UNIVERSIDAD POLITECNICA DE VALENCIA

DEPARTAMENTO DE MATEMATICA APLICADA

SOLUCIONES NUMERICAS CONTINUAS DE ECUACIONES DIFERENCIALES MATRICIALES CON COTAS DE ERROR A PRIORI

TESIS DOCTORAL

PRESENTADA POR:

ENRIQUE PONSODA MIRALLES

DIRIGIDA POR:

LUCAS JÓDAR SÁNCHEZ

VALENCIA, 1994.

LUCAS JODAR SANCHEZ, PROFESOR CATEDRATICO DEL DEPARTAMENTO DE MATEMATICA APLICADA DE LA UNIVERSIDAD POLITECNICA DE VALENCIA,

CERTIFICA: Que la presente memoria, "SOLUCIONES NUMERICAS CONTINUAS DE ECUACIONES DIFERENCIALES MATRICIALES CON COTAS DE ERROR A PRIORI" ha sido realizada bajo su dirección en el Departamento de Matemática Aplicada de la Universidad Politécnica de Valencia, por D. Enrique Ponsoda Miralles, y constituye su tesis para optar al Grado de Doctor en Ciencias Matemáticas.

Y para que conste, en cumplimiento de la legislación vigente, presento ante el Departamento de Matemática Aplicada de la Universidad Politécnica de Valencia la referida tesis, firmando el presente certificado en Valencia, a 4 de Noviembre de 1993.

Fdo. Lucas Jódar Sánchez.

INDICE

CAPITULO I

MOTIVACION Y PRELIMINARES

I. A. INTRODUCCIÓN.

El problema de la computación de soluciones aproximadas de ecuaciones diferenciales matriciales aparecen directamente en la formulación de modelos matemáticos en problemas tecnológicos, como por ejemplo las ecuaciones de tipo Riccati en teoría de control [REI1], [REI2], [JOD3]; al resolver problemas escalares de ecuaciones en derivadas parciales mediante el método de semidiscretización [REKT], [VEMU]; o incluso en la solución numérica de problemas de contorno vectoriales mediante el método del tiro (shooting), [MARZ].

La vectorización de un problema matricial es un recurso demasiado imperfecto para que no merezca ser discutido actualmente. Entre los defectos del método de vectorización, es decir, del enfoque vectorial de un problema matricial, se encuentran los siguientes:

(a) se aumenta el volúmen computacional,

(b) se pierde el significado físico de las magnitudes que aparecen en la formulación original del problema,

(c) en el análisis del error, las constantes que aparecen en las cotas del error suelen estar dilatadas, con lo que la calidad de las cotas empeora (gravemente en ocasiones),

(d) se desperdicia el uso de lenguajes algebraicos simbólicos que permiten trabajar con matrices.

Estas razones justifican y motivan el uso de métodos matriciales para resolver problemas matriciales, que es el ánimo de este proyecto de tesis. Los bloques temáticos de esta memoria son los siguientes:

(I) Soluciones numéricas contínuas de problemas de valores iniciales matriciales.

(II) Aplicaciones de los métodos de valores iniciales matriciales a la computación aproximada de funciones inversas.

(III) Soluciones analítico-numéricas de sistemas de ecuaciones en derivadas parciales de segundo órden.

En el bloque I se construyen soluciones numéricas discretas de problemas de valores iniciales matriciales en una malla de puntos mediante el uso de métodos multipaso matriciales. Posteriormente, utilizando funciones B-splines matriciales se construyen vía interpolación, aproximaciones numéricas contínuas. Las cotas de error de los métodos multipaso se transporta para la obtención de cotas de error de la aproximación contínua de todo el intervalo de definición.

Dentro del bloque I se estudian diferentes tipos deecuaciones diferenciales matriciales con coeficientes variables, así, en el capítulo 2 se estudian ecuaciones de tipo lineal

$$X'(t) = A(t)X(t) + X(t)B(t) + L(t) \ ; \quad X(0) = C \ ; \quad 0 \le t \le b$$

en el capítulo 3 tratamos el caso cuadrático de tipo Riccati

$$W'(t) = C(t) - D(t)W(t) - W(t)A(t) - W(t)B(t)W(t) \ ; \quad W(0) = W_0$$
$$0 \le t \le b$$

y en el capítulo 4 se trata el problema general

$$Y'(t) = f(t,Y(t)) \; ; \quad Y(0) = \Omega \; ; \quad 0 \leq t \leq b$$

donde $f: [0,b] \times \mathbb{C}^{r\times q} \longrightarrow \mathbb{C}^{r\times q}$ es acotada, contínua y satisface la condición de Lipschitz

$$\| f(t,P) - f(t,Q) \| \leq L \| P - Q \|$$

El bloque temático II se incluye en el capítulo 4 y esencialmente se apoya en la demostración de que bajo ciertas condiciones, la función inversa es localmente la solución de un problema diferencial matricial de valores iniciales.

El bloque temático III trata de la construcción de soluciones analítico-numéricas de problemas mixtos para sistemas de ecuaciones en derivadas parciales del tipo

$$u_t(x,t) - A\, u_{xx}(x,t) - B\, u(x,t) = G(x,t) \; ; \quad 0 \leq x \leq p \; ; \quad t \geq 0$$

donde A y B son matrices cuadradas complejas.

Este bloque se incluye en el capítulo 5, produciendo los siguientes niveles de respuesta:

(i) Obtención de una solución en serie exacta mediante un método de separación de variables matricial.

(ii) Truncación de la serie infinita atendiendo a la obtención de cotas de error prefijadas en dominios acotados predeterminados.

El caso donde B=0 ha sido tratado en la tesis de Matilde Legua [LEGU]. El problema aquí tratado para $B \neq 0$ involucra el estudio de problemas no triviales de álgebra lineal como son la localización del espectro de un haz de matrices $A+\lambda B$. Entonces se puede acotar la evolución de la norma de funciones matriciales del tipo $\| \exp (t (A+\lambda B)) \|$ que permite la viabilidad del desarrollo del capítulo.

En el capítulo 1 se incluyen algunos preliminares que facilitan la presentación de los resultados desarrollados posteriormente.

La clasificación temática de esta memoria atendiendo a la clasificación de la A.M.S. del año 1991 es la siguiente:

15A60, 15A18, 15A22, 34A34, 34A50, 35C10, 65F15, 65M15.

I. B. Cálculo Matricial.

I. B. (i). GENERALIDADES.

Si B es una matriz de $\mathbb{C}^{r\times r}$, representaremos por $\sigma(B)$ al conjunto de todos lo valores propios de B. Al radio espectral de B lo denotaremos por $\rho(B)$. Recordemos que la matríz B se dice que es de clase M si todos los valores propios z de B, tales que $|z|$ coincide con el radio espectral, tienen unos *bloques de Jordan* correspondientes de dimensión 1x1, [ORT1, pág. 24].

A la matriz identidad de $\mathbb{C}^{r\times r}$ la representaremos por I.

Si C es una matríz de $\mathbb{C}^{r\times q}$, denotaremos mediante $\|C\|$ a la 2-norma, definida como la raíz cuadrada del máximo del conjunto:

$$\left\{ |z|;\ z \text{ es valor propio de } C^H C \right\}$$

donde C^H representa la matriz conjugada y traspuesta de C, ver [ORT2, pág. 41].

La ∞-norma de C, que representaremos mediante $\|C\|_\infty$ se define:

$$\|C\|_\infty = \max_{1\le i\le r} \sum_{j=1}^{q} |c_{ij}|$$

Por [GOLU, pág. 14-15] tenemos que:

$$q^{-(1/2)} \| C \|_{\infty} \le \| C \| \le r^{(1/2)} \| C \|_{\infty} \qquad (1.3)$$

$$\max_{\substack{1 \le i \le r \\ 1 \le j \le q}} | c_{ij} | \le \| C \| \le (rq)^{(1/2)} \max_{\substack{1 \le i \le r \\ 1 \le j \le q}} | c_{ij} | \qquad (1.4)$$

siendo $C \in \mathbb{C}^{r\times q}$.

Si P y Q son matrices de $\mathbb{C}^{r\times r}$ y P es inversible, entonces por el *lema de Banach* [GOLU, pág. 28], si:

$$\| P - Q \| < \| P^{-1} \|^{-1}$$

se tiene que Q es inversible y:

$$\| P^{-1} - Q^{-1} \| < \| P \| \| Q \| \| P - Q \|$$

Teniendo en cuenta la desigualdad:

$$\| Q^{-1} \| \le \| Q^{-1} - P^{-1} \| + \| P^{-1} \|$$

podemos escribir:

$$\| P^{-1} - Q^{-1} \| \leq \left(1 - \| P^{-1} \| \| P - Q \|\right)^{-1} \| P^{-1} \|^2 \| P - Q \| \quad (1.5)$$

Recordemos también dos resultados que serán utilizados a lo largo de esta memoria. En primer lugar, la *fórmula de Sherman-Morrison-Woodbury* [ORT2, pág. 50], que establece que dada una matríz inversible $A \in \mathbb{R}^{nxn}$ y las matrices $U, V \in \mathbb{R}^{nxm}$, $m \leq n$, entonces la matríz $A+UV^T$ es inversible si y sólo si $I+V^TA^{-1}U$ es inversible, cumpliéndose además que:

$$\left(A + UV^T \right)^{-1} = A^{-1} - A^{-1} U \left(I + V^TA^{-1}U \right)^{-1} V^T A^{-1}$$

El otro resultado al que hacíamos referencia, trata sobre una *estimación de la norma de la inversa de una matríz* [FREE, pág. 255], que puede enunciarse como sigue:

Sea $A(x)=A(x_1,x_2,...,x_m)$ una función matricial real de tamaño *nxn*, cuyas entradas son funciones definidas en un dominio cerrado $D \subset E^m$ que contiene al orígen, de modo que $A(0)$ es conocida. Supongamos que $A(x)$ no es singular en D y tiene primera derivada parcial contínua con respecto a $x_1,x_2,...,x_m$ en D, con:

$$\left\| \frac{\partial A(x)}{\partial x_i} \right\| \leq M_i \,, \qquad i = 1,\, 2, \ldots,\, m,$$

donde $\| \cdot \|$ representa una norma matricial particular. Entonces, para cualquier rectángulo m-dimensional $D' \subset D$ de la forma:

$$D' = \left\{ -\ell_i \leq x_i \leq \ell_i \,, \quad i = 1,\, 2, \ldots,\, m \right\},$$

con:

$$\left\| \left[A(0) \right]^{-1} \right\| \sum_{i=1}^{m} M_i \, \ell_i < 1$$

tenemos:

$$\left\| \left[A(x) \right]^{-1} \right\| \leq \left\| \left[A(0) \right]^{-1} \right\| \left[1 - \left\| \left[A(0) \right]^{-1} \right\| \sum_{i=1}^{m} M_i \ell_i \right]^{-1} ; \quad x \in D'$$

Por último, señalemos que si x es un número real, representaremos por [x] a su parte entera.

I. B. (ii). POLINOMIOS MATRICIALES.

Para una mayor claridad en la presentación de los resultados de la sección siguiente relativos a métodos multipasos matriciales, recordemos algunos conceptos y propiedades sobre polinomios matriciales, cuyas demostraciones pueden encontrarse en [GOH1] y [GOH2].

Sea $L(z)$ el polinomio matricial:

$$L(z) = z^k I + z^{k-1} A_{k-1} + \dots + A_0 \qquad (1.6)$$

donde $A_i \in \mathbb{C}^{r\times r}$ para $0\le i\le k-1$, y el determinante de $L(z)$ no es idénticamente nulo. Un número complejo z_0 se dice que es un *valor propio* de $L(z)$ si $\det(L(z)) = 0$. Si z_0 es un valor propio de $L(z)$, entonces por el teorema S1.10 de [GOH1, pág. 331], $L(z)$ admite la siguiente representación:

$$L(z) = E_{z_0}(z) \begin{bmatrix} (z-z_0)^{k_1} & & 0 \\ & \ddots & \\ 0 & & (z-z_0)^{k_r} \end{bmatrix} F_{z_0}(z) \qquad (1.7)$$

donde $E_{z_0}(z)$ y $F_{z_0}(z)$ son polinomios matriciales inversibles en z_0, y $k_1 \leq k_2 \leq \ldots \leq k_r$ son enteros no negativos, los cuales coinciden con los grados de los divisores elementales de L(z) correspondientes a z_0. La representación (1.7) se llama *forma local de Smith* para L(z) en z_0. Los enteros k_j, para $1 \leq j \leq r$, están unívocamente determinados por L(z) y z_0, y se llaman *multiplicidades parciales* de L(z) en z_0. Notar que algunos de los números k_j pueden ser repetidos y que k_j aparece tantas veces como $(z-z_0)^{k_j}$ sea divisor elemental de L(z).

De acuerdo con el capítulo 7 de [GHO2], diremos que W(z) es una función racional matricial *rxr*, si cada entrada $w_{ij}(z)$ para $1 \leq i,j \leq r$ es una función racional escalar de la forma:

$$w_{ij}(z) = \frac{p_{ij}(z)}{q_{ij}(z)}$$

donde $p_{ij}(z)$ y $q_{ij}(z)$ son polinomios escalares y $q_{ij}(z)$ no es idénticamente nulo. Si W(z) es una *función racional matricial* de tamaño *rxr* con determinante no idénticamente cero, diremos que $z_0 \in \mathbb{C}$ es un *polo* de W(z), si z_0 es un polo de, al menos, una de las entradas de W(z). Asimismo, $z_0 \in \mathbb{C}$ es un polo de W(z) si, y sólo si, z_0 es un valor propio de la función matricial $\left(W(z)\right)^{-1}$.

El siguiente resultado es una versión matricial del lema 5.5 de [HENR, pág 242], donde se prueba para el caso escalar. La demostración puede ser obtenida trabajando componente a componente, usando resultados escalares y el teorema 7.2.3 de [GOH2, pág. 223].

TEOREMA 1.1. *Supongamos que el polinomio matricial* $L(z)$ *definido en* (1.6) *tiene todos sus valores propios en el disco* $|z|\leq 1$ *y que para aquellos valores propios* z_0 *de* $L(z)$ *tales que* $|z_0| = 1$, *se tiene que las multiplicidades parciales de* $L(z)$ *en* z_0 *son simples, es decir,* $k_j(z_0)=1$ *para* $1\leq j\leq r$. *Entonces:*

$$\left[I + z A_{k-1} + \ldots + z^k A_0 \right]^{-1} = \sum_{m\geq 0} \gamma_m z^m \quad ; \; |z|<1; \; \gamma_m \in \mathbb{C}^{r\times r} \qquad (1.8)$$

con:

$$\Gamma = \sup_{m\geq 0} || \gamma_m || < +\infty \qquad (1.9)$$

y:

$$\gamma_m + \gamma_{m-1} A_{k-1} + \ldots + \gamma_{m-k} A_0 = \begin{cases} I & \text{si } m=o \\ 0 & \text{si } m>0 \end{cases} \qquad (1.10)$$

donde hemos asumido que $\gamma_m = 0$ *para* $m < 0$.

NOTA 1.1. Por el teorema 5.1.1 de [GOH2, pág.145], el conjunto de los valores propios de L(z) coincide con el espectro de la matríz compañera:

$$C = \begin{bmatrix} 0 & I & 0 & \dots & 0 & 0 \\ 0 & 0 & I & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 0 & I \\ -A_0 & -A_1 & -A_2 & \dots & -A_{k-2} & -A_{k-1} \end{bmatrix} \tag{1.11}$$

y así, las hipótesis del *Teorema 1.1* sobre el polinomio matricial L(z), conducen a que la matríz C definida en (1.11), tenga radio espectral $\rho(C) = 1$ y a que C sea de clase M. Para el caso escalar, $r = 1$, las hipótesis del *Teorema 1.1* significan que el polinomio L(z) tiene todas sus raíces en el disco $|z| \le 1$, y que todas sus raíces z_0 tales que $|z_0| = 1$ son raíces simples.

I. B. (iii). CÁLCULO DIFERENCIAL MATRICIAL.

Por motivos de claridad en la presentación de los resultados que más adelante necesitaremos, comenzamos esta sección recordando los conceptos de *producto de Kronecker* de matrices y el *operador* *vector columna*. Sean, $A = [\ a_{ij}\ ] \in \mathbb{C}^{mxn}$ y $B = [\ b_{ij}\ ] \in \mathbb{C}^{rxs}$, entonces, el *producto de Kronecker* de A y B, representado por $A \otimes B$, es la matriz de los bloques definida como:

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix}$$

El operador vector columna, actuando sobre una matriz $A \in \mathbb{C}^{mxn}$ se define como:

$$\mathrm{vec}(A) = \begin{bmatrix} A_{.1} \\ \vdots \\ A_{.n} \end{bmatrix} ; \quad \text{siendo} \quad A_{.k} = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix}$$

Si $A \in \mathbb{C}^{mxn}$, $Y \in \mathbb{C}^{nxr}$, y $B \in \mathbb{C}^{rxs}$, entonces, por [GRAH, pág.25], tendremos:

$$\mathrm{vec}\ (\ A\ Y\ B\ ) = (\ B^T \otimes A\ )\ \mathrm{vec}Y$$

donde B^T es la traspuesta de la matríz B. Si $Y = [\ y_{ij}\ ] \in \mathbb{C}^{pxq}$

y $X = [\, x_{rs} \,] \in \mathbb{C}^{m\times n}$ entonces, por [GRAH, pág. 62 y 81] se verifica que:

$$\frac{\partial Y}{\partial x_{rs}} = \begin{bmatrix} \dfrac{\partial y_{11}}{\partial x_{rs}} & \cdots & \dfrac{\partial y_{1q}}{\partial x_{rs}} \\ \vdots & & \vdots \\ \dfrac{\partial y_{p1}}{\partial x_{rs}} & \cdots & \dfrac{\partial y_{pq}}{\partial x_{rs}} \end{bmatrix}$$

$$\frac{\partial Y}{\partial X} = \begin{bmatrix} \dfrac{\partial Y}{\partial x_{11}} & \cdots & \dfrac{\partial Y}{\partial x_{1n}} \\ \vdots & & \vdots \\ \dfrac{\partial Y}{\partial x_{m1}} & \cdots & \dfrac{\partial Y}{\partial x_{mn}} \end{bmatrix}$$

Recordemos la regla para la *derivada de un producto de matrices con respecto a una matríz*, cuya demostración puede encontrarse en [GRAH, pág.84]. Sean $X \in \mathbb{C}^{m\times n}$, $Y \in \mathbb{C}^{n\times v}$, $Z \in \mathbb{C}^{p\times q}$, entonces se tiene que:

$$\frac{\partial (XY)}{\partial Z} = \frac{\partial X}{\partial Z}\left(I_q \otimes Y \right) + \left(I_p \otimes X \right)\frac{\partial Y}{\partial Z}$$

Donde I_q e I_p son las matrices identidad para $\mathbb{C}^{q\times q}$ y $\mathbb{C}^{p\times p}$ respectivamente. Finalmente, recordemos la expresión de la *regla de la cadena para la derivada de una matríz con respecto a una matríz* [GRAH, pág. 88], [VETT]. Si $X \in \mathbb{C}^{m\times n}$, $Y \in \mathbb{C}^{u\times v}$, $Z \in \mathbb{C}^{p\times q}$, tenemos:

$$\frac{\partial Z}{\partial X} = \left[\frac{\partial [\text{vec}Y]^T}{\partial X} \otimes I_p \right] \left[I_n \otimes \frac{\partial Z}{\partial \text{vec}Y} \right]$$

I. C. Métodos Multipasos Matriciales.

I. C. (i). DEFINICIONES Y PROPIEDADES GENERALES.

En esta sección completaremos algunos resultados sobre métodos multipasos matriciales lineales, dados recientemente en [JOD4], donde fue introducido el concepto. Para una mayor claridad en la exposición, recordemos las siguientes definiciones:

DEFINICION 1.1. ([JOD4]) *Consideremos el problema de valores iniciales* (1.1) *bajo la hipótesis* (1.2). *Sean* $k\geq 1$, $t_n=nh$, $h>0$, *y* $f_n = f(t_n,Y_n) \in \mathbb{C}^{r\times q}$ *; entonces , un método k-pasos matricial es una relación de la forma:*

$$Y_{k+n}+A_{k-1}\,Y_{n+k-1}+\ \ldots\ +A_0\,Y_n = h\left\{B_k f_{n+k}+\ \ldots\ +B_0 f_n\right\} \tag{1.12}$$

donde $A_j \in \mathbb{C}^{r\times r}$ *y* $B_i \in \mathbb{C}^{r\times r}$ *para* $0\leq i\leq k$; $0\leq j\leq k-1$.

El concepto de *consistencia* lo introducimos ahora del modo siguiente:

DEFINICION 1.2. ([JOD4]) *Se dice que el método k-pasos matricial lineal* (1.12) *es consistente si las matrices* A_j *y* B_i *satisfacen* *las ecuaciones:*

$$A_0 + A_1 + \dots + A_{k-1} + I = 0$$

y:

$$B_0 + \dots + B_k = kI + (k-1)\,A_{k-1} + \dots + 2A_2 + A_1$$

donde asumimos que $A_j = 0$ *si* $j \geq k$.

En [JOD4] se define la *cero-estabilidad* para métodos k-pasos matriciales con $k>1$. Introduzcamos otra definición que coincida con la dada en [JOD4] para $k\geq 2$, pero que sea apropiada también en el caso $k=1$.

DEFINICION .1.3. *Diremos que el método k-pasos matricial lineal* (1.12) *es cero-estable si el polinomio matricial* $L(z)$ *dado en* (1.6) *tiene todos los valores propios en el disco* $|z| \leq 1$ *y, para todo valor propio* z_0 *tal que* $|z_0| = 1$, *todas las multiplicidades parciales de* $L(z)$ *en* z_0 *son simples, es decir:*

$$k_j(z_0) = 1 \quad \textit{para} \quad 1 \leq j \leq k$$

EJEMPLO 1.1. Consideremos el método unipaso matricial:

$$Y_{n+1} + A_0 Y_n = h \left\{ B_1 f_{n+1} + B_0 f_n \right\} ; \quad n \geq 0 \qquad (1.13)$$

En este caso, $L(z) = z\, I - A_0$. Los valores propios de $L(z)$ son los de A_0, y por la *Definición 1.3* y el teorema 5.1.1 de [GOH2], el método (1.13) es cero estable si $\rho(A) \leq 1$ y cuando $\rho(A) = 1$, entonces A_0 es una matríz de clase M. En particular, tomando $A_0 = -I$, este método es cero-estable y consistente. Si tomamos $A_0 = -I$ y $B_0 = B_1 = I/2$, se obtiene el análogo del *método trapezoidal escalar*.

Por la *Nota 1.1*, es fácil ver que la *Definición 1.3* es equivalente a la definición de cero-estabilidad dada en [JOD4] para el caso k>1. Consideremos ahora el método k-paso matricial (1.12) y su operador lineal en diferencias asociado Φ definido por:

$$\Phi(Y(t);h) = \sum_{j=0}^{k} \left[A_j Y(t+jh) - h\, B_j Y'(t+jh) \right] ; \quad A_k = I \qquad (1.14)$$

donde $Y(t)$ es una matríz arbitraria de $\mathbb{C}^{r\times q}$ cuyas entradas son funciones contínuamente diferenciables en $[a,b]$, e I representa la matríz identidad de $\mathbb{C}^{rxr}$. Desarrollando la función $Y(t+jh)$ y su derivada $Y'(t+jh)$ en *serie de Taylor* alrededor de t, y reagrupando términos en (1.14), se obtiene:

$$\Phi(Y(t);h) = C_0Y(t) + C_1hY'(t) + \ldots + C_sh^sY^{(s)}(t)+\ldots \qquad (1.15)$$

donde C_s son matrices de $\mathbb{C}^{rxr}$.

DEFINICION 1.4. *El operador en diferencias* (1.14) *y el método k-pasos matricial asociado* (1.12), *se dice que son de órden* p *si, en* (1.15), $C_0= C_1= \ldots = C_p= 0$ con $C_{p+1} \neq 0$.

Cálculos sencillos ofrecen la siguiente expresión para las matrices C_s en términos de las matrices coeficientes A_j, B_j:

$$C_0 = A_0 + A_1 + \ldots + A_{k-1} + I$$

$$C_1 = A_1 + 2A_2 + \ldots + kI - (B_0 + B_1 + \ldots + B_k)$$

$$\vdots$$

$$C_s = \frac{1}{s!} (A_1+2^sA_2+\ldots+k^sI) - \frac{1}{(s-1)!} (B_1 + 2^{s-1} B_2 + \ldots + k^{s-1} B)$$

$$s = 2,3, \ldots \qquad (1.16)$$

Así, por ejemplo, el método unipaso matricial de la forma (1.13) con $A_0 = -I$ y $B_0 = B_1 = I/2$ es de órden $p=2$ porque se tiene que $C_0 = C_1 = C_2 = 0$ y $C_3 = -I/12$.

De manera análoga al caso escalar, es fácil demostrar, ver [LAMB, pág. 49-52] y [HENR, pág.285], que si el método (1.12) es de órden p, entonces:

$$\| \Phi(Y(t);h) \| \le h^{p+1} G D \tag{1.17}$$

donde Y(t) es la solución teórica del problema (1.1) y:

$$G = \| C_{p+1} \| ; \qquad D \ge \max_{0\le t\le b} \| Y^{(p+1)}(t) \| \tag{1.18}$$

Uno de nuestros principales objetivos es encontrar cotas para el error global de discretización del método (1.12) en el punto $t_n = nh$, que denotaremos por e_n y que viene definido por:

$$e_n = Y(t_n) - Y_n \tag{1.19}$$

donde $Y(t_n)$ es el valor de la solución teórica del problema (1.1) en t_n, e Y_n es el valor aproximado que ofrece el método (1.12).

De modo similar al caso escalar, para encontrar *cotas* *superiores para el error global de discretización*, es necesario estudiar el desarrollo de las soluciones de ciertas ecuaciones en diferencias lineales no homogéneas. El siguiente resultado, cuya demostración puede encontrarse en [JOD5], trata de estos aspectos:

TEOREMA 1.2. *Consideremos la ecuación en diferencias matricial no homogénea:*

$$Z_{m+k} + A_{k-1} Z_{m+k-1} + \dots + A_0 Z_m =$$
$$= h \left\{ B_{k,m} \| Z_{m+k} \| + B_{k-1,m} \| Z_{m+k-1} \| + \dots + B_{0,m} \| Z_m \| \right\} + \Lambda_m \tag{1.20}$$

donde $A_j \in \mathbb{C}^{r\times r}$ *para* $0 \le j \le k-1$, $B_{s,m} \in \mathbb{C}^{r\times q}$ *para* $0 \le s \le k$; $h > 0$; m *es un entero no negativo;* $\Lambda_m \in \mathbb{C}^{r\times q}$; *y supongamos que el método* (1.12) *es cero-estable. Sean* B_* , B *y* Λ *constantes positivas tales que:*

$$\| B_{k,p} \| + \| B_{k-1,p} \| + \dots + \| B_{0,p} \| \le B_* \quad ; \quad \| \Lambda_p \| \le \Lambda \quad ;$$

$$\| B_{k,p} \| \le B \quad ; \quad 0 < h < B^{-1} \quad ; \quad 0 \le p \le N \tag{1.21}$$

Entonces, cada solución de (1.20) *para la cual:*

$$\| Z_p \| \le Z \quad ; \qquad 0 \le p \le k-1 \tag{1.22}$$

satisface:

$$\| Z_n \| \quad \le \quad K_* \ \exp \left(nhL_* \right) \qquad ; \quad 0 \le n \le N \tag{1.23}$$

donde:

$$L_* = \Gamma_* \, B_* \quad ; \qquad K_* = \Gamma_* \, (\, N\Lambda + kAZ \,) \tag{1.24}$$

$$A = \| A_0 \| + \ldots + \| A_{k-1} \| + 1 \quad ; \quad \Gamma = (\, 1 - hB \,)^{-1} \tag{1.25}$$

y Γ *está definido en el Teorema 1.1.*

Pensando en las posteriores aplicaciones , supongamos que la sucesión $\left\{ Y_n \right\}$ es una solución exacta de la ecuación (3.1). Sin embargo, supondremos que $\left\{ Y_n \right\}$ satisface tan sólo que:

$$Y_{m+k} + A_{k-1} Y_{m+k-1} + \ldots + A_0 Y_m - h \left\{ B_k f_{m+k} + \ldots + B_0 f_m \right\} = \Theta_m K_1 h^{q+1}$$

$$m \geq 0 \qquad (1.26)$$

donde K_1 y q son constantes no negativas, y Θ_m son matrices de $\mathbb{C}^{r\times q}$ cumpliendo que $\| \Theta_m \| \leq 1$. Supondremos también que los valores de partida $Y_j = \Omega_j(h)$; $0 \leq j \leq k-1$, son matrices en $\mathbb{C}^{r\times q}$ tales que están sujetas a un cierto error y sea $\xi = \xi(h)$ definido por:

$$\xi = \xi(h) = \max_{0 \leq j \leq k-1} \| \Omega_j(h) - Y(jh) \| \qquad (1.27)$$

Por el *Teorema 1.2* y los anteriores comentarios, se establece el siguiente resultado:

TEOREMA 1.3. *Consideremos un método k-pasos matricial cero-estable del tipo* (1.12) *de órden* $p \geq 1$. *Sea* Γ *la constante definida en Teorema 1.1 y sean* h, A, Γ^*, B^*, G *y* D *definidas anteriormente en* (1.18), (1.24), (1.25), *y:*

$$\Gamma^* = \Gamma \left(1 - h\, L\, \| B_k \| \right)^{-1} \quad ; \quad A = 1 + \| A_{k-1} \| + \ldots + \| A_0 \| \; ;$$

$$h < \left(\| B_k \| \; L \right)^{-1} \tag{1.28}$$

entonces, el error global de discretizacion e_n *está acotado superiormente mediante la desigualdad:*

$$\| e_n \| \le \Gamma^* \left[Ak\xi + t_n \left(h^p\, G\, D\, t_n + K_1 h^q \right) \right] \exp \left(L\, \Gamma^*\, B^*\, t_n \right) \tag{1.29}$$

donde L *es la constante de Lipschitz definida en* (1.2), *y* ξ *viene dado por* (1.27).

Nótese que *desde el punto de vista práctico, es necesario encontrar una constante adecuada D que satisfaga* (1.18), *cuestión que abordaremos para cada problema concreto en su sección correspondiente*. Por otro lado, en el siguiente apartado dentro de esta misma sección, probaremos este teorema en el caso de métodos unipaso matriciales, que son lo que aplicaremos en la resolución de los problemas que nos ocupan.

I. C. (ii). MÉTODOS LINEALES UNIPASO.

En este apartado, demostraremos algunos de los resultados presentados anteriormente en el caso general, para un tipo concreto de métodos unipaso matriciales, expuestos en *Ejemplo 1.1* , que englobann la generalización al caso matricial de la *regla de los trapecios* y del *método de Euler* , que serán los que vamos a utilizar en lo que sigue. Obtendremos así las expresiones concretas que necesitaremos para evaluar el error de discretización. Supondremos $\xi = 0$ pues en la ecuación (1.1) asumimos que el valor inicial Ω es conocido exactamente al tratarse de un dato del problema.

Así, para la construcción de soluciones numéricas discretas para el problema de valor inicial (1.1) bajo la hipótesis (1.2), utilizaremos métodos unipaso matriciales de la forma:

$$Y_{n+1} + A_0 Y_n = h \left\{ B_1 f_{n+1} + B_0 f_n \right\} ; \qquad n \geq 0 \qquad (1.13)$$

donde B_0 , B_1 son matrices de $\mathbb{C}^{r\times r}$ e Y_n , $f_n = f(t_n, Y_n)$ son matrices de $\mathbb{C}^{r\times q}$; $t_n = nh \in [0,b]$; $h>0$; y:

$$B_0 + B_1 = I \tag{1.30}$$

Consideremos la ecuación matricial en diferencias no homogénea, correspondiente de la expresión (1.20) para el caso que estamos considerando:

$$Z_{m+1} - Z_m = h\left\{ B_{1,m} \,\|Z_{m+1}\| + B_{0,m} \,\|Z_m\| \right\} + \Lambda_m \tag{1.31}$$

donde Λ_m, $B_{1,m}$, $B_{0,m}$ son matrices de $\mathbb{C}^{rxq}$; h>0; m es un entero no negativo; y sea $\left\{ Z_m \right\}$ una sucesión de matrices, solución de (1.31). Si escribimos esta ecuación para $m = n-p-1$, tendremos:

$$Z_{n-p}-Z_{n-p-1} = h\left\{B_{1,n-p-1} \,\|Z_{n-p}\| + B_{0,n-p-1} \,\|Z_{n-p-1}\|\right\} + \Lambda_{n-p-1}$$

Considerando la última ecuación para $p = 0,1,2,\ldots n-1$, y sumando las ecuaciones resultantes, se sigue que la suma del primer miembro de la igualdad es $S_n = Z_n - Z_0$ y el segundo miembro de la igualdad toma el valor:

$$h \left\{ B_{1,n-1} \| Z_n \| + \left(B_{0,n-1}+B_{1,n-2}\right) \| Z_{n-1} \| +\ldots+ B_{0,0} \| Z_0 \| \right\} +$$

$$+ \Lambda_{n-1} + \Lambda_{n-2} + \ldots + \Lambda_0 \qquad (1.32)$$

Consideremos las constantes positivas B, B_*, Λ, Γ_* y K_* definidas por:

$$\| B_{1,p} \| \leq B;\ \| B_{1,p} \| + \| B_{0,p} \| \leq B_* \ ;\ \| \Lambda_p \| \leq \Lambda;\ 0 \leq p \leq N \qquad (1.33)$$

$$0<h<B^{-1}\ ;\ \Gamma_* = (1-hB)^{-1}\ ;\ L_* = \Gamma_* B_*\ ;\ K_* = \Gamma_* \left(N\Lambda + \| Z_0 \|\right) \qquad (1.34)$$

Igualando la expresión (1.32) a $Z_n - Z_0$, tomando normas y aplicando la desigualdad $\| B_{0,0} \| + \| B_{1,0} \| \leq B_*$ se sigue que:

$$\| Z_n \| \leq h B \| Z_n \| + h B_* \sum_{m=0}^{n-1} \| Z_m \| + N\Lambda + \| Z_0 \| \qquad (1.35)$$

Por tanto:

$$(1-hB)\, \| Z_n \| \;\le\; h\, B_* \sum_{m=0}^{n-1} \| Z_m \| \;+\; \Big(N\Lambda + \| Z_0 \| \Big)$$

y por (1.33),(1.34), se sigue que:

$$\| Z_n \| \;\le\; h\, L_* \sum_{m=0}^{n-1} \| Z_m \| \;+\; K_* \tag{1.36}$$

Por (1.36) y [HENR, pág. 246] tenemos:

$$\| Z_n \| \;\le\; K_* \,(1+hL_*)^n \quad ; \quad 0 \le n \le N \tag{1.37}$$

Teniendo en cuenta que $(1+hL_*)^n \le \exp(nhL_*)$, queda probado el resultado siguiente:

TEOREMA 1.4 *Consideremos el método unipaso matricial definido por* (1.13) *y* (1.30), *y sean las constantes* Λ, B, B_* , K_* *y* L_* *definidas en* (1.33)-(1.34). *Entonces, para cualquier sucesión de matrices* $\{ Z_n \}$ *que sean solución de* (2.1), *se tiene que:*

$$\| Z_n \| \;\le\; K_* \exp\Big(n\, h\, L_* \Big) \quad ; \quad 0 \le n \le N \tag{1.38}$$

Introduzcamos el operador en diferencias Φ asociado al método (1.13) y definido por:

$$\Phi(Y(t);h) = Y(t+h) - Y(t) - h\left(B_1 Y'(t+h) + B_0 Y'(t)\right) \qquad (1.39)$$

donde $Y(t)$ es una matriz arbitraria de tamaño *rxq*, siendo sus entradas funciones contínuamente diferenciables en $[0,b]$ que toman valores en $\mathbb{C}$. Desarrollando la función $Y(t+jh)$ y su derivada $Y'(t+jh)$ en *serie de Taylor* alrededor de t, y reagrupando términos en (1.39), tenemos:

$$\Phi(Y(t);h) = C_0 Y(t) + C_1 hY'(t) + \ldots + C_s h^s Y^{(s)}(t) + \ldots \qquad (1.40)$$

donde C_s es una matriz de $\mathbb{C}^{rxr}$ que puede ser escrita en términos de los coeficientes matriciales:

$$C_0 = 0 \ ; \quad C_1 = I - \left(B_0 + B_1\right) = 0 \ ; \ \ldots \ ; \ C_s = \frac{1}{s!} I - \frac{1}{(s-1)!} B_1$$

$$s = 2, 3, \ldots \qquad (1.41)$$

Si escribimos (1.13) en la forma:

$$Y_{n+1} - Y_n - h\left\{ B_1 f_{n+1} + B_0 f_n \right\} = 0$$

y restando de esta ecuación $\Phi(Y(t_n);h)$ se sigue que:

$$e_{n+1} - e_n - h\left\{ B_1 \left(Y'(t_n+h) - f_{n+1} \right) + B_0 \left(Y'(t_n) - f_n \right)\right\} =$$

$$e_{n+1} - e_n - h\left\{ B_1\left(f(t_{n+1},Y(t_{n+1})) - f_{n+1}\right) + B_0\left(f(t_n,Y(t_n)) - f_n\right)\right\} =$$

$$=\Phi(Y(t_n);h) \qquad (1.42)$$

donde se ha aplicado la definición de error de discretización, dada en (1.19), y la expresión (1.1). Consideremos ahora la siguiente sucesión de matrices de $\mathbb{C}^{rxq}$ definidas por:

$$P_n = \begin{cases} \left(f(t_n,Y(t_n)) - f(t_n,Y_n)\right) \| e_n \|^{-1} & ; \quad \text{si} \quad e_n \neq 0 \\ 0 & ; \quad \text{si} \quad e_n = 0 \end{cases} \qquad (1.43)$$

De (1.42) y (1.43) tenemos:

$$e_{n+1}-e_n = h\left\{ B_1\, P_{n+1}\, ||\, e_{n+1}|| + B_0\, P_n\, ||\, e_n\, || \right\} + \Phi(Y(t_n);h) \quad (1.44)$$

Supongamos que el método (1.13) es de órden $p\geq1$, de acuerdo con la *definición 1.4*, y sean G Y D constantes que satisfacen (1.18). Por (1.17) tendremos que:

$$||\, \Phi(Y(t_n);h)\, || \;\leq\; h^{p+1}\, G\, D$$

Por otro lado, por (1.2) y (1.43) se sigue que $||\, P_n\, || \leq L$; $||\, B_1\, P_{n+1}\, || \;\leq\; L\, ||\, B_1\, ||$. Si llamamos B_* y N a los números positivos:

$$B_* = ||\, B_0\, || + ||\, B_1\, || \; ; \qquad N = [b/h] \qquad (1.45)$$

Aplicando el *teorema 1.5* a la acuación en diferencias (1.44), el siguiente resultado queda establecido:

TEOREMA 1.6. *Consideremos un método unipaso matricial del tipo (1.13) y (1.30), de órden* $p\geq1$ *, y sean* h *,* Γ^* *constantes positivas definidos como:*

$$h < \left(L \, \|B_1\| \right)^{-1} \quad ; \quad \Gamma^* = \left(1 - h \, L \, \| B_1 \| \right)^{-1} \tag{1.46}$$

donde L *es la constante de Lipschitz dada en* (1.2). *Si* G *y* D *son las constantes definidas mediante* (1.18), *entonces, el error global de dicretización* $e_n = Y(t_n) - Y_n$ *, está superiormente acotado por la desigualdad:*

$$\| e_n \| \le \Gamma^* \, h^p \, G \, D \, t_n \, \exp\left(L \, \Gamma^* \, B^* \, t_n \right) , \quad n \ge 0 \tag{1.47}$$

EJEMPLO 1.2. Consideremos el método unipaso matricial definido por (1.13) y (1.30), donde $B_0 = B_1 = I/2$:

$$Y_{n+1} - Y_n = \frac{h}{2} \left\{ f_{n+1} + f_n \right\} \tag{1.48}$$

Por (1.41) se tiene que $C_0 = C_1 = C_2 = 0$, y $C_3 = -I/12$. Así, (1.48) define un método unipaso de órden p=2. Las constantes que aparecen en el *Teorema 1.6* toman los siguientes valores:

$$G = \| C_3 \| = 1/12 \ ; \ B = \| B_0 \| + \| B_1 \| = 2 \ ; \ \Gamma^* = (1 - h\, L/2)^{-1}$$

$$D \geq \max \left\{ \| Y^{(3)}(t) \| \ ; \quad 0 \leq t \leq b \right\}$$

La desigualdad (1.47) toma la forma:

$$\| e_n \| \leq \frac{h^2}{12} D\, t_n\, (1-hL/2)^{-1} \exp\left(t_n\, L\, (1 - hL/2)^{-1} \right) \tag{1.49}$$

Nótese que este método es la generalización al caso matricial de la *regla de los trapecios*.

EJEMPLO 1.3. Consideremos el método unipaso matricial, donde $A_0 = I$ $B_0 = I$ y $B_1 = 0$:

$$Y_{n+1} - Y_n = h\, f_n \tag{1.50}$$

Por (1.41) se tiene que $C_0 = C_1 = 0$, y $C_2 = I/2$. Así, (1.50) define un método unipaso de órden p=1. Las constantes del *Teorema*

1.6 toman los valores:

$$G = \| C_2 \| = 1/2 \ ; \qquad \Gamma^* = 1.$$

$$D \geq \max \left\{ \| Y^{(2)}(t) \| \ ; \quad 0 \leq t \leq b \right\} \tag{1.51}$$

y el error de discretización e_n verifica:

$$\| e_n \| \leq \frac{h}{2} \ D \ t_n \ \exp \left(L \ t_n \right) \tag{1.52}$$

Nótese que este método corresponde a la generalización al caso matricial del *método de Euler*. Obsérvese que si bien este método puede resultar más sencillo al cálculo que el del *Ejemplo 1.2*, ofrece en cambio una cota de error menos precisa como puede observarse según la dependencia en h.

CAPITULO II

PROBLEMA LINEAL BILATERAL

CON COEFICIENTES VARIABLES

II. A. Caso Homogéneo.

II. A. (i). PRESENTACIÓN DEL PROBLEMA.

En esta sección, construiremos soluciones analíticas aproximadas para problemas de valor inicial referidos a ecuaciones diferenciales matriciales del tipo $X'(t) = A(t)X(t) + X(t)B(t)$, donde los coeficientes $A(t)$ y $B(t)$ son *funciones matriciales al menos dos veces contínuamente diferenciables*. Mediante el uso de métodos unipaso matriciales y funciones B-spline matriciales lineales que interpolen la solución numérica en un conjunto discreto de puntos, construiremos una solución aproximada cuyo error sea menor que un valor cualquiera $\varepsilon>0$ prefijado. Los resultados que presentamos a continuación, han sido aceptados para su publicación en la referencia [JOD7].

Consideremos pues, ecuaciones diferenciales matriciales del tipo:

$$X'(t) = A(t)X(t) + X(t)B(t); \quad X(0) = C; \quad 0\le t\le b \tag{2.1}$$

donde la incógnita $X(t)$ y los coeficientes $A(t)$, $B(t)$ son matrices complejas de tamaño rxr, es decir, elementos de $\mathbb{C}^{rxr}$. Estas ecuaciones aparecen en muchos campos de la ciencia y la ingeniería, principalmente en problemas de optimización en teoría de control lineal [BARN], [REI1], [THOM]. La ecuación (2.1) ha sido estudiada por muchos autores en el caso de coeficientes constantes, ver [BARR], [DAVI], [SERB], sin embargo, para el caso de coeficientes variables, esta ecuación ha recibido poco tratamiento numérico.

Por [REI1], [BARN, pág. 109], la solución teórica de la ecuación (2.1) viene dada por:

$$X(t) = Y(t)\, C\, Z(t)$$

donde $Y(t)$ es solución de la ecuación matricial:

$$Y'(t) = A(t)\, Y(t) \; ; \quad Y(0) = I \tag{2.2}$$

y $Z(t)$ es la solución de:

$$Z'(t) = Z(t)\, B(t) \; ; \quad Z(0) = I \tag{2.3}$$

Desgraciadamente, la solución exacta de los problemas (2.2) y (2.3) no son calculables analíticamente, lo cual motiva la búsqueda de alternativas que ofrezcan soluciones analíticas aproximadas y cotas de error en términos de los datos.

Si aplicamos el método unipaso matricial (1.48) para obtener las soluciones numéricas del problema (2.2) en un conjunto discreto de puntos del intervalo $[0,b]$, necesitamos determinar las constantes que aparecen en la expresión (1.49), en términos de los datos del problema, para evaluar el error de discretización cometido.

Consideremos la ecuación diferencial matricial (2.2), donde

A(t) es una matriz cuyas entradas son funciones dos veces contínuamente diferenciables. Tomando derivadas para la solución Y(t) de (2.2), se tiene que:

$$Y^{(2)}(t) = A'(t)\,Y(t) + (A(t))^2\,Y(t)$$

$$Y^{(3)}(t) = A^{(2)}(t)\,Y(t) + A'(t)\,A(t)\,Y(t) + (A(t))^3\,Y(t) +$$

$$+ \left[A'(t)\,A(t) + A(t)\,A'(t)\right] Y(t) =$$

$$= A^{(2)}(t)Y(t) + 2A'(t)A(t)Y(t) + A(t)A'(t)Y(t) + (A(t))^3Y(t)$$

Por [FLET, pág. 14], la solución teórica Y(t) de (2.2) satisface:

$$\| Y(t) \| \le \exp(tk_0); \quad 0\le t\le b \tag{2.4}$$

y si llamamos k_i, i = 0, 1, 2 a unas constantes positivas que satisfacen:

$$k_i \ge \max_{0\le t\le b} \| A^{(i)}(t) \| ; \quad i= 0,\ 1,\ 2. \tag{2.5}$$

tenemos:

$$\max_{0\le t\le b} \| Y^{(3)}(t) \| \le \exp(bk_0) \left\{k_0^3+3k_1k_0+k_2\right\} = D \tag{2.6}$$

Si $h < 2/k_0$, entonces, por los anteriores comentarios y lo expuesto en la parte dedicada a métodos multipasos matriciales del *capítulo I*, se sigue que el error global de discretización e_n en $t_n = nh$ definido en (1.19), donde una aproximación al valor exacto de la solución de (2.2) es el valor de Y_n obtenido mediante el método unipaso matricial (1.48), satisface:

$$\| e_n \| \le h^2 \, t_n \exp(bk_0) \left(1 - \frac{k_0 h}{2}\right)^{-1} \left(k_0^3 + 3k_1 k_0 + k_2\right) \cdot$$
$$\cdot \exp\left(\frac{t_n \left(1 - \frac{k_0 h}{2}\right)^{-1} k_0}{12}\right) \tag{2.7}$$

Como $h < 1/k_0$, resulta que:

$$1 - \frac{k_0 h}{2} > 1/2 \tag{2.8}$$

entonces, para $h < 1/k_0$, (2.7) toma la forma:

$$\| e_n \| \le \frac{h^2 t_n}{6} \exp(bk_0) \left\{k_0^3 + 3k_1 k_0 + k_2\right\} \exp\left(2k_0 t_n\right) \tag{2.9}$$

Si consideramos ahora el problema (2.3) y definimos las constantes q_i tales que:

$$q_i \geq \max_{0 \leq t \leq b} \| B^{(i)}(t) \| \quad ; \qquad i = 0,\ 1,\ 2, \tag{2.10}$$

y llamamos:

$$v_n = Z(t_n) - Z_n$$

al error global de discretización cuando aproximamos el valor exacto $Z(t_n)$ de la solución de (2.3) mediante la solución numérica Z_n calculado mediante el método unipaso:

$$Z_{n+1} - Z_n = \frac{h}{2} \left\{ g_{n+1} + g_n \right\} \tag{2.11}$$

donde:

$$g_n = G(t_n, Z_n)$$

y:

$$G(t,Z) = Z\ B(t)$$

entonces, para valores de h tales que:

$$h < \frac{1}{q_0} \tag{2.12}$$

se sigue que:

$$\| v_n \| \le \frac{h^2 t_n}{6} \exp(bq_0) \left\{q_0^3+3q_1q_0+q_2\right\} \exp\left(2q_0 t_n\right) \qquad (2.13)$$

II. A. (ii). Solución Aproximada Analítica Y Cota De Error.

Empecemos este apartado recordando algunos resultados sobre la interpolación con B-splines. Si estamos interesados en la construcción de una función aproximada que interpole los valores exactos $y_0, y_1, \ldots, y_N$ en los puntos $t_0, t_1, \ldots, t_N$ de una partición del intervalo $[0,b]$, calcularemos entonces el B-spline lineal definido por:

$$s(t) = \sum_{n=-1}^{N-1} b_{1n}(t) \; y_{n+1} \qquad (2.14)$$

donde $t \in [0,b]$; $t_{n+1} - t_n = h$; $h = [b/N]$; y:

$$b_{1n}(t) = h^{-1} \begin{cases} (t-t_n) & \text{para} \quad t_n \le t < t_{n+1} \\ (t_{n+2}-t) & \text{para} \quad t_{n+1} \le t < t_{n+2} \end{cases} \qquad (2.15)$$

con $b_{1n}(t) = 0$ para $t < t_n$ y para $t_{n+2} \le t$. Además,

b_{1n} es no negativo, satisfaciendo $b_{1n}(t) + b_{1,n-1}(t) = 1$ para todo $t \in [t_n, t_{n+2}]$; ver [HAMM, pág. 247-248]. Si consideramos el B-spline lineal construido en términos de los valores aproximados $\hat{y}_0$, $\hat{y}_1$, ... , $\hat{y}_N$; entonces tendríamos la nueva función aproximadora:

$$T(t) = \sum_{n=-1}^{N-1} b_{1n}(t)\ \hat{y}_{n+1} \tag{2.16}$$

de modo que:

$$| s(t) - T(t) | \le \max_{-1 \le n \le N-1} | y_{n+1} - \hat{y}_{n+1} | \sum_{n=-1}^{N-1} |b_{1n}(t)| =$$

$$= \max_{-1 \le n \le N-1} | y_{n+1} - \hat{y}_{n+1} | \ ; \quad \text{para} \quad 0=t_0 \le t \le t_N=b \tag{2.17}$$

Consideremos ahora el caso matricial. Suponiendo conocidos los valores exactos $Y(t_0)$, $Y(t_1)$, ... , $Y(t_N)$ de la función matricial $Y(t) \in \mathbb{C}^{r\times r}$, en los puntos $t_0 = 0$, t_1 , ... , $t_N = b$, pero desconociendo los valores de dicha matríz en el resto del intervalo, estamos interesados en construir una interpolación lineal con funciones B-spline matriciales, que denotaremos por $W(t)$, pero sustituyendo los valores exactos $Y(t_n)$ por sus correspondientes valores aproximados Y_n, para $n=0,1,2,\ldots,N$. Consideremos las funciones B-spline matriciales lineales definidas como:

$$V(t) = \sum_{n=-1}^{N-1} Y(t_{n+1})\, b_{1n}(t) \; ; \quad W(t) = \sum_{n=-1}^{N-1} Y_{n+1}\, b_{1n}(t) \; ;$$

$$0 = t_0 \le t \le t_N = b \tag{2.18}$$

entonces, tomando normas, tendremos:

$$\| V(t) - W(t) \| \le \max_{0 \le n \le N} \| Y(t_n) - Y_n \| \tag{2.19}$$

Sea $f(t)$ una función escalar dos veces contínuamente diferenciable en el intervalo $[0,b]$. Sea $s(t)$ el B-spline definido en (2.14) con $y_n = f(t_n)$ para $0 \le n \le N$, entonces por [HAMM, pág.257] se sigue que:

$$\max_{0 \le t \le b} |f(t)-s(t)| \le \frac{h^2}{8} \max_{0 \le t \le b} |f^{(2)}(t)| \tag{2.20}$$

Si $Y(t)$ es una matriz de $\mathbb{C}^{r\times r}$ cuyas entradas son funciones, y aplicamos (2.20) y (1.3), asumiendo que $Y(t)$ es dos veces contínuamente diferenciable, la función B-spline matricial lineal

$V(t)$ definida en (2.18), satisface:

$$\max_{0\le t\le b} \| Y(t) - V(t) \| \le r \frac{h^2}{8} \max_{0\le t\le b} \| Y^{(2)}(t) \| \qquad (2.21)$$

Nos interesamos ahora en la construcción de una aproximación analítica de la ecuación (2.1). Antes de nada, nótese que la solución numérica Y_n de (1.48), correspondiente al problema (2.2), procede de la relación:

$$Y_{n+1} - Y_n = \frac{h}{2} \left(A(t_{n+1})\, Y_{n+1} + A(t_n)\, Y_n \right) \; ; \qquad Y_0 = I$$

$$\left(I - \frac{h}{2} A(t_{n+1})\right) Y_{n+1} = \left(I + \frac{h}{2} A(t_n)\right) Y_n \; ; \; Y_0 = I \; ; \; n\ge 1 \qquad (2.22)$$

Si $h< 1/k_0$, estando k_0 definido en (2.5), entonces, por el *lema de perturbación*, [ORT2, pág.45], las matrices de los coeficientes de Y_{n+1} e Y_n en (2.22) son inversibles y se tiene que:

$$Y_n = \prod_{j=0}^{n-1} \left\{ \left(I - \frac{h}{2} A(t_{n-j}) \right)^{-1} \left(I + \frac{h}{2} A(t_{n-j-1}) \right) \right\}$$

$$Y_0 = I \ ; \quad n \geq 1 \tag{2.23}$$

De igual manera, si $h < 1/q_0$, donde q_0 viene definido en (2.10), la solución numérica Z_n de la ecuación (2.11) correspondiente al problema (2.3), procede de la relación:

$$Z_{n+1} - Z_n = \frac{h}{2} \left(Z_{n+1} B(t_{n+1}) + Z_n B(t_n) \right) \ ; \quad Z_0 = I$$

$$Z_{n+1} \left(I - \frac{h}{2} B(t_{n+1}) \right) = Z_n \left(I + \frac{h}{2} B(t_n) \right) \ ; \ Z_0 = I \ ; \quad n \geq 1$$

cuya solución es de la forma:

$$Z_n = \prod_{j=0}^{n-1} \left\{ \left(I + \frac{h}{2} B(t_j) \right) \left(I - \frac{h}{2} B(t_{j+1}) \right)^{-1} \right\}$$

$$Z_0 = I \ ; \quad n \geq 1 \tag{2.24}$$

Teniendo en cuenta que la solución exacta $X(t)$ del problema (2.1) viene dada por $X(t) = Y(t)\,C\,Z(t)$, de (2.23) y (2.24), proponemos la aproximación numérica de $X(t)$ en $t_n = nh$, definida por:

$$X_n = Y_n\,C\,Z_n \; ; \qquad X_0 = C \; ; \qquad 1 \le n \le N \tag{2.25}$$

donde Y_n viene dado por (2.23) y Z_n por (2.24), para $1 \le n \le N$ con $Nh = b$.

A partir de los valores aproximados de $X(t)$ en $t_n = nh$, dados por X_n definido según (2.25), construimos la función B-spline matricial lineal:

$$W(t) = \sum_{n=-1}^{N-1} X_{n+1}\, b_{1n}(t) \; ; \qquad 0 \le t \le b \tag{2.26}$$

donde b_{1n} está definido en (2.15). Si denotamos por $V(t)$ la función B-spline matricial lineal que interpola los valores teóricos exactos de la solución $X(t_n)$ en t_n para el problema (2.1):

$$V(t) = \sum_{n=-1}^{N-1} X(t_{n+1})\, b_{1n}(t) \;; \quad 0 \le t \le b, \tag{2.27}$$

entonces, de (2.19), resulta:

$$\| V(t) - W(t) \| \le \max_{0 \le n \le N} \| X(t_n) - X_n \| \tag{2.28}$$

Para poder obtener una cota superior de la expresión que aparece en el segundo miembro de la desigualdad (2.28), obsérvese que:

$$X(t_n) - X_n = Y(t_n)\, C\, Z(t_n) - Y_n\, C\, Z_n =$$

$$= \left(Y(t_n) - Y_n\right) C Z(t_n) + Y(t_n) C \left(Z(t_n) - Z_n\right) - \left(Y(t_n) - Y_n\right) C \left(Z(t_n) - Z_n\right) \tag{2.29}$$

Por [FLET, pág. 114] sabemos que:

$$\| Y(t) \| \le \exp(tk_0) \quad \text{y} \quad \| Z(t) \| \le \exp(tq_0) \quad \text{para } 0 \le t \le b, \tag{2.30}$$

por lo que tomando normas en (2.29), y utilizando (2.30), obtenemos:

$$\| X(t_n) - X_n \| \le \exp(t_n q_0) \| e_n \| \| C \| + \exp(t_n k_0) \| v_n \| \| C \| +$$

$$+ \| C \| \| v_n \| \| e_n \| \qquad (2.31)$$

donde $e_n = Y(t_n) - Y_n$ y $v_n = Z(t_n) - Z_n$ para $0 \le n \le N$, $t_n = n\,h$. De (2.9), (2.13) y (2.31) se sigue que:

$$\| X(t_n) - X_n \| \le \| C \| \frac{bh^2}{6} \left\{ \exp(3bq_0)\left(q_0^3+3q_1q_0+q_2\right) + \right.$$

$$\left. + \exp(3bk_0)\left(k_0^3+3k_1k_0+k_2\right) \right\} +$$

$$+ \| C \| \frac{b^2h^4}{36} \exp\left(3b(q_0+k_0)\right) \left(q_0^3+3q_1q_0+q_2\right) \left(k_0^3+3k_1k_0+k_2\right) , \qquad (2.32)$$

válido para $0 \le n \le N$ y para valores de $h<1/q_0$, $h < 1/k_0$.

Nótese que si $X(t)$ es la solución teórica de (2.1) y $A(t)$, $B(t)$ son *funciones dos veces contínuamente diferenciables* en $[0,b]$, entonces:

$$X^{(2)}(t) = A'(t)\, X(t) + A(t)\, X'(t) + X'(t)\, B(t) + X(t)\, B'(t) =$$

$$= A'(t)\, X(t) + (A(t))^2\, X(t) + A(t)\, X(t)\, B(t) + A(t)\, X(t)\, B(t) +$$

$$+ X(t)\, (B(t))^2 + X(t)\, B'(t) \qquad (2.33)$$

Teniendo en cuenta (2.5), (2.10), la relación $X(t) = Y(t)\, C\, Z(t)$, y (2.30), (2.33), se obtiene:

$$\| X^{(2)}(t) \| \le \| C \| \exp\left(b(k_0+q_0) \right) \left(k_1+k_0^2+2q_0k_0+q_0^2+q_1 \right) \quad 0\le t\le b \qquad (2.34)$$

De (2.21) y (2.27), se sigue que:

$$\| X(t) - V(t) \| \le h^2\alpha \,, \qquad 0\le t\le b \qquad (2.35)$$

donde:

$$\alpha = \frac{r}{8} \| C \| \exp\left(b(k_0+q_0) \right) \left(k_1+k_0^2+2q_0k_0+q_0^2+q_1 \right) \qquad (2.36)$$

Por (2.27), (2.28), (2.32), (2.35) y (2.36), la diferencia entre la solución teórica $X(t)$ del problema (2.1), y la función matricial $W(t)$ que interpola los valores aproximados X_n

obtenidos en (2.23), (2.24) y (2.25), está uniformemente acotado superiormente en $[0,b]$ por la desigualdad:

$$\| X(t) - W(t) \| \leq \| X(t) - V(t) \| + \| V(t) - W(t) \| \leq \beta h^2 + \gamma h^4 \tag{2.37}$$

donde:

$$\beta = \alpha + \frac{\| C \| \, b}{6} \left\{ \exp(3bq_0)\left(q_0^3+3q_1 q_0+q_2\right) + \exp(3bk_0)\left(k_0^3+3k_1 k_0+k_2\right) \right\}, \tag{2.38}$$

$$\gamma = \| C \| \frac{b^2}{36} \exp\left(3b(k_0+q_0)\right) \left(q_0^3+3q_1 q_0+q_2\right) \left(k_0^3+3k_1 k_0+k_2\right), \tag{2.39}$$

α está definida en (2.36) y $h < \min\left(\frac{1}{k_0}, \frac{1}{q_0}\right)$.

Dado un *error admisible* $\varepsilon>0$, si escogemos un número positivo h tal que:

$$h < \min\left(\frac{1}{k_0}, \frac{1}{q_0}\right) \quad \text{y} \quad \beta h^2 + \gamma h^4 \leq \varepsilon \tag{2.40}$$

entonces, tomando $N=[b/h]$, si consideramos la sucesión $\left\{ X_n \right\}_{n=0}^{N}$

definida mediante (2.23)-(2.25), y la función B-spline matricial lineal $W(t)$ definida por:

$$W(t) = h^{-1}\left\{ \left(t_{n+1} - t \right) X_n + \left(t - t_n \right) X_{n+1} \right\}$$

$$t_n = nh \ ; \quad t \in [t_n, t_{n+1}] \ ; \quad 0 \le n \le N-1, \tag{2.41}$$

hemos construído una aproximación de la solución del problema (2.1) cuyo error $E(t) = X(t) - W(t)$ satisface:

$$\| X(t) - W(t) \| \le \varepsilon \ , \quad \text{uniformemente para} \quad 0 \le t \le b \tag{2.42}$$

Tomando logaritmos, la desigualdad $\beta h^2 + \gamma h^4 \le \varepsilon$, se satisface bajo la condición $h<1$ y:

$$\ln(h) \le \frac{1}{4}\left\{ \ln(\varepsilon) - \ln(\beta+\gamma) \right\} \tag{2.43}$$

De los anteriores comentarios, concluimos la demostración del resultado siguiente:

TEOREMA 2.1. *Consideremos el problema* (2.1) *donde las funciones matriciales* $A(t)$, $B(t)$ *son dos veces contínuamente diferenciables en* $[0,b]$, *y sean* k_1 *y* q_1 *constantes definidas* *por* (2.5) *y* (2.10) *respectivamente, con* $i = 0, 1, 2$. *Dado un error admisible* $\varepsilon > 0$, *sean* β, γ *las constantes definidas en* (2.38) *y* (2.39) *y sea* h *un número con* $0<h<1$ *satisfaciendo* $h < \min\left(\frac{1}{k_0}, \frac{1}{q_0}\right)$ *y* (2.43). *Sea* $N = [b/h]$ *y sean* $\left\{ Y_n \right\}$, $\left\{ Z_n \right\}$ *las sucesiones finitas definidas mediante* (2.23) *y* (2.24) *respectivamente, con* $0 \le n \le N$. *Entonces, la función matricial* $W(t)$ *definida mediante* (3.28), *interpola la sucesión* $\left\{ X_n \right\}_{n=0}^{N}$ *dada por* (2.41), *y define una aproximación analítica del problema* (2.1), *cuyo error está acotado superiormente por* ε *en todo el dominio* $[0,b]$.

II. B. Caso No Homogéneo.

II. B. (i). Planteamiento Del Problema.

Consideremos la ecuación:

$$X'(t) = A(t)X(t) + X(t)B(t) + L(t); \quad X(0) = C; \quad 0 \le t \le b \qquad (2.44)$$

donde la incógnita X(t) y los coeficientes A(t), B(t) y L(t) son matrices complejas de tamaño rxr, es decir, elementos de $\mathbb{C}^{r\times r}$. La ecuación (2.44), al igual que el caso homogéneo estudiado en la sección anterior, ha sido estudiada por muchos autores en el caso de coeficientes constantes, ver [BARR], [DAVI], [SERB].

De [PORT], [BARN, pág. 109], sabemos que la solución local teórica de la ecuación (2.44) viene dada por:

$$X(t) = Y(t)CZ(t) + Y(t)\left\{\int_0^t Y^{-1}(s)L(s)Z^{-1}(s)\, ds\right\} Z(t) \qquad (2.45)$$

donde Y(t) y Z(t) son las soluciones de los problemas de valores iniciales:

$$Y'(t) = A(t)Y(t); \; Y(0)=I \qquad (2.2)$$

$$Z'(t) = Z(t)B(t); \; Z(0)=I \qquad (2.3)$$

estando $X(t)$ definida en las cercanías de $t=0$, donde las matrices $Y(t)$ y $Z(t)$ son inversibles. Ahora, además de que la solución exacta de los problemas (2.2) y (2.3) no son calculables analíticamente en la práctica, tenemos el *inconveniente añadido de que el intervalo donde* $Y(t)$ *y* $Z(t)$ *son inversibles no está bien determinado.* Todo ello motiva la búsqueda de alternativas que ofrezcan soluciones contínuas aproximadas y cotas al error para el problema (2.44), válidas para todo el intervalo donde la solución está bien definida.

Consideraremos el problema (2.44), donde $A(t)$, $B(t)$ *son funciones dos veces contínuamente diferenciables, y* $L(t)$ *es contínua en el dominio* $0\le t\le b$. En primer lugar, hallaremos las soluciones numéricas en un conjunto discreto de puntos del intervalo $[0,b]$. Dado que las ecuaciones a tratar son idénticas a las resueltas en la sección anterior, podemos utilizar la información de la que disponemos, considerando las mismas definiciones para las constantes k_i y q_i , $i=0,1,2$, que en (2.5) y (2.10) respectivamente; idénticas expresiones para los valores aproximados Y_n y Z_n dados en (2.23) y (2.24) respectivamente; y las mismas cotas para los errores de discretización e_n y v_n dados en (2.9) y (2.13) para cuando utilizamos los métodos lineales unipaso dados por (1.48) y (2.11).

En el siguiente apartado de esta sección, construiremos una solución numérica contínua partiendo de la solución numérica obtenida en el conjunto discreto de puntos e interpolando mediante funciones B-spline matriciales lineales. Proponemos también una cota global del error en términos del espaciado h y de los datos del problema. Así, dado un error admisible $\varepsilon>0$, construiremos la solución aproximada en un intervalo, de forma que el error está acotado superiormente por ε, uniformemente en todo el intervalo.

Este tipo de cuestiones son las que diferencian este caso del anterior homogéneo: consideraciones sobre inversas, determinación del intervalo de trabajo, y, evidentemente, un cálculo más complejo al tratarse de un caso más general.

II. B. (ii). Solución Numérica Contínua y Cota de Error.

Para empezar, notemos que Y_n dada en (2.23), es inversible para $h < \frac{1}{k_0}$ porque por el *lema de perturbación* [ORT2, pág. 45], los factores que aparecen en (2.23) son matrices inversibles. De forma análoga, si $h < \frac{1}{q_0}$, las matrices Z_n dadas en (2.24) que resuelven el problema (2.3), son también, por el mismo motivo, inversibles para $n\geq 0$.

También ahora deberemos tener presente en todo momento, las expresiones dadas anteriormente para la interpolación con funciones B-spline matriciales lineales.

El siguiente resultado determina, en términos de los datos, un intervalo donde la solución teórica (2.45) del problema (2.44) está bien definida.

TEOREMA 2.2. *Sean* k_0 *y* q_0 *definidas en* (2.5) *y* (2.10) *respectivamente, y sea* $r_0 = \max\left\{ k_0 , q_0 \right\}$. *Entonces se verifican los siguientes resultados:*

(i) *Sea* δ *un número positivo,* $\delta \le b$, *tal que:*

$$\ln(\delta) + r_0\delta < -\ln(r_0) \tag{2.46}$$

Entonces, la solución teórica Y(t) *del problema* (2.2) *y* Z(t) *del problema* (2.3) *son ambas inversibles en el intervalo* $0 \le t \le \delta$.

(ii) *Consideremos una partición del intervalo* $[0,\delta]$ *de la forma* $t_0=0$, $t_1=h$, ..., $t_N=Nh=\delta$. *Sean* $\left\{Y_n\right\}_{0\le n\le N}$, $\left\{Z_n\right\}_{0\le n\le N}$ *las sucesiones de matrices definidas en* (2.23) *y* (2.24) *respectivamente, para* $h < \dfrac{1}{r_0}$. *Sean* $S_Y(t)$ *y* $S_Z(t)$ *las funciones matriciales definidas por:*

$$S_Y(t)=h^{-1}\left\{(t_{n+1}-t)Y_n+(t-t_n)Y_{n+1}\right\};\ t_n\le t<t_{n+1};0\le n\le N-1, \tag{2.47}$$

$$S_Z(t)=h^{-1}\left\{(t_{n+1}-t)Z_n+(t-t_n)Z_{n+1}\right\};\ t_n\le t<t_{n+1};0\le n\le N-1, \tag{2.48}$$

para valores $h < \dfrac{1}{r_0}$ *que satisfacen:*

$$h<h_0=\left[\frac{(1/2)\left(1-\delta r_0\exp(\delta r_0)\right)}{\frac{\delta}{6}(r_0^3+3r_1r_0+r_2)\exp(3\delta r_0)+\frac{r}{8}(r_0^2+r_1)}\right]^{1/2}$$

$$r_i = \max\left\{k_i, q_i\right\} ;\quad i = 0,\ 1,\ 2. \qquad (2.49)$$

Entonces, $S_Y(t)$ *y* $S_Z(t)$ *son inversibles en el intervalo* $[0,\delta]$.

DEMOSTRACIÓN. Por [FLET, pág. 114] tenemos:

$$\| Y(t) \| \le \exp(tk_0) \quad y \quad \| Z(t) \| \le \exp(tq_0) \quad \text{para } 0\le t\le b$$

por lo que:

$$\| Y'(t) \| \le k_0\exp(tk_0) \quad y \quad \| Z'(t) \| \le q_0\exp(tq_0) \quad \text{para } 0\le t\le b$$

Nótese que si δ es un número positivo tal que:

$$\delta\ r_0 \exp(\ \delta\ r_0) < 1$$

entonces, para $|t|\le\delta$, obtenemos:

$$\| Y'(t) \|\ \delta < 1 \quad y \quad \| Z'(t) \|\ \delta < 1$$

Como $Y(0)=Z(0)=I$, y:

$$\| Y(t) - I \| < 1 \ ; \quad \| Z(t) - I \| < 1 \qquad \text{para } |t| \le \delta$$

por el *lema de perturbación* [ORT2, pág. 45], las matrices Y(t) y Z(t) son inversibles en el intervalo $[0,\delta]$. Es más, por el teorema de [FREE, pág. 255], se sigue que:

$$\begin{aligned} \| Y^{-1}(t) \| &\le \left(1-r_0\delta\exp(\delta r_0)\right)^{-1} \ ; \quad 0\le t\le\delta \\ \| Z^{-1}(t) \| &\le \left(1-r_0\delta\exp(\delta r_0)\right)^{-1} \ ; \quad 0\le t\le\delta \end{aligned} \tag{2.50}$$

Esto prueba la parte *(i)* del *Teorema 2.2*. Para probar la parte *(ii)*, introduzcamos las funciones B-spline matriciales lineales:

$$T_Y(t)=h^{-1}\left\{(t_{n+1}-t)Y(t_n)+(t-t_n)Y(t_{n+1})\right\}; \ t_n\le t<t_{n+1};0\le n\le N-1 \tag{2.51}$$

$$T_Z(t)=h^{-1}\left\{(t_{n+1}-t)Z(t_n)+(t-t_n)Z(t_{n+1})\right\}; \ t_n\le t<t_{n+1};0\le n\le N-1 \tag{2.52}$$

que interpolan los valores teóricos $Y(t_n)$ y $Z(t_n)$ de las soluciones de los problemas (2.2) y (2.3) respectivamente. Nótese que por (2.9) y (2.13), se sigue que:

$$\| Y(t_n)-Y_n \| \le \frac{h^2\delta}{6} \exp(3\delta k_0) \left(k_0^3+3k_1k_0+k_2\right); \quad 0\le n\le N-1, \qquad (2.53)$$

$$\| Z(t_n)-Z_n \| \le \frac{h^2\delta}{6} \exp(3\delta q_0) \left(q_0^3+3q_1q_0+q_2\right); \quad 0\le n\le N-1. \qquad (2.54)$$

De (2.19) , y (2.51)-(2.54), se tiene que:

$$\| T_Y(t)-S_Y(t) \| \le \frac{h^2\delta}{6} \exp(3\delta k_0) \left(k_0^3+3k_1k_0+k_2\right); \quad 0\le t\le \delta \qquad (2.55)$$

$$\| T_Z(t)-S_Z(t) \| \le \frac{h^2\delta}{6} \exp(3\delta q_0) \left(q_0^3+3q_1q_0+q_2\right); \quad 0\le t\le \delta \qquad (2.56)$$

Para la expresión teórica de la segunda derivada de la solución teórica Y(t) del problema (2.2), así como para la de Z(t) de (2.3), tendremos:

$$\begin{aligned} \max_{0\le t\le\delta} \| Y^{(2)}(t) \| &\le \exp(\delta k_0) \left(k_0^2+k_1\right) \\ \max_{0\le t\le\delta} \| Z^{(2)}(t) \| &\le \exp(\delta q_0) \left(q_0^2+q_1\right) \end{aligned} \qquad (2.57)$$

De (2.21), (2.55) y (2.56), (2.57) queda:

$$\| Y(t) - T_Y(t) \| \le \frac{rh^2}{8} \exp(k_0\delta) \left(k_0^2+k_1\right); \quad 0\le t\le\delta \qquad (2.58)$$

$$\| Z(t) - T_Z(t) \| \le \frac{rh^2}{8} \exp(q_0\delta) \left(q_0^2+q_1\right); \quad 0\le t\le\delta \qquad (2.59)$$

Nótese que de (2.50), (2.55), (2.58), para $h<h_0$ se obtiene:

$$\| S_Y(t)-Y(t)\| \le \frac{1}{2} \left(1-\delta r_0 \exp(\delta r_0)\right) \le \frac{1}{2} \left(\| Y^{-1}(t)\|\right)^{-1}; \quad 0\le t\le\delta \qquad (2.60)$$

y por (2.60) y el *lema de perturbación* [ORT2, pág. 45], la matríz $S_Y(t)$ es inversible para $0\le t\le\delta$. De forma análoga, por (2.50), (2.56) y (2.59), para $h<h_0$ se tiene:

$$\| S_Z(t)-Z(t)\| \le \frac{1}{2} \left(1-\delta r_0 \exp(\delta r_0)\right) \le \frac{1}{2} \left(\| Z^{-1}(t)\|\right)^{-1} \quad ; \quad 0\le t\le\delta$$

y así, $S_Z(t)$ es inversible en $0\le t\le\delta$, con lo que queda el resultado probado. □

De los resultados anteriores y de la expresión (2.45), observamos fácilmente que una solución numérica del problema (2.44) puede ser construida en la forma:

$$\hat{X}(t) = S_Y(t)\ C\ S_Z(t) + S_Y(t) \left\{ \int_0^t S_Y^{-1}(s)\ L(s)\ S_Z^{-1}(s)\ ds \right\} S_Z(t)$$

$$0 \le t \le \delta \qquad (2.61)$$

Dada la partición $0=t_0$; $t_1=h;\ldots;$ $t_N=Nh=\delta$ del intervalo $[0,\delta]$, notar que para t, en el intervalo $t_n \le t < t_{n+1}$, $\hat{X}(t)$ puede ser escrita en la forma:

$$\hat{X}(t) = S_Y(t)CS_Z(t) + S_Y(t)\left\{ \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} S_Y^{-1}(s)L(s)S_Z^{-1}(s)\ ds \right\}S_Z(t) +$$

$$+ S_Y(t) \left\{ \int_{t_n}^{t} S_Y^{-1}(s)L(s)S_Z^{-1}(s)\ ds \right\} S_Z(t) \ ; \quad t_n \le t < t_{n+1} \ ; \quad 1 \le n \le N-1$$

$$(2.62)$$

Para conseguir un cálculo efectivo de $\hat{X}(t)$, interesa obtener una expresión explícita para $S_Y^{-1}(s)$ y $S_Z^{-1}(s)$. Puesto que $S_Y(s)$ y $S_Z(s)$ son inversibles en $[0,\delta]$, por la *fórmula de*

Sherman-Morrison-Woodbury [ORT2, pág. 50], para $t_n \le t < t_{n+1}$ se tiene que:

$$S_Y^{-1}(t) = \frac{hY_n^{-1}\left[I - \dfrac{t-t_n}{t_{n+1}-t}\, Y_{n+1}\left(I + \dfrac{t-t_n}{t_{n+1}-t_n}\, Y_n^{-1}\, Y_{n+1}\right)^{-1} Y_n^{-1}\right]}{t_{n+1} - t} \tag{2.63}$$

$$S_Z^{-1}(t) = \frac{hZ_n^{-1}\left[I - \dfrac{t-t_n}{t_{n+1}-t}\, Z_{n+1}\left(I + \dfrac{t-t_n}{t_{n+1}-t_n}\, Z_n^{-1}\, Z_{n+1}\right)^{-1} Z_n^{-1}\right]}{t_{n+1} - t} \tag{2.64}$$

Es interesante hacer notar que tanto $S_Y^{-1}(t)$ como $S_Z^{-1}(t)$ no presentan problemas en cuanto a acotación se refiere, en $t=t_n$ y $t=t_{n+1}$. En efecto, $S_Y^{-1}(t)$ puede ser escrito en la forma:

$$S_Y^{-1}(t) = \frac{hY_n^{-1}\left[I - (t-t_n)Y_{n+1}\left((t_{n+1}-t)I + (t-t_n)Y_n^{-1}\, Y_{n+1}\right)^{-1} Y_n^{-1}\right]}{t_{n+1} - t} \tag{2.65}$$

Nótese que (2.65) puede interpretarse como el cociente de dos funciones analíticas que se anulan en $t=t_{n+1}$. La derivada del numerador de (2.65), toma la forma:

$$h\, Y_n^{-1} \left[-Y_{n+1} \left((t_{n+1}-t)\, I \; + \; (t-t_n)\, Y_n^{-1}\, Y_{n+1} \right)^{-1} + \right.$$

$$+ \; (t-t_n)\, Y_{n+1} \left((t_{n+1}-t)\, I + (t-t_n)\, Y_n^{-1}\, Y_{n+1} \right)^{-1} \left(-I + Y_n^{-1}\, Y_{n+1} \right) \cdot$$

$$\left. \cdot \left((t_{n+1}-t)\, I \; + \; (t-t_n)\, Y_n^{-1}\, Y_{n+1} \right)^{-1} Y_n^{-1} \right]$$

y su límite cuando t tiende a t_{n+1} con $t<t_{n+1}$ es $-Y_{n+1}^{-1}$. Así, por *el teorema de L'Hopital* [FLET, pág. 39], se tiene que:

$$\lim_{t \longrightarrow t_{n+1}} S_Y^{-1}(t) = Y_{n+1}^{-1} \qquad \text{con } t<t_{n+1}$$

Teniendo en cuenta la expresión de $S_Y^{-1}(t)$ para $t_{n+1} \le t < t_{n+2}$, es fácil comprobar que $S_Y^{-1}(t)$ tiende a Y_{n+1}^{-1} cuando t tiende a t_{n+1} con $t>t_{n+1}$. Así, $S_Y^{-1}(t)$ está acotado en los extremos del intervalo $[t_n, t_{n+1}]$, es contínuo e interpola los valores de Y_n^{-1}. La misma conclusión es válida para $S_Z^{-1}(t)$.

Llamemos $R_n(t)$ al integrando de (2.62) en el intervalo $t_n \le t < t_{n+1}$:

$$R_n(t) = S_Y^{-1}(t)\, L(t)\, S_Z^{-1}(t) \ ; \quad t_n \le t < t_{n+1} \ ; \quad 0 \le n \le N-1, \tag{2.66}$$

cuya expresión explícita puede escribirse de la forma:

$$\begin{aligned} R_n(t) = {} & \left(\frac{h}{t_{n+1}-t} \right)^2 Y_n^{-1} \Bigg\{ L(t) + \\ & - \left(\frac{t-t_n}{t_{n+1}-t} \right) \Bigg[Y_{n+1} \left(I + \frac{t-t_n}{t_{n+1}-t} Y_n^{-1} Y_{n+1} \right)^{-1} Y_n^{-1}(t)\, L(t) + \\ & + L(t)\, Z_n^{-1} Z_{n+1} \left(I + \frac{t-t_n}{t_{n+1}-t} Z_n^{-1} Z_{n+1} \right)^{-1} \Bigg] + \\ & + \left(\frac{t-t_n}{t_{n+1}-t} \right)^2 \Bigg[Y_{n+1} \left(I + \frac{t-t_n}{t_{n+1}-t} Y_n^{-1} Y_{n+1} \right)^{-1} Y_n^{-1}(t)\, L(t)\, Z_n^{-1} Z_{n+1}^{-1} \cdot \\ & \cdot \left(I + \frac{t-t_n}{t_{n+1}-t} Z_n^{-1} Z_{n+1} \right)^{-1} \Bigg] \Bigg\} Z_n^{-1} \end{aligned} \tag{2.67}$$

Con esta notación y teniendo en cuenta (2.61), proponemos una solución aproximada del problema (2.44), definida por:

$$\hat{X}(t) = S_Y(t)\, C\, S_Z(t) + S_Y(t) \left\{ \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} R_j(s)\, ds \right\} S_Z(t) +$$

$$+ S_Y(t) \left\{ \int_{t_n}^{t} R_n(s)\, ds \right\} S_Z(t) \quad ; \quad t_n \le t < t_{n+1} \quad ; \quad 1 \le n \le N-1 \tag{2.68}$$

De (2.45) y (2.61), el error de la solución aproximada $\hat{X}(t)$, viene dado por:

$$X(t) - \hat{X}(t) =$$

$$= \left[Y(t)-S_Y(t)\right] CZ(t) + Y(t)C\left[Z(t)-S_Z(t)\right] - \left[Y(t)-S_Y(t)\right] C \left[Z(t)-S_Z(t)\right] +$$

$$+ \left[Y(t)-S_Y(t)\right] \left\{ \int_0^t \left[Y^{-1}(s)-S_Y^{-1}(s)\right] L(s)\, Z^{-1}(s)\, ds \right\} \left[Z(t)-S_Z(t)\right] +$$

$$+ \left[Y(t)-S_Y(t)\right] \left\{ \int_0^t Y^{-1}(s)\, L(s) \left[Z^{-1}(s)-S_Z^{-1}(s)\right] ds \right\} \left[Z(t)-S_Z(t)\right] +$$

$$- \left[Y(t)-S_Y(t)\right] \left\{ \int_0^t \left[Y^{-1}(s)-S_Y^{-1}(s)\right] L(s) \left[Z^{-1}(s)-S_Z^{-1}(s)\right] ds \right\} \left[Z(t)-S_Z(t)\right] +$$

$$- \left[Y(t)-S_Y(t)\right] \left\{ \int_0^t \left[Y^{-1}(s)-S_Y^{-1}(s)\right] L(s)\, Z^{-1}(s) ds \right\} Z(t) +$$

$$-\left[Y(t)-S_Y(t)\right]\left\{\int_0^t Y^{-1}(s)\ L(s)\ \left[Z^{-1}(s)-S_Z^{-1}(s)\right]\ ds\right\} Z(t)+$$

$$+\left[Y(t)-S_Y(t)\right]\left\{\int_0^t \left[Y^{-1}(s)-S_Y^{-1}(s)\right]\ L(s)\ \left[Z^{-1}(s)-S_Z^{-1}(s)\right]\ ds\right\} Z(t)+$$

$$-Y(t)\left\{\int_0^t \left[Y^{-1}(s)-S_Y^{-1}(s)\right]\ L(s)\ Z^{-1}(s)\ ds\right\}\left[Z(t)-S_Z(t)\right]+$$

$$-\left[Y(t)-S_Y(t)\right]\left\{\int_0^t Y^{-1}(s)\ L(s)\ \left[Z^{-1}(s)-S_Z^{-1}(s)\right]\ ds\right\}\left[Z(t)-S_Z(t)\right]+$$

$$+Y(t)\left\{\int_0^t \left[Y^{-1}(s)-S_Y^{-1}(s)\right]\ L(s)\ \left[Z^{-1}(s)-S_Z^{-1}(s)\right]\ ds\right\}\left[Z(t)-S_Z(t)\right]+$$

$$+Y(t)\left\{\int_0^t \left[Y^{-1}(s)-S_Y^{-1}(s)\right]\ L(s)\ Z^{-1}(s)\ ds\right\} Z(t)+$$

$$+Y(t)\left\{\int_0^t Y^{-1}(s)\ L(s)\ \left[Z^{-1}(s)-S_Z^{-1}(s)\right]\ ds\right\} Z(t)+$$

$$-Y(t)\left\{\int_0^t \left[Y^{-1}(s)-S_Y^{-1}(s)\right]\ L(s)\ \left[Z^{-1}(s)-S_Z^{-1}(s)\right] ds\right\} Z(t)+$$

$$+\left[Y(t)-S_Y(t)\right]\left\{\int_0^t Y^{-1}(s)\ L(s)\ Z^{-1}(s)\ ds\right\} Z(t)+$$

$$+Y(t)\left\{\int_0^t \left[Y^{-1}(s)-S_Y^{-1}(s)\right]\ L(s)\ Z^{-1}(s)\ ds\right\}\left[Z(t)-S_Z(t)\right]+$$

$$-\left[Y(t)-S_Y(t)\right]\left\{\int_0^t Y^{-1}(s)\ L(s)\ Z^{-1}(s)\ ds\right\}\left[Z(t)-S_Z(t)\right] \qquad (2.69)$$

Nótese que para evaluar el error $X(t)-\hat{X}(t)$ en términos de los datos, necesitamos una cota de $\| Y^{-1}(s) - S_Y^{-1}(s) \|$ y de $\| Z^{-1}(s) - S_Z^{-1}(s) \|$ para $0 \le s \le \delta$.

Por el *lema de Banach* [GOLU, pág. 28], y por (2.60), para $0 \le s \le \delta$ se sigue que:

$$\| Y^{-1}(s) - S_Y^{-1}(s) \| \le \| Y^{-1}(s) \|\ \| S_Y^{-1}(s) \|\ \| Y(s) - S_Y(s) \| \le$$

$$\le \| Y^{-1}(s) \| \left\{ \| S_Y^{-1}(s)-Y^{-1}(s) \| + \| Y^{-1}(s) \| \right\} \| Y(s)-S_Y(s) \|$$

$$\| Y^{-1}(s)-S_Y^{-1}(s)\| \le \left(1-\| Y^{-1}(s)\|\ \| Y(s)-S_Y(s)\|\right)^{-1} \| Y^{-1}(s)\|^2\ \| Y(s)-S_Y(s)\| \le$$

$$\le 2\ \| Y^{-1}(s) \|^2\ \| Y(s)-S_Y(s) \| \le h^2 \rho_Y \qquad (2.70)$$

donde:

$$\rho_Y = 2\left(1-\delta r_0 \exp(\delta r_0)\right)^{-1}\left\{ \frac{\delta}{6}\left(k_0^3+3k_1k_0+k_2\right)\exp(3\delta k_0)+\frac{r}{8}\left(k_0^2+k_1\right)\right\} \qquad (2.71)$$

De manera análoga, se obtiene:

$$\| Z^{-1}(s) - S_Z^{-1}(s) \| \leq h^2 \rho_Z \tag{2.72}$$

donde:

$$\rho_Z = 2\left(1-\delta r_0 \exp(\delta r_0)\right)^{-1} \left\{ \frac{\delta}{6}\left(q_0^3 + 3q_1 q_0 + q_2\right) \exp(3\delta q_0) + \frac{r}{8}\,(q_0^2 + q_1) \right\} \tag{2.73}$$

Sea $\mathfrak{L}$ definida mediante:

$$\mathfrak{L} = \max_{0 \leq t \leq b} \| L(t) \| \tag{2.74}$$

Por (2.50) y (2.69)-(2.74) se tiene:

$$\| X(t) - \hat{X}(t) \| \leq \alpha_1 h^2 + \alpha_2 h^4 + \alpha_3 h^6 + \alpha_4 h^8 ; \quad 0 \leq t \leq \delta \tag{2.75}$$

donde:

$$\alpha_1 = \| C \| \left\{ \rho_Y \exp(\delta q_0) + \rho_Z \exp(\delta k_0) \right\} + 4\exp(\delta r_0) \mathfrak{L} \rho_Y \delta \left\{ 1 - \delta r_0 \exp(\delta r_0) \right\}^{-1} \tag{2.76}$$

$$\alpha_2 = \rho_Y \rho_Z \left\{ \| C \| + \delta l \exp(\delta r_0) + \delta \mathfrak{L} \left(1 - \delta r_0 \exp(\delta r_0)\right)^{-2} \right\} ; \quad \alpha_4 = \mathfrak{L}\delta \tag{2.77}$$

$$\alpha_3 = \delta\mathfrak{L} \left\{ 2\left(1-\delta r_0 \exp(\delta r_0)\right)^{-1} + \exp(\delta q_0)\rho_Y^2\rho_Z + \exp(\delta k_0)\rho_Y\rho_Z^2 \right\} \tag{2.78}$$

Ahora nos encontramos en una situación óptima para construir una solución numérica contínua del problema (2.44), con una precisión prefijada. En efecto, dado un error admisible $\varepsilon>0$, es suficiente tomar $h < \min\left\{ h_0 , \frac{1}{r_0} \right\}$ satisfaciendo la condición:

$$\alpha_1 h^2 + \alpha_2 h^4 + \alpha_3 h^6 + \alpha_4 h^8 < \varepsilon$$

o, más directamente, $0<h<1$ tal que:

$$\ln(h) < \frac{1}{8} \ln\left(\frac{\varepsilon}{\alpha_1+\alpha_2+\alpha_3+\alpha_4} \right) \tag{2.79}$$

Entonces, se sigue que:

$$\| X(t) - \hat{X}(t) \| < \varepsilon \text{ ; uniformemente para } 0\le t\le\delta \tag{2.80}$$

y el siguiente resultado queda demostrado:

TEOREMA 2.3. *Consideremos el problema* (2.44) *donde* $A(t)$, $B(t)$ *son funciones matriciales dos veces contínuamente diferenciables, y* $L(t)$ *es contínua en* $[0,b]$. *Sea* δ *definida en Teorema 2.2.*

Entonces, la solución aproximada $\hat{X}(t)$ *del problema* (2.44), *definida por* (2.68) *y la solución exacta* $X(t)$, *están bien definidas en* $[0,\delta]$. *Sea* $\varepsilon>0$ *un error admisible, sean* r_0 *y* h_0 *definidas en* (2.49) , *sean* ρ_Y , ρ_Z , $\mathfrak{L}$, *y* α_i *con* $i = 1, 2, 3, 4$, *las constantes definidas en* (2.71)-(2.78), *y sea* h *tal que satisface la desigualdad* (2.79) *y también que* $h < \min\left\{ h_0 , \frac{1}{r_0} , 1\right\}$ *y* $Nh=\delta$ *para un entero positivo* N, *con* $t_i=ih$ *para* $0\le i\le N$. *Entonces* $\hat{X}(t)$ *es una solución numérica contínua del problema* (2.44), *cuyo error* $X(t)-\hat{X}(t)$ *satisface* (2.80).

CAPITULO III

ECUACION MATRICIAL DE RICCATI CON COEFICIENTES VARIABLES

III. A. Introducción

En este capítulo consideraremos *ecuaciones diferenciales matriciales de tipo Riccati, donde los coeficientes son dos veces contínuamente diferenciables*. El método a seguir será el habitual: en primer lugar, determinaremos un intervalo de existencia en términos de los datos del problema de valores iniciales. Seguidamente, construiremos una solución numérica discreta en un conjunto de puntos del intervalo determinado anteriormente, usando métodos matriciales lineales unipaso. Por último, usando funciones B-spline matriciales lineales, obtendremos una solución numérica contínua con cota del error cometido. Dado un error admisible $\varepsilon>0$, la solución numérica contínua será construida con un error uniformemente acotado superiormente por ε en el dominio de existencia.

Consideremos la ecuación diferencial matricial de Riccati no simétrica y con coeficientes variables, del tipo:

$$W'(t) = C(t)-D(t)W(t)-W(t)A(t)-W(t)B(t)W(t) \;;\quad W(0)=W_0 \qquad (3.1)$$

donde la incógnita es $W(t)$ y los coeficientes $C(t)$, $D(t)$, $A(t)$ y $B(t)$ son funciones matriciales, al menos dos veces contínuamente diferenciables, a valores en $\mathbb{C}^{r\times r}$. Si $A(t)^T$ representa la traspuesta de la matriz $A(t)$, la ecuación de Riccati (3.1) se dice que es simétrica cuando los coeficientes de (3.1) son matrices reales y además $D(t)=A(t)^T$. La ecuación de Riccati (3.1) aparece en diversos campos de la ciencia y la ingeniería, pero desgraciadamente ha recibido poco tratamiento numérico en lo que concierne al estudio del comportamiento y cotas de error a priori en términos de los datos, para las soluciones numéricas propuestas. Más aún, estos estudios suelen centrarse exclusivamente al caso de coeficientes constantes, en contra del hecho de que muchos sistemas reales son de coeficientes variables. Algunas excepciones pueden encontrarse en [DIEC], [JOD3], [KUNK], [KENN], [OSHM]. En [KENN] se propone la integración directa de (3.1) usando rutinas conocidas y un método basado en el *algoritmo de Davison-Maki* modificado. Los resultados numéricos son puestos a prueba en un conjunto de ejemplos concretos sin información de la cota de error en términos de los datos para el caso general. En [OSHM], la solución de la ecuación de Riccati simétrica es reconstruida comenzando por la previa determinación de los valores y vectores propios de la matriz solución. Los autores de [KUNK] consideran ecuaciones diferenciales matriciales de Riccati implícitas, y proponen un método numérico basado en la descomposición contínua de valores singulares de funciones matriciales y formas canónicas de Kronecker.

El estudio de la ecuación (3.1) está estrechamente ligado al sistema lineal asociado:

$$Y'(t)=S(t)Y(t); \quad Y(0) = \begin{bmatrix} I \\ W_0 \end{bmatrix}; \quad Y(t) = \begin{bmatrix} U(t) \\ V(t) \end{bmatrix};$$

$$S(t)=\begin{bmatrix} A(t) & B(t) \\ C(t) & -D(t) \end{bmatrix} \tag{3.2}$$

El método propuesto en [JOD3] está basado en la vectorización del problema (3.2) y métodos aproximados sucesivos. También [JOD3] ofrece cotas de error. Sin embargo, la existencia del intervalo donde está definido el límite que da la solución, es muy pequeño, y además, la especial estructura de la ecuación de Riccati no es tenida en cuenta cuando se aplican sucesivamente los métodos aproximados. Aparte de estos inconvenientes, el método propuesto en [JOD3] converge despacio y resulta computacionalmente caro. En [DIEC] se propone un interesante método de integración numérica para la ecuación de Riccati no simétrica y de coeficientes variables. El principal inconveniente de este método es que requiere la hipótesis de una estructura dicotómica para el sistema lineal asociado (3.2) y, aunque se proponen cotas de error para la solución numérica, éstas están dadas en términos de propiedades geométricas y complicadas constantes relacionadas con esta dicotomía, que no pueden ser conocidas en la práctica.

En este capítulo intentaremos explotar la estructura matricial del problema (3.1) evitando la reformulación estándar del problema como un sistema de ecuaciones vectoriales. El capítulo se organiza como sigue. En ésta primera sección, trataremos la construcción de soluciones numéricas para problemas de valores iniciales del tipo (1.1) bajo la condición de Lipschitz (1.2), tal y como hemos venido aplicando en los capítulos anteriores, por lo que utilizaremos los resultados allí obtenidos. Esta información es usada en la construcción de las soluciones numéricas discretas del problema de Riccati de valores iniciales (3.1). Obtenemos también el intervalo de existencia para la solución del problema (3.1) y una cota superior para el error de la solución numérica en un conjunto de puntos, en términos de los datos. La sección siguiente trata sobre la construcción de una solución numérica contínua con cota de error, mediante la interpolación lineal usando funciones B-spline matriciales.

Por [REI2, pág. 28], la solución de (3.1) viene dada por:

$$W(t) = V(t)\ [U(t)]^{-1} = [0\ I]\ Y(t) \left\{ [I\ 0]\ Y(t) \right\}^{-1} \tag{3.3}$$

donde $U(t)$ y $V(t)$ son bloques correspondientes a la matríz solución del problema (3.2).

El siguiente lema, que extiende el resultado del teorema 1 de [JOD2], determina un intervalo de existencia para la solución del problema (3.1), tan adecuado como permita una cota superior para la norma de $[U(t)]^{-1}$ en dicho intervalo.

LEMA 3.1. *Sean:*

$$k_0 = \max_{0 \le t \le b} \| S(t) \| ; \qquad q_0 = \max_{0 \le t \le b} \| [\ A(t) \quad B(t)\ ] \|$$

y sea δ *un número positivo, menor o igual que* b, *satisfaciendo:*

$$\delta\, k_0 + \ln(\delta) < -\ln \left(q_0 \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| \right) \tag{3.4}$$

Entonces, la solución local W(t) *del problema* (3.1) *está definida por* (3.3) *en el intervalo* $[0,\delta]$. *Es más, si* Y(t) *es la solución de* (3.2), *entonces* $U(t) = [I\ 0]\ Y(t)$ *es inversible en* $[0,\delta]$ *y:*

$$\| [U(t)]^{-1} \| < (1-M\delta)^{-1} \ ; \ M = q_0 \exp(\delta k_0) \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| ; \quad 0 \le t \le \delta \tag{3.5}$$

DEMOSTRACIÓN. Puesto que $U(0) = I$, por el *lema de perturbación* [ORT2, pág. 45], U(t) es inversible si $\| U(t) - I \| < 1$. Nótese que:

$$Y(t) - Y(0) = \int_0^t Y'(s)\, ds = \int_0^t S(s)\, Y(s)\, ds$$

$$U(t) - I = U(t) - U(0) = [I\ 0]\left\{Y(t)-Y(0)\right\} = \int_0^t [A(s)\ B(s)]\ Y(s)\ ds$$

Tomando normas en la última expresión, por la definición de q_0 resulta que:

$$\| U(t) - I \| \leq t\, q_0 \max_{0 \leq s \leq t} \| Y(s) \| \qquad (3.6)$$

Por [FLET, pág. 114], la solución $Y(t)$ del problema (3.2) satisface:

$$\| Y(t) \| \leq \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| \exp(\delta k_0) \ ; \qquad 0 \leq t \leq \delta \qquad (3.7)$$

De (3.6) y (3.7), tomando un número positivo $\delta \leq b$ tal que:

$$\delta\, q_0 \exp(\delta k_0) \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| < 1 \qquad (3.8)$$

entonces, $U(t)$ es inversible en $0 \leq t \leq \delta$. Tomando logaritmos en (3.8) se obtiene (3.4) y la inversibilidad de $U(t)$ en $0 \leq t \leq \delta$. Nótese que:

$$U'(t) = [I\ 0]\ Y'(t) = [I\ 0]\ S(t)\ Y(t) = [A(t)\ B(t)]\ Y(t)$$

y tomando normas en esta última expresión, se sigue que:

$$\| U'(t) \| \le q_0 \exp(\delta k_0) \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| ; \quad 0 \le t \le \delta$$

Ahora, por [FREE, pág. 255] y por (3.8), podemos concluir (3.5). □

Determinado el intervalo de trabajo, el siguiente paso consiste en la aplicación del método lineal unipaso para la obtención de soluciones numéricas en un conjunto discreto de puntos de dicho intervalo.

Consideremos el método unipaso matricial (1.48), asociado al problema de valor inicial (3.2), y dado por:

$$Y_{n+1} - Y_n = \frac{h}{2} \left\{ S(t_{n+1})Y_{n+1} + S(t_n)Y_n \right\} ; \quad Y_0 = \begin{bmatrix} I \\ W_0 \end{bmatrix} \tag{3.9}$$

donde $t_n = nh$, $0 \le n \le N-1$, $N=[\delta/h]$ y δ definido en *Lema 3.1*. Recordar que en este caso tenemos que $C_1 = C_2 = 0$, y $C_3 = -I/12$, por lo que (3.9) define un método unipaso de órden p=2. Las constantes que aparecen para determinar el error de discretización, toman los siguientes valores:

$$G = \| C_3 \| = 1/12 \; ; \quad B = \| B_0 \| + \| B_1 \| = 1;$$

$$\Gamma^* = \left(1 - \frac{hL}{2} \right)^{-1} = \left[1 - \frac{hk_0}{2} \right]^{-1} \tag{3.10}$$

donde k_0 definido en el *Lema 3.1* es la constante de Lipschitz del problema de valor inicial (3.2), y:

$$D \geq \max_{0 \leq t \leq b} \| Y^{(3)}(t) \|$$

Nótese que la solución teórica Y(t) de (3.2) satisface:

$$Y^{(2)}(t) = S'(t)\, Y(t) + (S(t))^2\, Y(t)$$

$$Y^{(3)}(t) = S^{(2)}(t)Y(t)+2S'(t)S(t)Y(t)+S(t)S'(t)Y(t)+(S(t))^3Y(t) \tag{3.11}$$

Denotemos por k_i , $i = 0, 1, 2,$ las constantes positivas correspondientes a nuestro problema, definidas de manera que:

$$k_i \geq \max_{0\leq t\leq\delta} \| S^{(i)}(t) \| \quad ; \quad i= 0,\ 1,\ 2 \tag{3.12}$$

Entonces, por (3.7), (3.11) y (3.12), se sigue que:

$$\max_{0\leq t\leq\delta} \| Y^{(3)}(t)\| \leq \exp(\delta k_0) \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| \left\{k_0^3+3k_1k_0+k_2\right\} \tag{3.13}$$

Obsérvese que tomando $h<1/k_0$, la constante Γ^* definida en (3.10) satisface que $\Gamma^*<2$. Llamando e_n al error de discretización en $t_n=nh$, cuando se aproxima la solución teórica $Y(t_n)$ de (3.2) en t_n mediante el valor Y_n dado por el método (3.9), entonces, por (1.49) tenemos:

$$\| e_n \| \leq \frac{h^2 t_n}{6} \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| \exp\left(\delta k_0\right) \left\{k_0^3+ 3k_1k_0+ k_2\right\} \exp\left(2t_n k_0\right)$$

$$0\leq n\leq N;\ N=[\delta/h] \tag{3.14}$$

Resolviendo (3.9), es fácil ver que:

$$Y_n = \prod_{j=0}^{n-1} \left\{ \left(I - \frac{h}{2} S(t_{n-j}) \right)^{-1} \left(I + \frac{h}{2} S(t_{n-j-1}) \right) \right\} Y_0 \ ;$$

$$N \geq n \geq 1 \ ; \qquad Y_0 = \begin{bmatrix} I \\ W_0 \end{bmatrix} \tag{3.15}$$

Nótese que las matrices Y_n dadas en (3.15) están bien definidas porque para $h < 2/k_0$, por el *lema de perturbación*, las matrices $I - \frac{h}{2} S(t_{n-j})$ son inversibles para $0 \leq j \leq n-1$. Tomando los bloques correspondientes de Y_n, podemos escribir:

$$U_n = [I \ 0] \ Y_n \ ; \quad V_n = [0 \ I] \ Y_n ; \quad n \geq 1 \ ; \ U_0 = I \quad ; \ V_0 = W_0 \tag{3.16}$$

Denotemos por u_n y v_n las sucesiones de matrices definidas por:

$$\begin{aligned} u_n &= U(t_n) - U_n = [I \ 0] \left\{ Y(t_n) - Y_n \right\} \\ v_n &= V(t_n) - V_n = [0 \ I] \left\{ Y(t_n) - Y_n \right\} \end{aligned} \tag{3.17}$$

donde Y_n viene dado por (3.15). Tomando normas en (3.17), se sigue que:

$$\| u_n \| \le \| e_n \| \; ; \quad \| v_n \| \le \| e_n \| \; ; \quad 0 \le n \le N \tag{3.18}$$

Si tomamos un espaciado h tal que:

$$h < \left\{ 6\ (1-M\delta)\ \exp\left(-3\delta k_0\right) \left[\left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| \left\{ k_0^3 + 3k_1 k_0 + k_2 \right\} \delta \right]^{-1} \right\}^{1/2} \tag{3.19}$$

donde M está definida en (3.5). Entonces, por (3.14), (3.18) y (3.5), si h satisface (3.19) se verifica que:

$$\| u_n \| = \| U(t_n) - U_n \| \le 1-M\delta < \| [U(t_n)]^{-1} \|^{-1}; \quad 0 \le n \le N \tag{3.20}$$

y por el *lema de perturbación* se obtiene que la matriz U_n es inversible. De (3.16), tenemos que:

$$W_n = V_n \{ U_n \}^{-1} = \Big\{ [0 \ I] \ Y_n \Big\} \Big\{ [I \ 0] \ Y_n \Big\}^{-1} \ ; \quad 0 \le n \le N \qquad (3.21)$$

es una aproximación numérica al valor teórico $W(t_n)$ de la solución exacta de (3.1) en $t_n = nh$, $0 \le n \le N$, $N = [\delta/h]$.

III. B. Construcción De Una Solución Numérica Contínua, Con Precisión Predeterminada.

Teniendo en cuenta la notación de la sección anterior, y las ideas sobre interpolación lineal con funciones B-spline matriciales expuestas en el segundo apartado de la primera sección del capítulo II, las correspondientes función matriciales lineales que interpolan en el intervalo $[0,\delta]$ las sucesiones U_n, V_n definidas por (3.16), toman la forma:

$$S_U(t)= h^{-1}\left\{\left(t_{n+1}-t\right)U_n+\left(t-t_n\right)U_{n+1}\right\} = [I \ 0] \ S_Y(t) \tag{3.22}$$

$$S_V(t)= h^{-1}\left\{\left(t_{n+1}-t\right)V_n+\left(t-t_n\right)V_{n+1}\right\} = [0 \ I] \ S_Y(t) \tag{3.23}$$

con $t_n \le t < t_{n+1}$ y $Nh = t_N = \delta$.

Nótese que por la expresión de $Y^{(2)}(t)$ se verifica que:

$$\max_{0\le t\le\delta} \| U^{(2)}(t) \| \le \max_{0\le t\le\delta} \| Y^{(2)}(t) \| \le \exp(\delta k_0) \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| \left\{k_0^2+k_1\right\}$$

(3.24)

Con la notación previa, de (3.14), (3.18), (2.19), (2.21) y (3.24) se sigue que:

$$\| U(t) - S_U(t) \| \le \| U(t) - T_U(t) \| + \| T_U(t) - S_U(t) \| \le$$

$$\le h^2 \exp(\delta k_0) \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| \left\{ (r/8)(k_0^2+k_1)+(\delta/6)\exp(2\delta k_0)\{k_0^3+3k_1k_0+k_2\} \right\} =$$

$$= h^2 \gamma \qquad (3.25)$$

Sea M definido en (3.5) y sea h_0 un número positivo definido por:

$$h_0 = \left\{ \frac{1}{2} (1 - M\delta) \gamma^{-1} \right\}^{1/2} \qquad (3.26)$$

De (3.5) y (3.25), si $h<h_0$ se verifica que:

$$\| U(t)-S_U(t) \| \le \frac{1}{2}(1-M\delta) \le \frac{1}{2} \| [U(t)]^{-1} \|^{-1} \; ; \; 0\le t\le\delta \qquad (3.27)$$

y por (3.27) y el *lema de perturbación* [ORT2, pág. 45] se obtiene la inversibilidad de $S_U(t)$ para todo t en el intervalo $[0,\delta]$.

Las expresiones (3.3) y (3.21) indican la forma de obtener la solución analítica aproximada del problema (3.1) en $[0,\delta]$, definida por:

$$\hat{W}(t) = S_V(t)\,[S_U(t)]^{-1} =$$

$$= \left\{\left(t_{n+1}-t\right) V_n + \left(t-t_n\right) V_{n+1}\right\} \left\{\left(t_{n+1}-t\right) U_n + \left(t-t_n\right) U_{n+1}\right\}^{-1}$$

$$t_n \le t \le t_{n+1} \qquad (3.28)$$

Para determinar el error de $\hat{W}(t)$, escribamos:

$$W(t)-\hat{W}(t) = V(t)[U(t)]^{-1} - S_V(t)\,[S_U(t)]^{-1} =$$

$$= \left[V(t) - S_V(t) \right] \left\{ U(t) \right\}^{-1} + V(t) \left[U^{-1}(t) - \left\{S_U(t)\right\}^{-1} \right] -$$

$$- \left[V(t) - S_V(t) \right] \left[U^{-1}(t) - \left\{S_U(t)\right\}^{-1} \right] \qquad (3.29)$$

De una forma análoga a los comentarios previos de (3.24)-(3.25), es fácil ver que para $h<h_0$ y para $0\le t\le\delta$:

$$\| V(t) - S_V(t) \| \le \gamma\, h^2 \qquad (3.30)$$

donde γ está definido en (3.25). Teniendo en cuenta (3.8) y la relación $V(t) = [0 \; I] \; Y(t)$ se sigue que:

$$\| V(t) \| \le \exp(\delta k_0) \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| \quad ; \quad 0 \le t \le \delta \tag{3.31}$$

Si h_0 está definido por (3.26) y $h<h_0$, entonces por (3.27) obtenemos la inversibilidad de $S_U(t)$ en todo el intervalo $[0,\delta]$ y además, por (1.5), podemos escribir:

$$\| \{U(t)\}^{-1} - \{S_U(t)\}^{-1} \| \le$$

$$\le \left(1 - \|\{U(t)\}^{-1}\| \, \|U(t)-S_U(t)\| \right)^{-1} \| \{U(t)\}^{-1} \|^2 \, \| U(t) - S_U(t) \|$$

$$\le 2 \, \| \{U(t)\}^{-1} \|^2 \, \| U(t) - S_U(t) \| \tag{3.32}$$

Ahora, por (3.5), (3.25) y (3.32), podemos escribir:

$$\| \{U(t)\}^{-1} - \{S_U(t)\}^{-1} \| \le 2 \, (1-M\delta)^{-2} \, \gamma \, h^2 \tag{3.33}$$

donde γ está definido en (3.25). Teniendo en cuenta (3.5) y

(3.29)-(3.33), si $0 \le t \le \delta$ y $h < \min\left(h_0, \frac{1}{k_0} \right)$, se sigue que:

$$\| W(t) - \hat{W}(t) \| \le \alpha_1 h^2 + \alpha_2 h^4 \quad \text{para} \quad 0 \le t \le \delta \tag{3.34}$$

donde:

$$\alpha_1 = \gamma \, (1-M\delta)^{-1} \left\{ 1 + 2 \, (1-M\delta)^{-1} \exp\left(\delta k_0\right) \left\| \begin{bmatrix} I \\ W_0 \end{bmatrix} \right\| \right\}$$

$$\alpha_2 = 2 \, (1-M\delta)^{-2} \, \gamma^2 \tag{3.35}$$

Por tanto, dado un error admisible ε, tomando $h < \min\left\{ h_0 \, , \, \frac{1}{k_0} \, , \, 1 \right\}$ tal que:

$$\ln(h) < \frac{1}{4} \ln\left(\frac{\varepsilon}{\alpha_1 + \alpha_2} \right) \tag{3.36}$$

concluimos que $\hat{W}(t)$ definida en (3.28), es una solución aproximada del problema (3.1), cuyo error está uniformemente acotado superiormente por ε en todo el intervalo $[0,\delta]$. Así pues, el siguiente resultado queda probado.

TEOREMA 3.1. *Con la anterior notación, sea* $\varepsilon>0$ *y sean* δ *y* M *las constantes definidas en el Lema 3.1. Sea* h *un número positivo tal que* $Nh=\delta$, $h < \min\left\{ h_0 , \frac{1}{k_0}, 1 \right\}$ *y satisface (3.36). Sean* U_n *y* V_n *las sucesiones de matrices definidas en (3.21), donde* Y_n *viene dado por (3.15). Entonces,* $\hat{W}(t)$ *definido por (3.28) en el intervalo* $t_n=nh \le t \le t_{n+1}=(n+1)h$, *para* $0\le n\le N-1$, *es una solución aproximada contínua del problema (3.1), cuyo error es uniformemente acotado superiormente por* ε *en* $[0,\delta]$.

NOTA 3.1. Recordemos un importante comentario que ya apuntamos en el capítulo anterior. Es fácil ver, usando el *teorema de L'Hôpital* [FLET, pág. 39], que $S_U(t)$ definido en (3.22), es tal que su inversa $\left(S_U(t)\right)^{-1}$ es una función contínua en $[0,\delta]$, satisfaciendo:

$$\lim_{t \to t_{n+1}} \{S_U(t)\}^{-1} = \{U_{n+1}\}^{-1} \quad ; \quad \lim_{t \to t_n} \{S_U(t)\}^{-1} = \{U_n\}^{-1}$$

Esto indica que $\{S_U(t)\}^{-1}$ interpola la sucesión de matrices $\left\{ (U_n)^{-1} \right\}_{0\le n\le N}$ y que $\hat{W}(t)$ es una función interpoladora de $\left\{ W_n \right\}_{0\le n\le N}$ en $[0,\delta]$.

NOTA 3.2. Nótese que la longitud δ del intervalo de existencia predeterminado, donde se construye la solución numérica del problema (3.1), está expresada en términos de los datos por la desigualdad (3.4). Sin embargo, las constantes positivas γ, M, α_1 , α_2 y h_0 dependen del parámetro δ, y para un valor fijado de $h < \min\left\{ h_0 , \frac{1}{k_0} , 1 \right\}$, la expresión $\alpha_1 h^2 + \alpha_2 h^4$ que aparece en el segundo miembro de (3.35), *depende muy sensiblemente* del valor del parámetro δ. Un pequeño cambio en δ provoca un cambio importante en el error dado por (3.35). Este hecho se ilustra con el siguiente ejemplo.

EJEMPLO 3.1. Consideremos el problema (3.1) con los datos:

$$A(t) = D(t) = \frac{1}{2} t^2 \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} ; \qquad C(t) = \frac{1}{2} t^2 \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$B(t) = t^2 \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} ; \quad b=1 \ ; \quad r=2 \ ; \quad W_0=0$$

La función matricial S(t) del problema asociado (3.2), toma la forma:

$$S(t) = \frac{1}{2}\, t^2 \begin{bmatrix} 1 & -1 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & -1 & 1 \\ 0 & -1 & 0 & -1 \end{bmatrix} = \frac{1}{2}\, t^2\, S$$

con $S'(t) = tS$; $S^{(2)}(t) = S$.

Calculando tenemos:

$$k_0 = \frac{1}{2}\, \|S\| = 1.727673345$$

$$q_0 = \left\| \begin{bmatrix} 1/2 & -1/2 & 0 & 1 \\ 0 & 1/2 & 0 & 1 \end{bmatrix} \right\| = 2.618033989$$

$$k_1 = k_2 = \|S\| = 3.45534669$$

y la desigualdad (3.4), adquiere la forma:

$$1.727673345\, \delta + \ln(\delta) + 0.9624236502 < 0 \qquad (3.37)$$

Tomemos tres valores diferentes de δ que satisfagan (3.37):

$$\delta_1 = 0.24 \qquad \delta_2 = 0.22 \qquad \delta_3 = 0.20$$

Los valores correspondientes de γ y M, definidos en (3.25) y (3.5) resultan:

$\gamma_1=\gamma(\delta_1)=7.581998425$; $\gamma_2=\gamma(\delta_2)=6.810595563$; $\gamma_3=\gamma(\delta_3)=6.133829704$

$M_1=M(\delta_1)=3.963253398$; $M_2=M(\delta_2)=3.828648183$; $M_3=M(\delta_3)=3.698614608$

Las constantes α_1 y α_2 dadas en (3.35) asociadas a los valores elegidos de δ, toman los siguientes valores:

$$\delta = \delta_1 = 0.24 \begin{cases} \alpha_1 = 9787.146553 \\ \alpha_2 = 48241.00365 \end{cases} \qquad \delta = \delta_1 = 0.22 \begin{cases} \alpha_1 = 844.1945465 \\ \alpha_2 = 3730.363158 \end{cases}$$

$$\delta = \delta_1 = 0.20 \begin{cases} \alpha_1 = 279.39783 \\ \alpha_2 = 1110.763443 \end{cases}$$

Las tablas siguientes muestran la sensibilidad de la cota de error $\alpha_1 h^2 + \alpha_2 h^4$ para los diferentes valores de δ:

$\delta = 0.24$		
Número de puntos N	Tamaño h	Cota de error $\alpha_1 h^2 + \alpha_2 h^4$
10	$2.4 \cdot 10^{-2}$	5.653401622
100	$2.4 \cdot 10^{-3}$	$5.637556467 \cdot 10^{-2}$
200	$1.2 \cdot 10^{-3}$	$1.409359107 \cdot 10^{-2}$
500	$4.8 \cdot 10^{-4}$	$2.254961127 \cdot 10^{-3}$
1000	$2.4 \cdot 10^{-4}$	$5.637398015 \cdot 10^{-4}$

$\delta = 0.22$		
Número de puntos N	Tamaño h	Cota de error $\alpha_1 h^2 + \alpha_2 h^4$
10	$2.2 \cdot 10^{-2}$	0.4094640205
100	$2.2 \cdot 10^{-3}$	$4.085988991 \cdot 10^{-3}$
200	$1.1 \cdot 10^{-3}$	$1.021480863 \cdot 10^{-3}$
500	$4.4 \cdot 10^{-4}$	$1.634362040 \cdot 10^{-4}$
1000	$2.2 \cdot 10^{-4}$	$4.085902479 \cdot 10^{-5}$

$\delta = 0.20$		
Número de puntos N	Tamaño h	Cota de error $\alpha_1 h^2 + \alpha_2 h^4$
10	$2\cdot10^{-2}$	0.1119368542
100	$2\cdot10^{-3}$	$1.117609092\cdot10^{-3}$
200	$1\cdot10^{-3}$	$2.793989408\cdot10^{-4}$
500	$4\cdot10^{-4}$	$4.470368124\cdot10^{-5}$
1000	$2\cdot10^{-4}$	$1.117591498\cdot10^{-5}$

CAPITULO IV

PROBLEMA GENERAL. APLICACIONES.

IV. A. Problema De Valores Iniciales General.

IV. A. (i). PLANTEAMIENTO.

En este capítulo vamos a aplicar el método que hemos venido utilizando, para resolver problemas matriciales de valores iniciales del tipo (1.1), el más general posible:

$$Y'(t) = f(t,Y(t)); \quad Y(0)=\Omega\in\mathbb{C}^{r\times q} \; ; \quad 0\le t\le b$$

donde $f\colon [0,b]\times\mathbb{C}^{r\times q} \longrightarrow \mathbb{C}^{r\times q}$ es acotada, contínua, y satisface la condición de Lipschitz (1.2):

$$\| f(t,P) - f(t,Q) \| \le L \| P - Q \|$$

que asegura la existencia de una única solución $Y(t)$, la cual es contínuamente diferenciable en $[0,b]$, ver [FLET, pág. 99].

Los resultados que aquí exponemos, pueden ser encontrados en la referencia [JOD6].

Al aplicar el método unipaso (1.48), necesitamos determinar las constantes que aparecen en el error de discretización (1.49) en términos de los datos del problema. Para ello, notar que a la hora de calcular cotas de error para las derivadas de la solución teórica del problema (1.1), debemos utilizar los conceptos expuestos en el apartado de la segunda sección del capítulo I, dedicado a cálculo diferencial matricial, debido a que una de las variables de la función que aparece en (1.1) es matricial.

Introduzcamos la función:

$$f:\ [0,b] \times \mathbb{C}^{p\times q} \longrightarrow \mathbb{C}^{p\times q}$$

que denotaremos mediante f(t,P). Sean las funciones g_1 y g_2 definidas por:

$$g_1 : [0,b] \longrightarrow [0,b] \qquad\qquad g_2 : [0,b] \longrightarrow \mathbb{C}^{p\times q}$$

$$g_1(t) = t \qquad\qquad g_2(t) = Y(t)$$

donde Y(t) es la solución teórica de (1.1). Describamos ahora f(t,Y(t)) como la composición de las funciones f y (g_1,g_2), es decir:

$$\Psi:\ [0,b] \longrightarrow \mathbb{C}^{p\times q}$$

$$\Psi(t) = \Big[f\circ(g_1,g_2)\Big](t) = f\Big(g_1(t),g_2(t)\Big) = f(t,Y(t))$$

Así, Ψ es función de variable real t, y por el teorema 8.9.2 de [GRAH, pág. 170], su diferencial toma la forma:

$$D\Psi = D(f\circ(g_1,g_2)) = \Big((D_1f)(g_1,g_2)\Big)\cdot Dg_1 + \Big((D_2f)(g_1,g_2)\Big)\cdot Dg_2$$

donde hemos supuesto la existencia y continuidad de las derivadas parciales $D_1 f$ y $D_2 f$.

Por los comentarios anteriores se tiene que:

$$\Psi'(t)=\frac{\partial f(t,Y(t))}{\partial t} + \left([\mathrm{vec}\ f(t,Y(t))]^T \otimes I_p \right) \frac{\partial f(t,Y(t))}{\partial\ \mathrm{vec}\, \mathcal{Y}} \qquad (4.1)$$

y:

$$\Psi^{(2)}(t)=\frac{\partial^2 f(t,Y(t))}{\partial t^2} + \left([\mathrm{vec}\ f(t,Y(t))]^T \otimes I_p \right) \frac{\partial}{\partial t}\left(\frac{\partial f(t,Y(t))}{\partial\ \mathrm{vec}\, \mathcal{Y}} \right)$$

$$+ \left[\frac{\partial [\mathrm{vec} f(t,Y(t))]^T}{\partial t} \otimes I_p \right] \frac{\partial f(t,Y(t))}{\partial\ \mathrm{vec}\, \mathcal{Y}} +$$

$$+ \left([\mathrm{vec}\ f(t,Y(t)]^T \otimes I_p \right) \frac{\partial}{\partial\ \mathrm{vec}\, \mathcal{Y}} \left(\frac{\partial f(t,Y(t))}{\partial t} \right) +$$

$$+ \left([\mathrm{vec}\ f(t,Y(t)]^T \otimes I_p \right) \left(\frac{\partial [\mathrm{vec} f(t,Y(t))]^T}{\partial\ \mathrm{vec}\, \mathcal{Y}} \otimes I_p \right) \frac{\partial f(t,Y(t))}{\partial\ \mathrm{vec}\, \mathcal{Y}} +$$

$$+ \left([\mathrm{vec}\ f(t,Y(t)]^T \otimes I_p \right) \left([\mathrm{vec}\ f(t,Y(t)]^T \otimes I_{p^2 q} \right) \frac{\partial^2 f(t,Y(t))}{(\partial \mathrm{vec} \mathcal{Y})^2} \qquad (4.2)$$

donde hemos usado que:

$$\frac{d\,[\text{ vec } Y(t)\,]^T}{dt} = [\text{ vec } f(t,Y(t)\,]^T$$

y se supone también la existencia y continuidad de la segunda derivada parcial de $f(t,P)$.

Notar que la derivada de la solución teórica $Y(t)$ del problema (1.1) está relacionada con las derivadas de Ψ mediante las expresiones:

$$Y^{(2)}(t)=\Psi'(t) \quad \text{y} \quad Y^{(3)}(t)=\Psi^{(2)}(t) \tag{4.3}$$

De una forma análoga podríamos obtener expresiones para derivadas de $Y(t)$ de mayor órden.

IV. A. (ii). Construcción de Una Solución Numérica Contínua.

Suponer que hemos obtenido los valores aproximados Y_{n+1} de las soluciones teóricas $Y(t_{n+1})$ de la ecuación (1.1) mediante la aplicación del método (1.48). Queremos interpolar estos valores con el B-spline lineal dado en (2.47). Aplicando (2.15), podemos obtener una expresión más explícita para esta función de B-spline, dada por:

$$S_Y(t) = h^{-1} \left\{ \left(t_{n+1} - t\right) Y_n + \left(t - t_n \right) Y_{n+1} \right\} \tag{4.4}$$

$$t_n \leq t < t_{n+1} \quad ; \qquad 0 \leq n \leq N-1$$

Aplicando (2.19) y (2.21), pero con la notación seguida ahora, se sigue que:

$$|| Y(t) - S_Y(t) || \leq || Y(t) - T_Y(t) || + || T_Y(t) - S_Y(t) || \leq$$

$$\leq (rq)^{(1/2)} \frac{h^2}{8} \max_{0 \leq t \leq b} || Y^{(2)}(t) || + \max_{0 \leq n \leq N} || e_n || \tag{4.5}$$

siendo e_n el error global de discretización del método usado para obtener Y_n , y $T_Y(t)$ el B-spline lineal que interpolaría los teóricos valores exactos de la las soluciones de la ecuación (1.1).

El error de discretización e_n lo tenemos calculado según (1.49). Así, sustituyendo en (4.5), queda:

$$|| Y(t) - S_Y(t) || \leq \gamma h^2$$

uniformemente para $t \in [0,b]$, donde:

$$\gamma = (rq)^{(1/2)} \frac{D_2}{8} + \Gamma^* G D_3 b \exp\left(\Gamma^* B^* L b \right)$$

Donde, recordemos que:

$$D_3 \geq \max \| Y^{(3)}(t) \| \quad ; \quad D_2 \geq \max \| Y^{(2)}(t) \|$$

son calculables por (4.1)-(4.3), y el resto de constantes están expresadas en (1.18), (1.33), (1.43), (1.45) y (1.46).

IV. B. CÁLCULO DE FUNCIONES INVERSAS.

IV. B. (i). Funciones Inversas Como Soluciones De Un Problema Diferencial Matricial De Valores Iniciales.

En esta sección proponemos un procedimiento constructivo que ofrezca funciones contínuas aproximadas, y cotas de error para la inversa local de una función diferenciable de clase C^m, que actúa entre espacios de Banach de dimensión finita. Más concretamente, utilizaremos funciones del tipo $f: \Delta \in E \longrightarrow E$ las cuales sean contínuamente diferenciables de clase C^m en un conjunto abierto Δ, que contiene un disco centrado en el origen del espacio de Banach E de dimensión finita, satisfaciendo las condiciones:

$$f(0)=0, \text{ y } Df(0) \text{ es un isomorfismo} \tag{4.6}$$

Antes de nada, determinamos en términos de los datos, la región donde la función f, de clase C^m en Δ y que satisface (4.6), es inversible, y veremos también, que esta inversa queda expresada en términos de la solución de un problema diferencial de valores iniciales, dependiente de parámetros. Finalmente, en el siguiente apartado de esta misma sección, construiremos la aproximación de la función inversa con un margen de error dentro de un valor prefijado, utilizando las técnicas que hemos venido desarrollando a lo largo de esta memoria. Incluimos también un ejemplo ilustrativo.

Si f es una función diferenciable $f: E \longrightarrow E$, donde E es un espacio de Banach de dimensión finita, denotaremos por $\| Df(x) \|$ al supremo del conjunto:

$$\left\{ \left\| \left(Df(x)\right)(v) \right\| ; \quad v \in E ; \quad \| v \| \leq 1 \right\}$$

Al disco abierto de radio $r>0$ centrado en el orígen de E lo denotaremos por U_r y a su correspondiente disco cerrado por D_r. El conjunto de todas las funciones lineales $u: E \longrightarrow E$, dotado con el operador norma $\| u \| = \sup \left\{ \| u(x) \| ; \quad \| x \| \leq 1 \right\}$ es un espacio de Banach que denotaremos por $\mathcal{L}(E,E)$.

Establecidas esta serie de cuestiones previas, comencemos con un lema que determina la región donde la ecuación diferencial satisfecha por la función inversa, está bien definida.

LEMA 4.1. *Sea* E *un espacio de Banach de dimensión finita y sea* Δ *un conjunto abierto que contiene un disco* D_r *de radio* $r>0$ *centrado en el orígen. Sea* $f: \Delta \longrightarrow E$ *una función de clase* C^2 *tal que* $Df(0)$ *es un isomorfismo.*

(i) *Sea* $M_0 = \sup \left\{ \| D^2 f(y) \| ; \quad \| y \| \leq r \right\}$, *y sea* $r'>0$ *definido así:*

$$r' = \min \left\{ r \; , \; \left(2 \; \| (Df(0))^{-1} \| \; M_0 \right)^{-1} \right\}$$

entonces para $x \in U_r$ *,* $Df(x)$ *es un isomorfismo, y además:*

$$\sup \left\{ \| (Df(z))^{-1} \| \; ; \; \| z \| < r^{-1} \right\} \le 2 \| (Df(0))^{-1} \| \qquad (4.7)$$

(ii) *Consideremos la función* $H: U_{r'} \longrightarrow \mathcal{L}(E,E)$ *definida como* $H(x)=(Df(x))^{-1}$. *Entonces:*

$$\sup \left\{ \| H(x) \| \; ; \; \| x \| < r' \right\} \le 4 \; M_0 \left\| \left(Df(0) \right)^{-1} \right\|^2 \qquad (4.8)$$

donde M_0 *está definido en la parte (i) del lema.*

Sea $\gamma > 0$ *y sea* $G: U_r \times U_\gamma \longrightarrow E$ *definida como:*

$$G(x,v) = \left(Df(x) \right)^{-1} (v) \qquad (4.9)$$

Entonces G *satisface la condición de Lipschitz:*

$$\| G(x,v) - G(y,v) \| \le 4\,\gamma\, M_0 \;\| (Df(0))^{-1} \|^2 \;\| x - y \|$$

$$x,\ y \in U_r\ ,\qquad v \in U_\gamma \qquad\qquad (4.10)$$

DEMOSTRACIÓN. (i) Por el *teorema del valor medio* [DIEU, pág.158] si $x \in U_r$ se tiene que:

$$\| Df(x) - Df(0) \| \le \| x \|\ M_0$$

Por la definición de r', si $\| x \| < r'$ se tiene que:

$$\| Df(x) - Df(0) \| < \| (Df(0))^{-1} \|^{-1}$$

Por el *lema de perturbación* [DUNF, pág. 584] y la anterior desigualdad, se sigue la inversibilidad de Df(x).

Por el *lema de Banach*, el *teorema del valor medio* , y la definición de r', se tiene que:

$$\left\| \left(Df(x)\right)^{-1} - \left(Df(0)\right)^{-1} \right\| \le \left\| \left(Df(x)\right)^{-1} \right\| \left\| \left(Df(0)\right)^{-1} \right\| \left\| Df(x) - Df(0) \right\| <$$

$$< \left\| \left(Df(x)\right)^{-1} \right\| \left\| \left(Df(0)\right)^{-1} \right\| r'\ M_0$$

$$\left\| \left(Df(x)\right)^{-1} \right\| - \left\| \left(Df(0)\right)^{-1} \right\| < \left\| \left(Df(x)\right)^{-1} \right\| \; \left\| \left(Df(0)\right)^{-1} \right\| r' M_0$$

$$\left\| \left(Df(x)\right)^{-1} \right\| \left(1 - r' M_0 \left\| \left(Df(0)\right)^{-1} \right\| \right) < \left\| \left(Df(0)\right)^{-1} \right\|$$

$$\left\|\left(Df(x)\right)^{-1}\right\| < \left(1 - r'_0 M \|(Df(0))^{-1}\|\right)^{-2} \left\|\left(Df(0)\right)^{-1}\right\| < 2 \left\|\left(Df(o)\right)^{-1}\right\|$$

Entonces, (4.7) queda probado.

(ii) Nótese que $H(x) = (g_2 \circ g_1)(x)$, donde $g_1 : U_{r'} \longrightarrow L(E,E)$ y $g_2 : L(E,E) \longrightarrow L(E,E)$, vienen definidas por: $g_1(x) = Df(x)$ y $g_2(y) = y^{-1}$, respectivamente. Por el teorema 8.2.1. de [DIEU, pág. 149] se tiene que $H'(x) = g'_2(g_1(x)) \cdot g'_1(x)$. Así, por el teorema 8.2.2. de [DIEU, pág. 151] se sigue que:

$$\left\| g'_2\left(g_1(x)\right) \right\| = \left\| g'_2\left(Df(x)\right) \right\| \leq \left\| \left(Df(x)\right)^{-1} \right\|^2 \tag{4.11}$$

Teniendo en cuenta que $g'_1(x) = D^2f(x)$, de (4.11) se sigue que:

$$\| H'(x) \| \le \left\| \left(Df(x)\right)^{-1} \right\|^2 \| D^2f(x) \| \tag{4.12}$$

Por (2.12) y (4.7), se obtiene (4.8).

Notar que $G(x,v) = (Df(x))^{-1}(v) = H(x)(v)$. Por el *teorema del valor medio* y por (4.8), se establece que:

$$\| G(x,v) - G(x,y) \| = \| H(x)(v) - H(y)(v) \| = \| [H(x)-H(y)](v) \|$$

$$\le \| H(x) - H(y) \| \, \| v \| \le \| H(x) - H(y) \| \gamma \le$$

$$\le \sup \left\{ \| H'(z) \| , \ \| z \| < r' \right\} \gamma \| x - y \|$$

Por (4.12) y (4.6) obtenemos (4.10). □

TEOREMA 4.1. *Sea* E *un espacio de Banach de dimensión finita, y sea* Δ *un conjunto abierto de* E *que contiene al disco cerrado* D_r *de radio* $r>0$ *y centrado en el origen de E. Sea* $f\colon \Delta \longrightarrow E$ *una función diferenciable de clase* C^m, $m\ge 2$, *tal que* $f(0)=0$ *y* $Df(0)$ *es un isomorfismo.*

(i) *Sea* $r'>0$ *definido mediante el Lema 4.1 y consideremos el* *sistema diferencial dependiente de parámetros:*

$$\frac{dx}{dt} = G(x,v) \;;\quad x(0,v)=0 \;;\quad \| v \| < \gamma \tag{4.13}$$

donde $G(x,v)$ *está definida en* (4.9) *y* $\gamma>0$. *Sea* δ *definida por:*

$$\delta = 2 \, \| (Df(0))^{-1} \| \, r' \left[1 + 2\, r' \, M_0 \, \gamma \, \| (Df(0))^{-1} \| \right] \tag{4.14}$$

Entonces, el sistema (4.13) *está bien definido y tiene una única solución en el intervalo* $]-\delta,\delta[$.

(ii) *Sea* δ *definida en* (4.14) *y sea* $\varepsilon = \frac{\gamma\delta}{2}$. *Entonces la función* $f\colon U_{r'} \longrightarrow U_\varepsilon$ *admite inversa* $g\colon U_\varepsilon \longrightarrow U_{r'}$ *definida por:*

$$g(v) = x\left(\frac{\delta}{2} \,,\, \frac{2v}{\delta} \right) , \qquad v \in U_\varepsilon$$

donde $x(t,v)$ *es la solución de* (4.13).

DEMOSTRACIÓN. (i) Por el *Lema 4.1*, $Df(x)$ es un isomorfismo para $x \in U_{r'}$ y por lo tanto, el problema (4.13) queda bien definido.

Por los teoremas 10.7.1., 10.7.3. y 10.7.4. de [DIEU], para cada $x \in U_\gamma$ existe una única solución $t \longrightarrow x(t,v)$ del problema (4.13), de clase C^{m-1} y definida en $]-\delta,\delta[$ tal que:

$$df\left(x(t,v)\right) \frac{\partial x(t,v)}{\partial t} = v$$

Consecuentemente, $\dfrac{df(x(t,v))}{dt} = v$ y entonces:

$$f\left(x(t,v)\right) = tv + \Psi(v)$$

Tomando $t=0$, se sigue que $\Psi=0$, y así:

$$f\left(x(t,v)\right) = tv \tag{4.15}$$

(ii) Sea δ definida por (4.13) y sea $g(v) = x\left(\dfrac{\delta}{2}, \dfrac{2v}{\delta}\right)$ para $v \in U_\varepsilon$ con $\varepsilon = \frac{\gamma\delta}{2}$. Por (4.15) se tiene que:

$$(f\circ g)(v) = f\left(x\left(\frac{\delta}{2}, \frac{2v}{\delta}\right)\right) = v$$

y $f\circ g=id$. Así, g es la inversa de f por la derecha, de clase C^{m-1} donde $g: U_\varepsilon \longrightarrow U_r$. Por otro lado, como $(Df)^{-1}$ es de clase C^{m-1}, la identidad $Dg = (Df)^{-1} \circ g$ muestra que g es de clase C^m. Es más, por la igualdad $f\circ g=id$, se tiene que $Df(0) \circ Dg(0) = id$. Puesto que $Df(0)$ es un isomorfismo, concluimos que $Dg(0)$ es también un isomorfismo. Entonces, podemos aplicar la primera parte de la demostración para la función g, obteniendo su inversa por la derecha h, es decir, $g\circ h=id$ en una región apropiada de los alrededores del orígen. De las ecuaciones $f\circ g=id$, $g\circ h=id$, se sigue que f=h y el resultado queda probado. □

El siguiente ejemplo muestra la utilidad del anterior teorema en un caso donde la función inversa es conocida.

EJEMPLO 4.1. Sea $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$, la función $f(x) = Ax$, donde $x \in \mathbb{R}^n$ y A es una matriz cuadrada de órden n, inversible. Puesto que $f'(x) = A$, el sistema diferencial correspondiente a (4.13), toma la forma:

$$\frac{dx(t,v)}{dt} = A^{-1}(v) \quad , \qquad x\left(0\ ,\ v\right) = 0$$

Integrando, obtenemos $x(t,v) = A^{-1}\,v\,t$ y la solución está definida en toda la recta real. Tomando cualquier valor para r y calculando su correspondiente δ según el *Teorema 4.1*, la función inversa es:

$$g(v) = x\left(\frac{\delta}{2}\ ,\ \frac{2v}{\delta}\right) = A^{-1}\,v$$

IV. B. (ii). Cálculo De La Función Inversa Aproximada Y Cota De Error.

Recordemos la expresión de la *regla de la cadena* para cuando derivamos respecto a matrices, que ya indicamos en el capítulo I. Si $X \in \mathbb{C}^{m\times n}$, $Y \in \mathbb{C}^{u\times v}$, $Z \in \mathbb{C}^{p\times q}$, tenemos:

$$\frac{\partial Z}{\partial X} = \left[\frac{\partial [\mathrm{vec}Y]^T}{\partial X} \otimes I_p \right] \left[I_n \otimes \frac{\partial Z}{\partial \mathrm{vec}Y} \right]$$

Si consideramos la solución teórica $x(t,v)$ del problema (4.13) en el intervalo $[0,\delta^*]$ donde $\delta^* < \delta$ y δ está definido en (4.14), entonces, por la expresión dada para la *regla de la cadena*, la segunda derivada de la solución $x(t,v)$ del problema (4.13) toma la forma:

$$\frac{d^2x(t,v)}{dt^2} = \left\{ \left[\mathrm{vec}\left(Df(x(t,v)) \right)^{-1}(v) \right]^T \otimes I_n \right\} \frac{\partial (Df(x(t,v))^{-1}(v)}{\partial \mathrm{vec}X} \qquad (4.16)$$

donde n es la dimensión del espacio de Banach E. Teniendo en cuenta que por [LANC, pág. 439] se tiene que:

$$\| A \otimes B \| = \| A \| \| B \| \qquad (4.17)$$

por el *Lema 4.1*, *Teorema 4.1*, la expresión de *la regla de la cadena*, (4.16), y (4.17), tendremos que:

$$\sup \left\{ \left\| \frac{d^2x(t,v)}{dt^2} \right\| \; ; \; 0 \le t \le \delta^* < \delta \; ; \; \| v \| < \gamma \right\} \le$$

$$\le 8 \left\| \left(Df(0) \right)^{-1} \right\|^3 M_0 \gamma^2 \tag{4.18}$$

Aplicaremos esta vez el método unipaso definido en (1.50), correspondiente a la generalización al caso matricial del *método de Euler*. La justificación de esta decisión es la sencillez de cálculo, pues vamos a resolver un ejemplo concreto en el que las ecuaciones a las que conduce el *método de los trapecios* son implícitas y por ello, bastante más complejas que las que obtendremos con el *método de Euler*.

Entonces, para el problema (4.13), la constante D_2 que aparece en la expresión del error de discretización para este método, dado por (1.52), toma la forma:

$$D_2 = 8 \left\| \left(Df(0) \right)^{-1} \right\|^3 M_0 \gamma^2 \tag{4.19}$$

donde se ha tenido en cuenta (4.18), los anteriores comentarios y la expresión:

$$D_2 \geq \left\{ \| Y^{(2)}(t) \| \ ; \ t\in[0,b] \right\}$$

El siguiente resultado recapitula el proceso para la construcción aproximada de la función inversa, siendo su demostración una consecuencia directa de *Lema 4.1*, *Teorema 4.1*, (1.52) y (4.19).

TEOREMA 4.2. *Sea $\gamma>0$ y sea E un espacio de Banach de dimensión finita. Sea Δ un conjunto abierto en E, que incluye a un disco cerrado U_r de radio $r>0$, centrado en el orígen de E. Sea $f: \Delta \longrightarrow E$ una función diferenciable de clase C^m, $m\geq 2$, tal que $f(0)=0$ y $Df(0)$ es un isomorfismo. Sea $\delta^*<\delta$ donde δ está definido en (4.14). Sea $L = 4\,\gamma\, M_0 \| (Df(0))^{-1} \|^2$ y sea $D_2 = 8\,\gamma^2 \| (Df(0))^{-1} \|^3 M_0$. Consideremos el método unipaso:*

$$Y_{n+1} - Y_n = h \left[Df(Y_n)\right]^{-1}(v) \ ; \qquad Y_0 = 0 \ ; \qquad 0\leq n\leq N-1 \qquad (4.20)$$

donde $\gamma>0$, $h>0$, $v\in E$, $\| v \| < \dfrac{\gamma\delta^*}{2}$ y $N = [\delta^*/h] = 2p$ es un entero par.

Definimos la función inversa aproximada:

$$\hat{g}(\ .\ ,\ h\ ,\ \gamma\ ):\quad U_{\frac{\gamma\delta^*}{2}} \longrightarrow U_{r'}$$

mediante la expresión:

$$\hat{g}(v,h,\gamma) = Y_{\frac{N}{2}}\left(\frac{2v}{\delta^*}\right) \tag{4.21}$$

donde r' *viene dado por el Lema 4.1.*

El error de $\hat{g}$ *respecto de la función inversa teórica* f^{-1} *de* f *viene acotado superiormente por la desigualdad:*

$$\| f^{-1}(v) - \hat{g}(v,h,\gamma) \| \le h\,\gamma \left\| \left(Df(0)\right)^{-1} \right\| K(\gamma) \exp\left[K(\gamma)\right]$$

$$\| v \| < \frac{\gamma\delta^*}{2} \tag{4.22}$$

donde:

$$K(\gamma) = 4\,\gamma\, M_0 \left\| \left(Df(0)\right)^{-1} \right\|^3 \left(1 + 2\, r'\, M_0\, \gamma \left\| \left(Df(0)\right)^{-1} \right\| \right) r' \tag{4.23}$$

Dado un error admisible $\varepsilon>0$, tomando:

$$h < \varepsilon \cdot \left[\gamma \, K(\gamma) \left\| \left(Df(0)\right)^{-1} \right\| \exp[K(\gamma)] \right]^{-1}$$

la correspondiente función inversa $\hat{g}(.,h,v)$ *satisface:*

$$\| f^{-1}(v) - \hat{g}(v,h,\gamma) \| < \varepsilon \, , \qquad v \in U_{\frac{\gamma\delta}{2}}^{*} \tag{4.24}$$

NOTA 4.1. Téngase en cuenta que por *Teorema 4.2*, dado un error admisible $\varepsilon>0$, la cota de error (4.22) es buena en todo el dominio de definición de la función inversa f^{-1}, y la función inversa aproximada $\hat{g}(.,h,\gamma)$ depende del parámetro γ. Así, el tamaño requerido para h y el dominio de la inversa cambian si cambia el parámetro γ. Este hecho se ilustra con el siguiente ejemplo.

EJEMPLO 4.2.. Consideremos la función $f\colon \mathbb{R}^{2x2} \longrightarrow \mathbb{R}^{2x2}$ definida por:

$$f(X) = X^2 + X \tag{4.25}$$

Por el teorema 8.14 de [DIEU, pág.148], es fácil ver que:

$$(Df(X)): \mathbb{R}^{2x2} \longrightarrow \mathbb{R}^{2x2}$$
$$(Df(X))(V) = XV + VX + V; \quad V \in \mathbb{R}^{2x2}; \quad X \in \mathbb{R}^{2x2} \tag{4.26}$$

Es claro que $f(0)=0$ y que $(Df(0))(V)=V$. Luego $Df(0)$ es un isomorfismo y se satisface la condición (4.6). Tomemos $r=1$, entonces según la notación del *Lema 4.1*, *Teorema 4.1* y *Teorema 4.2*, tenemos que $(Df(0))^{-1}(V) = V$ y además:

$$\left\| \left(Df(0)\right)^{-1} \right\| = 1 \; ; \quad M_0 = 3 \; ; \quad L = 4 \left\| \left(Df(0)\right)^{-1} \right\|^2 M_0 = 12\,\gamma$$

$$\delta = \frac{1}{3}(1+\gamma) \; ; \quad D_2 = 24\,\gamma^2 \; ; \quad K(\gamma) = 2\,\gamma\,(1+\gamma)$$

Nótese que para calcular la constante δ que aparece en (4.14), necesitamos la expresión de $\left(Df(0)\right)^{-1}$, que por *Lema 4.1* está bien definida para $\| X \| < \frac{1}{6}$. Por (3.33), si X, T son matrices de $\mathbb{R}^{2x2}$ y $\| X \| < \frac{1}{6}$, se tiene que:

$$(Df(X))(T) \;=\; TX + XT + T \;=\; (X+I)\,T \;+\; TX$$

$$\Big(Df(X)\Big)^{-1}\Big(\,(X+I)\,T + \;TX\,\Big) \;=\; T$$

y por linealidad:

$$(X+I)\,\Big(Df(X)\Big)^{-1}(T) \;+\; \Big(Df(X)\Big)^{-1}(T)\;X \;=\; T$$

Si llamamos $A = \Big(Df(X)\Big)^{-1}(T)$, se sigue que A satisface la *ecuación matricial de Sylvester:*

$$(X+I)A + AX \;=\; T \tag{4.27}$$

Por [JAME] y [BOCH], o [MÜLL], podemos escribir:

$$A = \left(Df(X)\right)^{-1}(T) =$$

$$= \left[\left(1+trX\right)T + TX - X \right] \left[\left(1+trX+|X|\right)I + \left(2+trX\right)X + X^2 \right]^{-1} \qquad (4.28)$$

donde trX representa la traza de X y $|X|$ su determinante. El método unipaso matricial (1.48) toma la forma:

$$X_{n+1} = h \sum_{j=0}^{n} \left(Df(X_j)\right)^{-1}(V) \; ; \quad \| V \| < \frac{\delta^* \gamma}{2} \qquad (4.29)$$

Tomando N=10, la tabla 4.1 muestra los resultados obtenidos para diferentes valores del parámetro γ, donde:

$$E(\gamma,h) = 2\ h\ (1+\gamma)\ \gamma^2 \exp \left[\ 2\ \gamma\ (1+\gamma)\ \right]$$

La aproximación a la función inversa para los diferentes valores de h viene dada por $\hat{g}(V,h,\gamma) = \left(\frac{2}{\delta}\right) X_5$, donde X_5 se calcula a través de la expresión (4.29).

Usando Mathematica, [WOLF], para $h = 0.04$ y:

$$V = \begin{bmatrix} 0.01 & 0.01 \\ 0 & 0.01 \end{bmatrix}$$

obtenemos el valor de la aproximación a la inversa en V:

$$\hat{g}(V,0.04,0.2) = \left(\frac{2}{0.5}\right) X_5 = 5\cdot 10^{-3} \begin{pmatrix} 1.99681 & 1.99363 \\ 0 & 1.99681 \end{pmatrix}$$

γ	δ	h	Cota de Error: $E(\gamma,h)$
0.1	0.366666	0.036666	$1.0051685\cdot 10^{-3}$
0.2	0.400000	0.040000	$6.205725\cdot 10^{-3}$
0.3	0.433333	0.043333	$2.212012\cdot 10^{-2}$
0.4	0.466666	0.046666	$6.407588\cdot 10^{-2}$
0.5	0.500000	0.050000	$1.680633\cdot 10^{-1}$

Tabla 4.1. Ejemplo 4.2.

CAPITULO V

ECUACIONES DIFERENCIALES EN DERIVADAS PARCIALES

V. A. Introducción

En este capítulo proponemos soluciones analítico-numéricas para problemas de valores iniciales y de contorno, relacionados con ecuaciones diferenciales en derivadas parciales. Dado un error admisible ε y un dominio finito G, la aproximación de la solución, construída en términos de los datos, será hecha de modo que el error está uniformemente acotado superiormente por ε en todo el dominio G.

Muchos sistemas físicos no pueden ser descritos por una única ecuación diferencial en derivadas parciales, pero sí como un sistema de dichas ecuaciones. Ejemplos de esto pueden encontrarse en difusión de calor [CANN], [HILL], [LEE], magnetohidrodinámica [BUTS], [SEZG], mecánica [MORS], modelos de armamento [GOPO], problemas de conducción nerviosa y neuronal [MAS1], [MAS2], etc...

Podemos encontrar métodos basados en la transformación de un sistema de ecuaciones diferenciales en derivadas parciales en un conjunto desacoplado de ecuaciones, [COUR], [ZACH], sin embargo, esta aproximación mediante desacoplamiento, presenta serios inconvenientes, ver [CHIN]

Nuestro objetivo es construir soluciones analítico-numéricas para sistemas de ecuaciones diferenciales en derivadas parciales de segundo órden, del tipo:

$$u_t(x,t)-Au_{xx}(x,t)-Bu(x,t)=G(x,t); \quad 0<x<p; \quad t>0 \tag{5.1}$$

$$u(0,t)=u(p,t)0=; \quad t>0 \tag{5.2}$$

$$u(x,0)=F(x); \quad 0\le x\le p \tag{5.3}$$

donde $u=(u_1,u_2,...,u_m)^T$, $F(x)$ y $G(x,t)$ son vectores n-dimensionales, elementos de $\mathbb{C}^m$, y A, B son matrices complejas de tamaño mxm, es decir, elementos de $\mathbb{C}^{mxm}$, tales que:

$$\text{Cada valor propio de } \left(\frac{A+A^H}{2}\right) \text{ es positivo,} \tag{5.4}$$

donde A^H representa la matríz conjugada traspuesta de A.

Nótese que en general, el sistema (5.1)-(5.3) no puede ser transformado en un conjunto de ecuaciones desacopladas si las matrices A y B no son simultáneamente diagonalizables. Para el caso en que B=0 y A es una matríz cuyos valores propios tienen parte real positiva, el anterior problema está estudiado en [JOD1].

El capítulo se organiza de la siguiente manera. En la sección siguiente tratamos sobre la construcción de una serie convergente, solución del problema (5.1)-(5.3), basándonos en un teorema referido a cotas para los valores propios del haz de matrices $B+\lambda A$. En la última sección, contestaremos la siguiente cuestión: dado $0<t_0<t_1$ y un error admisible $\varepsilon>0$, construir en términos de los datos del problema, una solución analítico-numérica de (5.1)-(5.3), cuyo error con respecto a la serie infinita solución del problema, sea menor que ε uniformemente en el dominio $D(t_0,t_1)=[0,p]\times[t_0,t_1]$.

Denotaremos, como habitualmente hemos hecho, mediante $\sigma(D)$ al conjunto de todos los valores propios de la matríz $D\in\mathbb{C}^{m\times m}$, y por $\rho(D)$ a su radio espectral que venía definido por el máximo del conjunto:

$$\left\{|z|;\ z\in\sigma(D)\right\}$$

Utilizaremos la norma de Frobenius [GOLU, pág. 14] que representaremos por $\|D\|_F$ para distinguirla de la 2-norma que representábamos por $\|D\|$. Recordemos que por [GOLU, pág. 15] se tiene que:

$$\|D\| \le \|D\|_F \le m^{1/2}\|D\|$$

Finalmente, si $B \in \mathbb{C}^{mxm}$ existe una matríz unitaria $Q \in \mathbb{C}^{mxm}$ tal que:

$$Q^H B Q = D + N$$

donde D es una matríz diagonal y $N \in \mathbb{C}^{mxm}$ es estrictamente triangular superior. Esta es la llamada *descomposición de Schur* de B, y además, por [GOLU, pág. 192-193], se tiene que:

$$\| N \|_F \leq \| B \|_F$$

Si denotamos por:

$$\alpha(B) = \max \left\{ \mathrm{Re}(z);\ z \in \sigma(B) \right\} \ ; \quad M_s(t) = \sum_{k=0}^{m-1} \frac{\| N\, t \|^k}{k!} \qquad (5.5)$$

entonces, por [GOLU, p.396] se tiene que:

$$\| e^{tB} \| \leq e^{t\alpha(B)} \ M_s(t) \qquad (5.6)$$

V. B. Solución Como Serie Infinita

Basándonos en el método de funciones propias del caso escalar, escogemos una serie como candidata a solución de (5.1)-(5.3) de la forma:

$$u(x,t)=\sum_{n\geq 1} T_n(t)\ \sin\left(\frac{n\pi x}{p}\right) \tag{5.7}$$

donde $T_n(t)$ es una función vectorial a determinar, que toma valores en $\mathbb{C}^m$. Supongamos que $T_n(t)$ verifica la ecuación diferencial:

$$T'_n(t) - \left[B-\left(\frac{n\pi}{p}\right)^2 A\right] T_n(t) = B_n(t)\ ;\quad t<0\ ;$$

$$B_n(t)=\frac{2}{p}\int_0^p G(x,t)\sin\left(\frac{n\pi x}{p}\right)dx \tag{5.8}$$

donde $B_n(t)$ es la transformada seno finita de la función $G(.,t)$, para un valor fijado $t>0$. La solución de (5.8) toma la forma [FLET, pág. 220]:

$$T_n(t)=\exp\left[B-(n\pi/p)^2A\right] C_n+\int_0^t \exp\left[(t-s)\left(B-(n\pi/p)^2A\right)\right]B_n(s)\ ds \qquad (5.9)$$

donde C_n es un vector arbitrario de $\mathbb{C}^m$. Para satisfacer la condición inicial (5.3), tomemos:

$$C_n = \frac{2}{p}\int_0^p F(x)\ \sin\left(\frac{n\pi x}{p}\right) dx\ ;\quad n\geq 1 \qquad (5.10)$$

Así, la posible solución del problema (5.1)-(5.3), puede ser escrita en la forma:

$$u(x,t)=\sum_{n\geq 1} T_n(t)\sin\left(\frac{n\pi x}{p}\right) = \sum_{n\geq 1} U_n(x,t) + \sum_{n\geq 1} V_n(x,t) \qquad (5.11)$$

donde:

$$V_n(x,t)=\left\{\int_0^t \exp\left[(t-s)\left(B-(n\pi/p)^2A\right)\right]B_n(s)\ ds\right\}\sin\left(\frac{n\pi x}{p}\right);\ n\geq 1 \quad (5.12)$$

$$U_n(x,t) = \exp\left[t\left(B-(n\pi/p)^2 A\right)\right] C_n \sin\left(\frac{n\pi x}{p}\right); \; n\geq 1 \tag{5.13}$$

Asumamos de momento que la solución en forma de serie u(x,t) definida por (5.7)-(5.8) es convergente, y sus términos son dos veces parcialmente diferenciables respecto a x, y una vez respecto a t, y además:

$$G(x,t)=\sum_{n\geq 1} B_n(t) \sin\left(\frac{n\pi x}{p}\right) \tag{5.14}$$

entonces, se tiene que:

$$u_t(x,t)-Au_{xx}(x,t)-Bu(x,t)-G(x,t)=$$

$$=\sum_{n\geq 1}\left\{T'_n(t)-\left[B-\left(\frac{n\pi}{p}\right)^2 A\right]T_n(t)-B_n(t)\right\} \sin\left(\frac{n\pi x}{p}\right) = 0$$

$$u(0,t)=u(p,t)=0$$

y:

$$u(x,0) = \sum_{n\geq 1} T_n(0) \sin\left(\frac{n\pi x}{p}\right) =\sum_{n\geq 1} C_n \sin\left(\frac{n\pi x}{p}\right) \tag{5.15}$$

Teniendo en cuenta (5.10) y (5.15), si cada componente f_j de $F=(f_1,f_2,\ldots,f_m)^T$ satisface una condición suficiente que garantice la convergencia de la serie seno de Fourier de $f_j(x)$ en $x\in[0,p]$, se obtiene:

$$u(x,0) = \sum_{n\geq 1}\left\{ \frac{2}{p}\int_0^p F(x)\, \sin\left(\frac{n\pi x}{p}\right)\, dx\right\} \sin\left(\frac{n\pi x}{p}\right) = F(x) \qquad (5.16)$$

Para poder probar las anteriores propiedades sobre convergencia, es necesario encontrar una cota para los valores propios de las matrices $B-(n\pi/p)^2A$, con $n\geq 1$. El siguiente *Teorema* está relacionado con esta cuestión:

TEOREMA 5.1. *Sean* B, A *matrices complejas de* $\mathbb{C}^{m\times m}$ *satisfaciendo la condición* (5.4) *y sea* $\gamma(A)$ *definida mediante:*

$$\gamma(A)=\min\left\{z;\ z\in\sigma\left(\frac{A+A^H}{2}\right)\right\} \qquad (5.17)$$

Entonces se tiene que:

$$\mathrm{Re}(z) < \rho\left(\frac{B+B^H}{2}\right) - \left(\frac{n\pi}{p}\right)^2 \gamma(A); \quad z\in\sigma\left[B-(n\pi/p)^2 A\right]; \quad n\geq 1 \qquad (5.18)$$

DEMOSTRACIÓN. Consideremos la descomposición:

$$A=H_1+iH_2; \quad B=S_1+iS_2$$

donde:

$$H_1=\frac{A+A^H}{2} \;; \quad S_1=\frac{B+B^H}{2} \;; \quad H_2=\frac{A-A^H}{2i} \;; \quad S_2=\frac{B+B^H}{2i}$$

y sea:

$$B-(n\pi/p)^2 A = S_1-(n\pi/p)^2 H_1 + i\left[S_2-(n\pi/p)^2 H_2\right]$$

Puesto que S_1 y $-(n\pi/p)^2 H_1$ son matrices hermíticas, por [BAUE, pág. 246] se tiene que:

$$\sigma\left[S_1-(n\pi/p)^2 H_1\right] \subset \bigcup_{i=1}^{m} G_i \qquad (5.19)$$

donde:

$$G_i = \left\{ z\in\mathbb{C};\ \left| z+(n\pi/p)^2\lambda_i(H_1) \right| \le \rho(S_1) \right\} \tag{5.20}$$

y:

$$\sigma(H_1)=\left\{ \lambda_i(H_1);\ 1\le i\le n \right\}$$

Ahora, aplicando el *Teorema de Bendixon* [STOE, pág.359] a la matríz $B-(n\pi/p)^2A$, se tiene que:

$$\mathrm{Re}(z) \le z_{max}\left(S_1-(n\pi/p)^2H_1 \right);\quad z\in\sigma\left(B-(n\pi/p)^2A \right);\quad n\ge 1 \tag{5.21}$$

donde $z_{max}\left(S_1-(n\pi/p)^2H_1 \right)$ representa el máximo valor propio de la matríz hermítica $S_1 - (n\pi/p)^2 H$. Por (5.19)-(5.21), la desigualdad (5.18) queda demostrada. □

Nótese que las expresiones para las derivadas parciales de las funciones vectoriales $U_n(x,t)$, $V_n(x,t)$ definidas en (5.12)-(5.13), toman la forma:

$$\frac{\partial V_n(x,t)}{\partial x} = \frac{n\pi}{p}\left\{\int_0^t \exp\left[(t-s)\left(B-(n\pi/p)^2A\right)\right]B_n(s)ds\right\}\cos\left(\frac{n\pi x}{p}\right) \qquad (5.22.a)$$

$$\frac{\partial^2 V_n(x,t)}{\partial x^2} =-\left(\frac{n\pi}{p}\right)^2 V_n(x,t) \qquad (5.22.b)$$

$$\frac{\partial V_n(x,t)}{\partial t} = B_n(t)\sin\left(\frac{n\pi x}{p}\right)+ \left(B-(n\pi/p)^2A\right)V_n(x,t) \qquad (5.22.c)$$

$$\frac{\partial U_n(x,t)}{\partial x} = \frac{n\pi}{p}U_n(x,t) \qquad (5.23.a)$$

$$\frac{\partial^2 U_n(x,t)}{\partial x^2} =-\left(\frac{n\pi}{p}\right)^2 U_n(x,t) \qquad (5.23.b)$$

$$\frac{\partial U_n(x,t)}{\partial t} = \left(B-(n\pi/p)^2A\right)U_n(x,t) \qquad (5.23.c)$$

Para probar la convergencia de las series (5.11)-(5.12), en tanto convergen los correspondientes términos de las derivadas parciales de $U_n(x,t)$, $V_n(x,t)$, usaremos un argumento local basado en el siguiente resultado:

TEOREMA 5.2. *Sean* A, B *matrices de* $\mathbb{C}^{m\times m}$ *con* $m\geq 2$, *satisfaciendo* (5.4) *y supongamos que para cada* $t>0$, $G(.,t)$ *es una función contínua de la variable* $x\in[0,p]$, *admitiendo derivadas parciales contínuas hasta, al menos, segundo órden con respecto a la variable* x , *y:*

$$\text{(i)}\quad G(0,t) = G(p,t) = 0\ ;\qquad \frac{\partial G(0,t)}{\partial x} = \frac{\partial G(p,t)}{\partial x} = 0$$

$$\text{(ii)}\ G(x,t);\ \frac{\partial G(0,t)}{\partial x};\ \frac{\partial^2 G(p,t)}{\partial x^2}\quad \textit{son acotados en}\quad [0,p]\times\,]0,\infty[$$

Sean $U_n(x,t)$, $V_n(x,t)$ *y* $B_n(t)$ *definidos por* (5.13), (5.12) *y* (5.8) *respectivamente, y sea* C_n *definida en* (5.10) *donde* $F(x)$ *es una función acotada, localmente integrable en* $[0,p]$. *Si* τ, θ *y* M *son las constantes positivas definidas como:*

$$\tau = \| B \|_F \;; \qquad \theta = \left(\frac{\pi}{p}\right)^2 \| A \|_F \;;$$

$$M = \sup\left\{ \left\| \frac{\partial^2 G(x,t)}{\partial x^2} \right\| \;;\; 0 \le x \le p;\; t>0 \right\} \tag{5.24}$$

entonces, para $(x,t) \in D(t_0,t_1) = [0,p]\times[t_0,t_1]$, *con* $t_1>t_0>0$, *se tiene que:*

(iii) $$\| B_n(t) \| \le 2M \left(\frac{p}{n\pi}\right)^2 ; \quad n\ge 1$$

(iv) $$\| U_n(x,t) \| \le \| C_n \| \exp\left[-(n\pi/p)^2 \gamma(A) t_0\right] \cdot$$

$$\cdot \left[\sum_{k=0}^{m-1} (t_1)^k \frac{(\tau+n^2\theta)^k}{k!} \right] \cdot \exp\left[t_1 \rho\left(\frac{B+B^H}{2} \right) \right]$$

(v) $$\| V_n(t) \| \le \frac{1}{n^4} \exp\left[-(n\pi/p)^2 \gamma(A) t_0\right] H(n)$$

donde H(n) *es la función racional en la variable* n, *definida por:*

$$H(n)=\left\{1+\sum_{k=0}^{m-1}\left[t_1^2(\tau+n^2\theta)\right]^k\left(\sum_{j=1}^{k}\left[(n\pi/p)^2\gamma(A)t_1\right]^{-j}\right)\right\}\frac{2Mp^4\exp\left[t_1\rho\left(\frac{B+B^H}{2}\right)\right]}{\pi^4\gamma(A)}$$

cuyo mayor exponente es n^{2m-4} *con coeficiente:*

$$\gamma=\frac{2(t_1)^{2m-3}\;p^6\;M\exp\left[t_1\rho\left(\frac{B+B^H}{2}\right)\right]}{\left(\gamma(A)\right)^2\;\pi^6}$$

DEMOSTRACIÓN. Por las hipótesis (i), (ii) y [ZYGM, pág. 70], se obtiene (iii). Consideremos *la descomposición de Schur* para la matríz $B-(n\pi/p)^2A$:

$$Q_N^H\left[B-(n\pi/p)^2A\right]Q_N = D_N + N_N$$

con:

$$\| N_n \| \le \| N_n \|_F \le \| B \|_F + (n\pi/p)^2 \| A \|_F = \tau + n^2\theta$$

$$\tau = \| B \|_F \; ; \quad \theta = (\pi/p)^2 \| A \|_F$$

Si $\alpha_n(A,B)=\max\left\{\mathrm{Re}(z);\ z\in\sigma\left[B-(n\pi/p)^2A\right]\right\}$, por *Teorema 5.1* tenemos:

$$\alpha_n(A,B) \le \left[\rho\left(\frac{B+B^H}{2}\right)\right] - \left[(n\pi/p)^2\gamma(A)\right];\ n\ge 1 \qquad (5.25)$$

Por (5.5), (5.6) y (5.25), para $(x,t)\in D(t_0,t_1)$, se tiene:

$$\left\|\exp\left[t\left(B-(n\pi/p)^2A\right)\right]\right\| \le \exp\left[t\left(\rho\left(\frac{B+B^H}{2}\right)-(n\pi/p)^2\gamma(A)\right)\right]\sum_{k=0}^{m-1}(t_1)^k\frac{\| N_n\|^k}{k!}$$

$$\le \exp\left[t_1\rho\left(\frac{B+B^H}{2}\right)\right]\exp\left[-(n\pi/p)^2\ t_0\ \gamma(A)\right]\sum_{k=0}^{m-1}(t_1)^k\ \frac{(\tau+n^2\theta)^k}{k!} \qquad (5.26)$$

Por (5.26) se obtiene (iv).

Sea $0 \le s \le t \le t_1$, entonces por (5.5), (5.6) y (5.25), se tiene que:

$$\left\| \exp \left[(t-s) \left(B-(n\pi/p)^2 A \right) \right] \right\| \le$$

$$\le \exp\left[t_1 \rho\left(\frac{B+B^H}{2} \right) \right] \exp\left[-(n\pi/p)^2 \, (t-s) \, \gamma(A) \right] \sum_{k=0}^{m-1} (t-s)^k \, \frac{(\tau+n^2\theta)^k}{k!} \qquad (5.27)$$

Por (iii), (5.13), (5.27) y usando la fórmula [GRAD, pág. 92]:

$$\int x^n \, e^{ax} \, dx = e^{ax} \left\{ \frac{x^n}{a} + \sum_{j=1}^{n} (-1)^j \, \frac{n(n-1)\dots(n-j+1)}{a^{j+1}} \, x^{n-j} \right\}$$

podemos escribir:

$$\| V_n(x,t) \| \le \int_0^t \left\| \exp \left[(t-s) \left(B-(n\pi/p)^2 A \right) \right] \right\| \, \| B_n(s) \| \, ds \le$$

$$\le 2M \, (p/n\pi)^2 \exp\left[t_1 \rho\left(\frac{B+B^H}{2} \right) \right].$$

$$\cdot \int_0^t \exp\left[-(n\pi/p)^2\ (t-s)\ \gamma(A)\right] \sum_{k=0}^{m-1} (t-s)^k \frac{(\tau+n^2\theta)^k}{k!} ds \le$$

$$\le 2\ M\ (p/n\pi)^2 \exp\left[t_1 \rho\left(\frac{B+B^H}{2}\right)\right] \left\{ \sum_{k=0}^{m-1} \frac{(\tau+n^2\theta)^k}{k!} \cdot \right.$$

$$\left. \cdot \int_0^t \exp\left[-(n\pi/p)^2\ (t-s)\ \gamma(A)\right]\ (t-s)^k\ ds + \frac{1}{\gamma(A)\ (n\pi/p)^2} \right\} \le$$

$$\le 2M(p/n\pi)^2 \exp\left[-(n\pi/p)^2\ (t_0)\ \gamma(A)\right] \left\{ \sum_{k=0}^{m-1} \frac{(\tau+n^2\theta)^k}{k!} \cdot \right.$$

$$\left. \cdot \sum_{j=1}^{k} \frac{k(k-1)\ldots(k-j+1)}{\left(\gamma(A)(n\pi/p)^2\right)^{j+1}}\ t_1^{2k-j} + \frac{1}{\gamma(A)\ (n\pi/p)^2} \right\} \exp\left[t_1 \rho\left(\frac{B+B^H}{2}\right)\right] =$$

$$= \frac{2Mp^4}{n^4 \gamma(A) \pi^4} \exp\left[-\left(\frac{n\pi}{p}\right)^2 t_0\ \gamma(A)\right] \left\{ 1 + \sum_{k=0}^{m-1} \frac{(\tau+n^2\theta)^k}{k!} \cdot \right.$$

$$\cdot \sum_{j=1}^{k} \frac{\left(t_1(\tau+n^2\theta)\right)^k}{\left(\gamma(A)(n\pi/p)^2 t_1\right)^j} \Bigg\} \exp\left[t_1 \rho\left(\frac{B+B^H}{2}\right)\right]$$

Con lo que queda demostrado el teorema. □

COROLARIO 5.1. *Supongamos que* A *es una matríz de* $\mathbb{C}^{m\times m}$ *satisfaciendo* (5.4)*, Sea* $B\in\mathbb{C}^{m\times m}$ *y sea* G(x,t) *una función que satisface las hipótesis del Teorema 5.2. Sea* F(x) *una función contínua en* [0,p] *tal que* F(0)=F(p)=0 *y cada una de sus componentes* f_j *para* $1\le j\le m$ *satisface alguna de las condiciones:*

(a) $f_j(x)$ *es una variación local acotada en cualquier punto* $x\in[0,p]$.

(b) $f_j(x)$ *admite derivadas laterales* $\left(f'_j\right)_R(x)$ *y* $\left(f'_j\right)_L(x)$ *en cualquier punto* $x\in[0,p]$.

Entonces u(x,t) *definida en* (5.11) *es una solución del problema* (5.1)-(5.3).

DEMOSTRACIÓN. Por las hipótesis impuestas a F(x) y por (5.4), la serie:

$$U_n(x,0) = \sum_{n\geq 1} C_n \sin\left(\frac{n\pi x}{p}\right)$$

converge a F(x) para cada $x\in[0,p]$, ver el Corolario 1 de [CHUR, pág. 57]. Por otro lado, si $t_1>t_0>0$ y $D(t_0,t_1) = [0,p]\times[t_0,t_1]$ por *Teorema 5.2*, la serie (5.11) y cada una de las series cuyo término general sea uno de los que aparecen en (5.22)-(5.23), son uniformementes convergentes en $D(t_o,t_1)$. Por el *teorema de derivación para series funcionales* [APOS, Teorema 9.14], u(x,t) definido en (5.11) es dos veces parcialmente diferenciable en sus términos respecto a x, y verifica (5.1)-(5.3) en cualquier punto (x,t_0), con $t_0>0$, $0\leq x\leq p$. □

V.C. Solución Analítico-Numérica y Cota de Error.

La solución en serie $u(x,t)$ que proporciona el *Corolario 5.1* presenta dos dificultades para su cálculo. En primer lugar, porque $u(x,t)$ es una serie infinita, y después porque el término general incluye exponenciales de matrices cuyo cálculo no es una tarea fácil [MOLE]. En esta sección pretendemos superar estos inconvenientes en dos etapas. Primero, truncando la serie infinita en una suma parcial y entonces, cada exponencial de matrices, sustituirla por una suma finita apropiada de su desarrollo en *serie de Taylor.*

TEOREMA 5.3. *Sea* N_1 *una constante positiva tal que:*

$$\| F(x) \| \le N_1$$

para cada $x \in [0,p]$. *Sean* τ *y* θ *definidas en* (5.24), *sea* γ *definida en Corolario 5.1 y sea* $D(t_0,t_1)=[0,p]x[t_0,t_1]$, $\psi =(\pi/p)^2 \gamma(A)\, t_0$

(i) *Sea* $H(n)$ *definida en Teorema 5.2 y sea* n_1 *el primer entero positivo* n *que satisface las desigualdades:*

$$H(n)\le 2\gamma n^{2m-4} \; ; \quad \frac{\ln(n)}{n(n\psi-1)} < \frac{1}{2m-8} \; ; \quad n(2n\psi-1) > 2m-8 \; ; \; n\ge n_1 \qquad (5.28)$$

si $V_n(x,t)$ *viene definida por* (5.12) *y* $(x,t)\in D(t_0,t_1)$, *entonces se tiene que:*

$$\sum_{n>n_1} \left\| V_n(x,t) \right\| \le 2\,\gamma\,(1-e^{-1})^{-1}\, e^{-n_1} \qquad (5.29)$$

(ii) *Sea* $0\le k\le m-1$ *y sea* $t_k>0$ *tal que:*

$$2\,k\,\ln\left(1+t_k^2\theta\tau^{-1}\right) < t_k\psi \quad \text{and} \quad \frac{1}{2}\,t_k\left(t_k\psi-2\right) > k\,\ln(\tau) \qquad (5.30)$$

Si $n_2\ge\max(t_0,t_1,\ldots,t_{m-1})$ *y satisface:*

$$\frac{\theta(m-1)}{\tau+n_2^2\theta} + \frac{1}{2n^2} < \psi \qquad (5.31)$$

y si $U_n(x,t)$ *está definido por* (5.13) *y* $(x,t)\in D(t_0,t_1)$, *entonces se tiene que:*

$$\sum_{n>n_2} \left\| U_n(x,t) \right\| \le$$

$$\leq 2N_1(1-e^{-1})^{-1}\exp\left[t_1\rho\left(\frac{B+B^H}{2}\right)\right]\left(\sum_{k=0}^{m-1}\frac{t_1^k}{k!}\right)e^{-n_2} \tag{5.32}$$

DEMOSTRACIÓN. (i) Puesto que $\lim_{n\to\infty}\frac{\ln(n)}{n(n\psi-1)} = 0$ y $\lim_{n\to\infty}\frac{H(n)}{n^{2m-4}} = \gamma$ la existencia del primer entero positivo que satisface (5.28) está garantizada. Nótese que para $n\geq n_1$ se tiene que:

$$\exp\left[-(n\pi/p)^2\gamma(A)t_0\right]\frac{H(n)}{n^4}\leq 2\gamma\exp\left[-n^2\psi+(2m-8)\ln(n)\right]=2\gamma\exp\left(h(n)\right) \tag{5.33}$$

donde $h(n)=(2m-8)\ln(n)-\psi n^2$, y nótese que la sucesión $\left\{h(n)+n\right\}$ es decreciente y negativa para $n\geq n_1$. En efecto, sea:

$$h(t) = (2m-8)\ln(t) - \psi\, t^2$$

y nótese que $h(t)+t<0$, si y sólo si:

$$(2m-8)\ \ln(t) - \psi\, t^2 + t < 0\ ; \qquad \ln(t) < \frac{\psi\, t^2 - t}{2m - 8}$$

y $h'(t)+1<0$, si y sólo si:

$$\frac{2m - 8}{t} - 2t\psi + 1 < 0\ ; \qquad 2m-8 < t\ (2t\psi - 1)$$

Así, para $n\geq n_1$ se obtiene $h(n+1)+n+1 < h(n)+n < 0$ y por (5.33) podemos escribir:

$$\sum_{n\geq n_1} \exp\left[-(n\pi/p)^2 t_0 \gamma(A)\right] \frac{H(n)}{n^4} \leq 2\gamma \sum_{n\geq n_1} \exp\Big(h(n)\Big) \leq$$

$$\leq 2\gamma \sum_{n\geq n_1} e^{-n} = \frac{2\gamma e^{-n_1}}{1-e^{-1}} \tag{5.34}$$

Por (5.34) y *Teorema 4.2-(v)* la parte (i) queda demostrada.

(ii) Sea $h_k(t) = k\ \ln(\tau+t^2\theta) - \psi\ t^2$ y nótese que $h_k(t)+t$ es decreciente y negativa si $h_k'(t)+1 < 0$ y $h_k(t) < -t$.

Podemos escribir:

$$\exp\left[-(n\pi/p)^2 t_0 \gamma(A)\right](\tau+n^2\theta)^k = e^{-n^2\psi}\ e^{k\ln(\tau+n^2\theta)} = \exp\left(h_k(n)\right) \qquad (5.35)$$

Estamos interesados en determinar el primer entero positivo n_2 tal que $\left\{h_k(n)+n\right\}$ es decreciente y negativa para $n\geq n_2$, i.e.:

$$h_k(n+1)+n+1 < h_k(n)+n < 0 \quad \text{for } n\geq n_2 \quad \text{and } 0\leq k\leq m-1 \qquad (5.36)$$

Nótese que $h_k(t)=k\ln(\tau)+k\ln\left(1+t^2\theta\tau^{-1}\right)-t^2\psi$ y como:

$$\lim_{t\to\infty} \frac{\ln\left(1+t^2\theta\tau^{-1}\right)}{t^2} = 0$$

podemos encontrar un número positivo t_k tal que:

$$2k\ln\left(1+t^2\theta\tau^{-1}\right) < \psi\ t^2 \quad \text{and} \quad \frac{1}{2}\left(t\psi-2\right) > k\ \ln(\tau) \quad \text{for} \quad t\geq t_k \qquad (5.37)$$

Las condiciones $h'_k(t)+1<0$ y $h_k(t)<0$ pueden escribirse en la forma:

$$\frac{2tk\theta}{\tau+t^2\theta} - 2t\psi < -1\ , \qquad \text{or} \qquad \frac{\theta k}{\tau+t^2\theta} + \frac{1}{2t} < \psi \qquad (5.38)$$

y:

$$k\ \ln\left(1+t^2\theta\tau^{-1}\right) + k\ \ln(\tau) < t^2\psi \qquad (5.39)$$

Nótese que si t_k satisface (5.30), se tiene que:

$$k\ \ln\left(1+t_k^2\theta\tau^{-1}\right) + k\ \ln(\tau) - t^2\psi \ < \ k\ \ln(\tau) + \frac{t_k\psi}{2} - t_k^2\psi \ < \ 0$$

Así, por (5.39), (5.37) y (5.30), obtenemos que:

$$h_k'(t)+1<0 \quad y \quad h_k(t)<0 \quad \text{para} \quad t\geq t_k$$

Nótese que si $n_2 \geq \max(t_0, t_1, \ldots, t_{m-1})$ y satisface (5.31), entonces n_2 satisface también (5.38) para $k=0,1,\ldots,m-1$. Como $\| C_n \| \leq 2N_1$ para $n\geq 1$, por (5.35) y *Teorema 2-(iv)*, si $(x,t)\in D(t_0,t_1)$ se tiene que:

$$\sum_{n\geq n_2} \left\| U_n(x,t) \right\| \leq$$

$$\leq 2N_1 \exp\left[t_1 \rho\left(\frac{B+B^H}{2}\right)\right] \left(\sum_{k=0}^{m-1} \frac{t_1^k}{k!}\right) \left(\sum_{n\geq n_2} \exp\left(h_k(n)\right)\right) \leq$$

$$\leq 2N_1 \exp\left[t_1 \rho\left(\frac{B+B^H}{2}\right)\right] \left(\sum_{k=0}^{m-1} \frac{t_1^k}{k!}\right) \frac{e^{-n_2}}{(1-e^{-1})}$$

Así, la parte (ii) del *Teorema 5.3* queda demostrada. □

COROLARIO 5.2. *Sea* $\varepsilon>0$, *y sean* n_1 *y* n_2 *enteros positivos* *definidos en Teorema 5.3, y* $n_3 \geq \max\left\{ n_1, n_2 \right\}$ *tal que:*

$$n_3 \geq \ln \left\{ \frac{2\left[N_1 \exp\left[t_1 r \dfrac{B+B^H}{2}\right]\left(\displaystyle\sum_{k=0}^{m-1} \frac{t_1^k}{k!}\right) + \gamma\right]}{\varepsilon\ (1-e^{-1})} \right\} \tag{5.40}$$

entonces:

$$u(x,t,n_3) = \sum_{n=1}^{n_3} \left[U_n(x,t) + V_n(x,t) \right] \tag{5.41}$$

es una solución aproximada del problema (5.1)-(5.3) *cuyo error respecto a la solución exacta* u(x,t) *dada en Corolario 5.1, está acota superiormente por* $\varepsilon/2$*, uniformemente para* $(x,t)\in D(t_0,t_1)$.

DEMOSTRACIÓN. Por *Teorema 5.2*, para $(x,t)\in D(t_0,t_1)$ se tiene que:

$$\sum_{n>n_3} \left[\| U_n(x,t) \| + \| V_n(x,t) \| \right] \leq$$

$$\leq e^{-n_3}\left[2N_1\left(1-e^{-1}\right)^{-1}\exp\left[t_1\rho\left(\frac{B+B^H}{2}\right)\right]\left(\sum_{k=0}^{m-1}\frac{t_1^k}{k!}\right)+2\gamma\left(1-e^{-1}\right)\right]$$

Con lo que el resultado queda establecido. □

Ahora, sustituyamos cada exponencial de matrices de (5.41) por su serie de Taylor truncada de grado q, obteniendo la aproximación finita:

$$u(x,t,n_3,q)=\sum_{n=1}^{n_3}\sum_{j=0}^{q}\frac{t^j[B-(n\pi/p)^2A]^j}{j!}\,C_n\sin(n\pi x/p)+$$

$$+\sum_{n=1}^{n_3}\left[\int_0^t\left\{\sum_{j=0}^{q}\frac{(t-s)^j[B-(n\pi/p)^2A]^j}{j!}\right\}B_n(s)\,ds\right]\sin(n\pi x/p)=$$

$$=\sum_{n=1}^{n_3}\sum_{j=0}^{q}\frac{[B-(n\pi/p)^2A]^j}{j!}\left\{C_nt^j+\frac{2}{p}\int_0^p\int_0^t G(u,s)\left(t-s\right)^j\sin\left(\frac{n\pi u}{p}\right)duds\right\}\sin\left(\frac{n\pi x}{p}\right)$$

(5.42)

Ahora, determinemos el entero positivo q, tal que la diferencia entre $u(x,t,n_3)$ dada por *Corolario 5.2* y $u(x,t,n_3,q)$ verifique:

$$\| u(x,t,n_3) - u(x,t,n_3,q) \| \le \varepsilon \;; \quad \forall (x,t)\in D(t_0,t_1) \tag{5.43}$$

Por Teorema 11.2.4 de [GOLU, pág. 390], si $(x,t)\in D(t_0,t_1)$ y q es un entero positivo, se tiene que:

$$\left\| \exp\left[t\left(B-(n\pi/p)^2 A \right) \right] - \sum_{j=0}^{q} \frac{t^j [B-(n\pi/p)^2 A]^j}{j!} \right\| \le$$

$$\le \frac{m}{(q+1)!} \exp\left[t_1 \rho\left(\frac{B+B^H}{2} \right) \right] \sum_{k=0}^{m-1} (t_1)^k \frac{(\tau+n^2\theta)^k}{k!} \tag{5.44}$$

Usando la desigualdad:

$$\| C_n \| \le 2\, N_1 \;; \quad \forall n\ge 1 \tag{5.45}$$

y por (5.44)-(5.45), podemos escribir:

$$\left\| \sum_{n=1}^{n_3} \left\{ U_n(x,t) - \sum_{j=0}^{q} \frac{t^j[B-(n\pi/p)^2A]^j}{j!} C_n \sin\left(\frac{n\pi x}{p}\right) \right\} \right\| \le$$

$$\le \frac{2N_1 m}{(q+1)!} \exp\left[t_1 \rho\left(\frac{B+B^H}{2}\right)\right] \sum_{n=1}^{n_3} \sum_{j=0}^{q} (t_1)^k \frac{(\tau+n^2\theta)^k}{k!} \tag{5.46}$$

De manera análoga, por (5.26), si $0 \le s \le t \le t_1$ tenemos:

$$\left\| \exp\left[(t-s)\left(B-(n\pi/p)^2A\right)\right] - \sum_{j=0}^{q} \frac{(t-s)^j \, [B-(n\pi/p)^2A]^j}{j!} \right\| \le$$

$$\le \frac{m}{(q+1)!} \exp\left[(t-s) \, \rho\left(\frac{B+B^H}{2}\right)\right] \sum_{k=0}^{m-1} (t-s)^k \frac{(\tau+n^2\theta)^k}{k!} \tag{5.47}$$

Teniendo en cuenta que para $t_0 \le t \le t_1$ se tiene:

$$\int_0^t (t-s)^k \exp\left[(t-s)\rho\left(\frac{B+B^H}{2}\right)\right] ds \le \frac{t_1^{k+1}}{k+1} \exp\left[t_1\rho\left(\frac{B+B^H}{2}\right)\right] \quad (5.48)$$

por (5.47), (5.48) y *Teorma 5.2-(iii)* obtenemos:

$$\left\| \int_0^t \left[\exp\left[(t-s)\left(B-(n\pi/p)^2A\right)\right] - \sum_{j=0}^{q} \frac{(t-s)^j \ [B-(n\pi/p)^2A]^j}{j!}\right] B_n(s)\ ds \right\|$$

$$\le \frac{mN}{(q+1)!} \exp\left[t_1\rho\left(\frac{B+B^H}{2}\right)\right] \sum_{k=0}^{m-1} (t_1)^{k+1} \frac{(\tau+n^2\theta)^2}{n^2\ (k+1)!} \quad (5.49)$$

$$\left\| \sum_{n=1}^{n_3} \left\{ \int_0^t \left[\exp\left[(t-s) \left(B-(n\pi/p)^2A\right) \right] - \right.\right.\right.$$

$$\left.\left.\left. - \sum_{j=0}^{q} \frac{(t-s)^j \ [B-(n\pi/p)^2A]^j}{j!} \right] B_n(s)\ ds \right\} \sin\left(\frac{n\pi x}{p}\right) \right\| \le$$

$$\leq \frac{mN}{(q+1)!} \exp\left[t_1 \rho\left(\frac{B+B^H}{2}\right)\right] \sum_{n=1}^{n_3} \sum_{k=0}^{m-1} (t_1)^{k+1} \frac{(\tau+n^2\theta)^k}{(k+1)!} \qquad (5.50)$$

Por (5.41), (5.42), (5.46) y (5.50), para $(x,t)\in D(t_0,t_1)$ se tiene que:

$$\| u(x,t,n_3) - u(x,t,n_3,q) \| \leq$$

$$\leq \frac{m(N+2N_1)}{(q+1)!} \exp\left[t_1 \rho\left(\frac{B+B^H}{2}\right)\right] \sum_{n=1}^{n_3} \sum_{k=0}^{m-1} \left[1+ \frac{t_1}{k+1}\right] t_1^k \frac{(\tau+n^2\theta)^k}{k!} \qquad (5.51)$$

Nótese que si q_0 es el primer entero positivo q que satisface:

$$(q+1)! \geq 2\left\{\varepsilon m(N+2N_1)\exp\left[t_1 \rho\left(\frac{B+B^H}{2}\right)\right] \sum_{n=1}^{n_3} \sum_{k=0}^{m-1} \left[1+ \frac{t_1}{k+1}\right] t_1^k \frac{(\tau+n^2\theta)^k}{k!}\right\}$$

(5.52)

entonces, por (5.51)-(5.52), para $(x,t)\in D(t_0,t_1)$, se tiene que:

$$\| u(x,t,n_3) - u(x,t,n_3,q) \| < \frac{\varepsilon}{2} \tag{5.53}$$

Por *Corolario 5.2* y los anteriores comentarios, el siguiente resultado queda establecido:

TEOREMA 5.4. *Con la anterior notación, y bajo las hipótesis del Teorema 5.2, Corolario 5.1, y Teorema 5.3, dado* $\varepsilon>0$ *y* $t_1>t_0>0$, *la función* $u(x,t,n_3,q_0)$ *definida por:*

$$\sum_{n=1}^{n_3} \sum_{j=0}^{q} \frac{(t-s)^j \, [B-(n\pi/p)^2 A]^j}{j!} \left\{ C_n \, t^j + \right.$$

$$\left. + \frac{2}{p} \int_0^p \int_0^t G(u,s) \left(t-s\right)^j \sin\left(\frac{n\pi u}{p}\right) du \, ds \right\} \sin\left(\frac{n\pi x}{p}\right)$$

es una solución aproximada del problema (5.1)-(5.3), *cuyo error respecto a la solución en serie dada en* (5.11), *es uniformemente acotado superiormente por* ε *para* $(x,t)\in D(t_0,t_1)=[0,p]\times[t_0,t_1]$.

NOTA 5.1. Nótese que $u(x,t,n_3,q_0)$ está expresada únicamente en términos de los datos del problema, y los enteros n_3 y q_0 son dependientes directamente del dominio $D(t_0,t_1)$. Para el caso en el que B es la matríz nula, los resultados propuestos aquí contienen a los dados en [JOD1], en el sentido de que la condición (5.4) implica que $Re(z)>0$ para cada valor propio de la matríz A, ver *Teorema de Bendixon*, [STOE, pág. 395]. Es más, la cota de error que aquí se ofrece, no requiere información espectral usada en [JOD1] en relación a la matríz A, y en particular, el conocimiento del índice de los valores propios de A así como las proyecciones espectrales asociadas a A.

REFERENCIAS

[APOS] APOSTOL, T. M. Mathematical Analysis, Addison-Wesley, Reading M. A. 1977.

[BARN] BARNETT, S. Matrices in Control Theory. Van Nostrand, Reinhold. London. 1971.

[BARR] BARRAUD, A. Y. Nouveaux Dévelopements sur la Résolution Numérique de $X' = A_1X + XA_2 + D$, $X(0)=C$. R.A.I.R.O. Vol. 16, No.4. 1982, pp. 341-356.

[BAUE] BAUER, F. L.; HOUSEHOLDER, A. S. Absolute Norms and Characteristic Roots, Numerische Math., 3 (1961), 241-246.

[BOCH] BOCHER, M. Introduction to Higher Algebra. New York, McMillan. 1947.

[BUTS] BUTSENICKS, I.E.; SHCHERBININ, E.V. Magnetohydrodynamics Flow in Electrodinamically Coupled Rectangular Ducts I, Magn. Girodinam. 2 (1976), 35-40.

[CANN] CANNON, J. R.; KLEIN, R. E. On the Observability and Stability of the Temperature Distribution in a Composite Heat Conductor, SIAM J. Appl. Math., 24 (1973), 596-602.

[COUR] COURANT, R.; HILBERT, D. Methods of Mathematical Physics, Vol. II, Interscience, New York, 1962.

[CHIN] CHIN, H. H. Theory of Irreducible Linear Systems, Quart.

Appl. Math., XXVIII (1969), 87-104.

[CHUR] CHURCHILL, R. V.; BROWN, J. W. Fourier Series and Boundary Value Problems, McGraw-Hill, New York, 1978.

[DAVI] DAVIDSON, E. J. The Numerical Solution of $X' = A_1X + XA_2 + D$, $X(0)=C$. IEEE Trans Aut. Control. August 1975, pp. 566-567.

[DIEC] DIECI, L. Numerical Integration of the Differential Riccati Equations and Some Related Issues. SIAM J. Numerical Analysis, 29. Vol.3 . 1992, pp. 781-815.

[DIEU] DIEUDONNE, J. Fundamentos de Análisis Moderno. Ed. Reverté. Barcelona. 1966.

[DUNF] DUNFORD, N.; SCHWARTZ, J. Linear Operators, part I. New York. Interscience. 1957.

[FLET] FLETT, T. M. Differential Analysis. Cambridge Univ. Press, New York. 1980.

[FREE] FREEDMAN, H. I. An Explicit Estimate of the Norm of the Inverse of a Matrix. SIAM Review 11 (2). 1969, pp. 254-256.

[GOH1] GOHGBERG, I.; LANCASTER, P.; RODMAN, L. Invariant Subspaces of Matrices with Applications. John Wiley, New York. 1986.

[GOH2] GOHGBERG, I.; LANCASTER, P.; RODMAN, L. Matrix Polynomials. Academic Press, New York. 1982.

[GOLU] GOLUB, G. H.; VAN LOAN, C. F. Matrix Computations. John Hopkins Univ. Press. Baltimore, Maryland. 1985.

[GOPO] GOPOLSAMY, K. An Arms Race with deteriorating armaments, Math. Biosci., 37 (1977), 191-203.

[GRAD] GRADSHTEYN I. S.; RIYZHIK, I. M. Table of Integrals, Series and Products, Academic Press, New York, 1980.

[GRAH] GRAHAM, A. Kronecker Products and Matrix Calculus with Applications. John Wiley, New York. 1981.

[HAMM] HAMMERLIN, G.; HOFFMAN, K. H. Numerical Mathematics. Readings in Maths. Springer-Verlag, New York. 1991.

[HENR] HENRICI, P. Discrete Variable Methods in Ordinary Differential Equations. John-Wiley, New York. 1962.

[HILL] HILL, J. M.; WILLIS, J. R. Analytic Aspects of Heat Flow in Random Two-component laminates, Quart. Appl. Math., 46 No.2, (1988), 353-364.

[JAME] JAMESON, A. Solution of the Equation AX+XB=C by Inversion of an MxM or an NxN Matrix. SIAM J. Appl. Maths. 16. 1968, pp. 1020-1023.

[JOD1] JÓDAR, L. Computing Accurate Solutions for Coupled Systems of Second Order Partial Differential Equations II, Int. J. Computer Math., 46 (1992), 63-75.

[JOD2] JÓDAR, L. Rational Approximate Solutions and Error Bounds for the Non-Symmetric Riccati Matrix Differential Equation. Appl. Maths. and Comput. 50. 1992, pp.23-31.

[JOD3] JÓDAR, L.; ABOU-KANDIL, H. A Resolution Method for Riccati Differential Systems Coupled in their Quadratic Terms. SIAM J. Math. Anal., 19. 1988, pp. 1225-1230.

[JOD4] JÓDAR, L.; MORERA, J. L.; NAVARRO, E. On Convergent Linear Multi-Step Matrix Methods. Int. J.Computer Maths. 40. 1991, pp. 211-219.

[JOD5] JÓDAR, L.; RUBIO, G.; PONSODA E. Matrix Methods for the Numerical Solution of Matrix Differential Equations. (Submit for publication).

[JOD6] JÓDAR, L.; PONSODA E. Continuous Numerical Solutions and Error Bounds for Matrix Differential Equations. Int. Proc. First Int. Coll. Num. Anal. D. Bainov and V. Covachev Ed. VSP. Utrech. The Netherlands, 1993. pp. 73-88.

[JOD7] JÓDAR, L.; PONSODA E. Analytic Approximate Solutions and Error Bounds for Linear Matrix Differential Equations

Appearing in Control. Contr. and Cyb. Vol 21 (1992) No. 3/4.

[KELL] KELLER, J. B. Numerical Solution of Two Point Boundary Value Problem. CBMS-NSF Reg.Conf.Series in Appl.Maths., SIAM, Pennsylvania. 1973.

[KENN] KENNEY, S. S.;LEIPNIK, R. B. Numerical Integration of the Differential Matrix Riccati Equation. IEEE Trans.Aut.Control AC-30. 1985, pp. 962-970.

[KUNK] KUNKEL, P.; MEHRMANN, V. Numerical Solution of Differential-Algebraic Riccati Equation. Linear Algebra Appl. 137/138. 1990, pp. 39-66.

[LAMB] LAMBERT, J. D. Computational Methods in Ordinary Differential Equations. John-Wiley, New York. 1972.

[LANC] LANCASTER, P.; TISMENETSKY, M. The Theory of Matrices. Second Ed. New York. Academic Press. 1985.

[LEE] LEE, A. I. ; HILL, J. M. On the General Linear Coupled System for Diffusion in Media with Two Diffusitives, J. Math. Anal. Appl., 89 (1982), 530-557.

[LEGU] LEGUA, M. Ecuaciones Matriciales Diferenciales, en Diferencia y en Derivadas Parciales, Tesis Doctoral, Departamento de Matemática Aplicada. Universidad Politécnica de Valencia, 1991.

[MAS1] MASCAGNI, M. An Critical-Boundary Value Problem of Physiological Significance for Equations of Nerve Conduction, Com. Pure Appl. Math., XLII (1989), 213-227.

[MAS2] MASCAGNI, M. The Backward Euler Method for Numerical Solution of the Hodgkin-Huxley Equations of Nerve Conduction, SIAM J. Numer. Anal., 27 No.4 (1990), 941-962.

[MARZ] MARZULLI, P.; GHERI, G. Estimation of the Global Discretization Error in Shooting Methods for Linear Boundary Value Problems. J.Comput.Appl.Maths., 28, 309. 1989, pp. 301-314.

[MOLE] MOLER C. B.; VAN LOAN, C. F. Nineteen Dubious Ways to Compute the Exponential of a Matrix, SIAM Review 20 (1978), 801-836.

[MORS] MORSE P. M.; FESHBACH, H. Methods of Theoretical Physics, McGraw-Hill, Tokyo, 1953.

[MÜLL] MÜLLER, P. Chr. Solution of the Matrix Equations AX+XB=-Q and $S^TX+XS=-Q$. SIAM J. Appl. Maths. 18. N.3. 1970, pp. 682-687.

[ORT1] ORTEGA, J. M. Numerical Analysis, A Second Course. Academic Press, New York. 1972.

[ORT2] ORTEGA, J. M.; RHEINBOLT, W. C. Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New

York. 1970.

[OSHM] OSHMAN, Y.; BAR-ITZHACK, I. Y. Eiegenfactor Solution of the Matrix Riccati Equation. A Continuous Square Root Algorithm. Proc. 23. Conf. on Decision and Control. Las Vegas. 1984, pp. 503-508.

[PORT] PORTER, W. A. On the Matrix Riccati Equations. IEEE Trans. Aut. Control. AC-12. 1967, pp.746-749.

[REI1] REID, W. T. A Matrix Differential Equations of Riccati Type. Amer. J. Math, 68. 1946, pp. 237-246.

[REI2] REID, W. T. Riccati Differential Equations. Academic Press, New York. 1972.

[REKT] REKTORYS, K. The Method of Discretization in Time and Partial Differential Equations. D.Reidel Pub.Co. 1982.

[SERB] SERBIN, S. M.; SERBIN, C. A. A Time-Stepping Procedure for $X' = A_1X + XA_2 + D$, $X(0)=C$. IEEE Trans Aut. Control, Vol. AC-25, No. 6. 1980, pp.1138-1141.

[SEZG] SEZGIN, M. Magnetohydrodynamic Flow in an Infinite Channel, Int. J. Nummer. Methods Fluids, 6 (1986), 593-609.

[SIMA] SIMAAN, M.; CRUZ, J. B. On the Solution of Open-Loop Nash Riccati Equations in Linear Quadratic Differential Games.

Int.J.Control. 18, 57. 1973.

[STOE] STOER, J.; BULIRSCH, R. Introduction to Numerical Analysis, Springer-Verlag, New York, 1980.

[THOM] THOMAS, V.; BARRAUD, A. Y. Commande Optimale à Horizon Fluyant. R.A.I.R.O.J-1. 1974, pp.126-140.

[VEMU] VEMURI, V.; KARPLUS, W. J. Digital Computer Treatment of Partial Differential Equations, Prentice Hall, Englewood Cliffs, New Jersey, 1981.

[VETT] VETTER, W. J. Derivative Operations on Matrices. IEEE Trans. Aut. Control AC-15. 1970, pp. 241-244.

[WOLF] WOLFRAM, S. Mathematica, System for Doing Mathematics by Computer. Redwood City. Addison Wesley Publishing Co. 1989.

[ZACH] ZACHMANOGLOU, E. C.; THOE, D. W. Introduction to Partial Differential Equations with Applications, William and Wilkins, Baltimore, 1976.

[ZYGM] ZYGMUND, A. Trigonometric Series, Second Ed., Vols. I and II, Cambridge Univ. Press, Cambridge, 1977.