arch. Rational Mechanics Anal.* **65**, No. 1 (1977), pp. 25–46.
- [37] I. MULLER AND K. WILMANSKI, A Model for Phase Transitions in Pseudoelastic Bodies, *Il Nuovo Cimento* **57b**, No. 2 (1980), pp. 283–318.
- [38] M. NIEZGODKA and J. SPREKELS, Existence of Solutions for a Mathematical Model of Structural Phase Transitions in Shape Memory Alloys, *Math. Methods Appl. Sci.* **10** (1988), pp. 197–223.

-
- [39] M. NIEZGODKA and J. SPREKELS, Convergent Numerical Approximations of the Thermomechanical Phase Transitions in Shape Memory Alloys, *Numer. Math.* **58** (1991), pp. 759–778.
- [40] M. NIEZGODKA and J. SPREKELS, Global Solution to a Model of Structural Phase Transitions in Shape Memory Alloys, *J. Math. Anal. Appl.* **130** (1988), pp. 39–54.
- [41] J. W. NUNZIATO, On Heat Conduction in Materials with Memory, *Quart. Appl. Math.* **29** (1971), pp.187–204.
- [42] A. PAZY, “Semigroups of Linear Operators and Applications to Partial Differential Equations”, corrected 2nd printing, Springer-Verlag, New York/Berlin, 1983.
- [43] C. A. ROGERS, Intelligent Material Systems: The Dawn of a New Materials Age, *J. Intelligent Materials, Systems and Structures* **4** (1993), pp. 4–12.
- [44] C. A. ROGERS, C. LIANG and J. JIA, Behavior of Shape Memory Alloy Reinforced Composites, Part I: Model Formulations and Control Problems”, in *30th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Mobile, AL, April 1989.
- [45] M. A. SCHMERLING, M. A. WILKOV, A. E. SANDERS and J. E. WOOSLEY, A Proposed Medical Application of the Shape Memory Effect: A NiTi Harrington Rod for the Treatment of Scoliosis, in “Shape Memory Effects in Alloys”, Jeff Perkins (ed.), Plenum Press, 1975, pp. 563–574.
- [46] R. D. SPIES, A state-space approach to a one-dimensional mathematical model for the dynamics of phase transitions in pseudoelastic materials, *J. Math. Anal. Appl.* **190** (1995), pp. 58–100.
- [47] R. D. SPIES, Results on a mathematical model of thermomechanical phase transitions in shape memory materials, *J. Smart Materials Structures* **3** (1994), pp. 459–469.

-
- [48] J. SPREKELS, Automatic Control of One-Dimensional Thermomechanical Phase Transitions, *in* "Mathematical Models for Phase Change Problems", International Series of Numerical Mathematics, Vol. 88, pp. 89–98, Birkhäuser, Basel, 1989.
- [49] J. SPREKELS, Global Existence for Thermomechanical Processes with Nonconvex Free Energies of Ginzburg-Landau Form, *J. Math. Anal. Appl.* **141** (1989), pp. 333–348.
- [50] J. SPREKELS AND S. ZHENG, Global Solutions to the Equations of a Ginzburg-Landau Theory for Structural Phase Transitions in Shape Memory Alloys, *Physica D* **39** (1989), pp. 59–76.
- [51] H. TRIEBEL, "Interpolation Theory, Function Spaces, Differential Operators", North-Holland, Amsterdam, 1978.
- [52] K. WILMANSKY, Propagation of the Interface in the Stress-Induced Austenite-Martensite Transformation, *Ingenieur-Archiv.* **53** (1983), pp. 291–301.
- [53] S. ZHENG, Global Solutions to the Thermomechanical Equations with Non-convex Landau-Ginzburg Free Energy, *J. Appl. Math. Phys. (ZAMP)* (Jan. 1989), pp. 111–127.