Journal of Computational Mathematics, Vol.9, No.2, 1991, 163–169.

APPROXIMATE SOLUTIONS AND ERROR BOUNDS FOR SOLVING MATRIX DIFFERENTIAL EQUATIONS WITHOUT INCREASING THE DIMENSION OF THE PROBLEM*

Lucas Jódar
(*Department of Applied Mathematics, Polytechnical University of Valencia*)

Abstract

In this paper we present a method for solving initial value problems related to second order matrix differential equations. This method is based on the existence of a solution of a certain algebraic matrix equation related to the problem, and it avoids the increase of the dimension of the problem for its resolution. Approximate solutions, and their error bounds in terms of error bounds for the approximate solutions of the algebraic problem, are given.

§1. Introduction

Second order matrix differential equations of the type

$$X^{(2)}(t) + A_1X^{(1)}(t) + A_0X(t) = F(t); \quad X(0) = C_0, \quad X^{(1)}(0) = C_1 \tag{1.1}$$

where A_i, C_i, for $i = 0, 1$, and $F(t)$ are square complex matrices, elements of $\mathbb{C}_{p\times p}$, and F is continuous, appear in the theory of damped oscillatory systems and vibrational systems[6].

The standard method for solving equations of the type (1.1) is based on the consideration of the change $Y_1 = X; Y_2 = X^{(1)}$, and the equivalent first order extended linear system

$$\left(\frac{d}{dt}\right)\begin{bmatrix} Y_1(t) \\ Y_2(t) \end{bmatrix} = C_L \begin{bmatrix} Y_1(t) \\ Y_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ F(t) \end{bmatrix}; \quad C_L = \begin{bmatrix} 0 & I \\ -A_0 & -A_1 \end{bmatrix} \tag{1.2}$$

and the solution of problem (1.1) is given by

$$X(t) = [I, 0]\left\{ \exp(tC_L)\begin{bmatrix} C_0 \\ C_1 \end{bmatrix} + \int_0^t \exp((t-s)C_L)\begin{bmatrix} 0 \\ F(s) \end{bmatrix}\right\}ds. \tag{1.3}$$

The expression (1.3) for the solution of problem (1.1) has the inconvenience of the increase of the dimension of the problem and the aim of this paper is to present a method for solving (1.1), without increasing the dimension of the original problem, which provides approximate solutions and their error bounds in terms of data and a solution of the algebraic equation

$$X^2 + A_1X + A_0 = 0. \tag{1.4}$$

* Received March 30, 1988.

Explicit methods for solving equations of the type (1.4) may be found in [1, 4], and iterative methods for its resolution are given in [9–12, 14].

For the sake of clarity in the presentation of the paper we recall some concept and properties that will be used below. If A, B are matrices in $\mathbb{C}_{p\times p}$, we denote by $\|\ \|$ the operator norm that is defined by the expression

$$\|A\| = \sup_{\|x\|=1} \|Ax\|$$

where for a vector y in $\mathbb{C}_p$, the symbol $\|y\|$ means the Euclidean norm of y. From [3] and [8], it follows that

$$\|AB\| \leq \|A\|\ \|B\|$$

and

$$\|\exp(tA) - \exp(tB)\| \leq \exp(t\|B\|)(\exp(t\|A - B\|) - 1) \tag{1.5}$$

where t is a real number.

§2. Approximate Solutions and Error Bounds

We begin this section with a result that provides a sequence of approximations that converges to the unique solution of problem (1.1), without increasing the dimension of the problem and under the existence hypothesis of a solution of the algebraic equation (1.4).

Theorem 1. *Let X_0 be a solution of equation (1.4), and let $\{Z_n\}_{n\geq 1}$ be a sequence of matrices in $\mathbb{C}_{p\times p}$ that converges to X_0 in the operator norm, and let us suppose that F is a continuous function. The sequence of matrix functions $X_n(t)$, defined by*

$$X_n(t) = \exp(tZ_n)C_n(t) + \Big(\int_0^t \exp((t-s)Z_n)\exp(-s(Z_n + A_1))ds\Big) D_n(t), \tag{2.1}$$

$$C_n(t) = C_0 - \int_0^t \int_0^s \exp(-uZ_n)\exp((-u+s)(Z_n + A_1))F(s)duds,$$

$$D_n(t) = (C_1 - Z_nC_0) + \int_0^t \exp(s(Z_n + A_1))F(s)ds \tag{2.2}$$

where $n \geq 1$, is pointwise convergent to the unique solution of problem (1.1), given by

$$X(t) = \exp(tX_0)C(t) + \Big(\int_0^t \exp((t-s)X_0)\exp(-s(X_0 + A_1))ds\Big) D(t), \tag{2.3}$$

$$C(t) = C_0 - \int_0^t \int_0^s \exp(-uX_0)\exp((-u+s)(X_0 + A_1))F(s)duds,$$

$$D(t) = C_1 - X_0C_0 + \int_0^t \exp(s(X_0 + A_1))F(s)ds. \tag{2.4}$$

Proof. It is clear that $X_1(t) = \exp(tX_0)$ is a solution of the homogeneous operator differential equation

$$X^{(2)} + A_1X^{(1)} + A_0X = 0. \tag{2.5}$$

Let us suppose that we are looking for a solution of the type $X_2(t) = X_1(t)U(t)$, where $U(t)$ is a new unknown matrix function. Taking derivatives of $X_2(t)$, it follows that

$$\begin{aligned} X_2^{(1)}(t) &= X_1(t)(X_0U(t) + U^{(1)}(t)), \\ X_2^{(2)}(t) &= X_1(t)(X_0^2U(t) + 2X_0U^{(1)}(t) + U^{(2)}(t)). \end{aligned} \tag{2.6}$$

By imposing that $X_2(t)$ is a solution of (2.5), it follows that $U(t)$ must satisfy

$$\begin{aligned} X_2^{(2)} + A_1X_2^{(1)} + A_0X_2 &= X_1(t)U^{(2)}(t) + (X_0^2 + A_1X_0 + A_0)X_1(t)U(t) \\ &+ (2X_0 + A_1)X_1(t)U^{(1)}(t) = X_1(t)U^{(2)}(t) + (2X_0 + A_1)X_1(t)U^{(1)}(t) = 0. \end{aligned} \tag{2.7}$$

If we denote by $Z(t)$ the matrix function $Z(t) = X_1(t)U^{(1)}(t)$, it follows that (2.7) may be rewritten by the form

$$Z^{(1)}(t) + (X_0 + A_1)Z(t) = 0. \tag{2.8}$$

Solving (2.8), it follows that the solution of this equation that satisfies the initial condition $Z(0) = I$ is given by

$$Z(t) = \exp(-t(X_0 + A_1)).$$

Hence, $U^{(1)}(t) = (X_1(t))^{-1}Z(t) = \exp(-tX_0)\exp(-t(X_0 + A_1))$, and by imposing $U(0) = 0$, one gets that

$$U(t) = \int_0^t \exp(-sX_0)\exp(-s(X_0 + A_1))ds$$

and

$$X_2(t) = X_1(t)U(t) = \int_0^t \exp((t-s)X_0)\exp(-s(X_0 + A_1))ds \tag{2.9}$$

is a solution of (2.5), such that $X_2(0) = 0, X_2^{(1)}(0) = I$, because $X_2^{(1)}(t) = \exp(-t(X_0 + A_1)) + X_0\int_0^t \exp((t-s)X_0)\exp(-s(X_0 + A_1))ds$.

Let us suppose that in an analogous way to the scalar case we choose matrix functions $C(t), D(t)$, such that

$$W(t) = \begin{bmatrix} X_1(t) & X_2(t) \\ X_1^{(1)}(t) & X_2^{(1)}(t) \end{bmatrix}; \quad W(t)\begin{bmatrix} C^{(1)}(t) \\ D^{(1)}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ F(t) \end{bmatrix}. \tag{2.10}$$

If $C(t), D(t)$ satisfy (2.10), then the function

$$X(t) = X_1(t)C(t) + X_2(t)D(t) \tag{2.11}$$

satisfies

$$\begin{aligned} X^{(1)}(t) &= X_1^{(1)}(t)C(t) + X_2^{(1)}(t)D(t), \\ X^{(2)}(t) &= X_1^{(2)}(t)C(t) + X_2^{(2)}(t)D(t) + F(t) \end{aligned}$$

and thus

$$\begin{aligned} X^{(2)}(t) + A_1X^{(1)}(t) + A_0X(t) &= (X_1^{(2)} + A_1X_1^{(1)} + A_0X_1)C(t) \\ &+ (X_2^{(2)} + A_1X_2^{(1)} + A_0X_2)D(t) + F(t) = F(t). \end{aligned}$$

An easy computation yields that $V(t) = X_2^{(1)}(t) - X_1^{(1)}(t)(X_1(t))^{-1}X_2(t) = \exp(-t(X_0 + A_1))$ is invertible, and from lemma 1 of [5], it follows that $W(t)$ defined by (2.10) satisfies

$$(W(s))^{-1} = \begin{bmatrix} * & -\int_0^s \exp(-uX_0)\exp((-u+s)(X_0+A_1))du \\ * & \exp(s(X_0+A_1)) \end{bmatrix}. \tag{2.12}$$

From (2.10) and (2.12), it follows that

$$\begin{aligned} C(t) &= C(0) - \int_0^t \int_0^s \exp(-uX_0)\exp((-u+s)(X_0+A_1))F(s)duds, \\ D(t) &= D(0) + \int_0^t \exp(s(X_0+A_1))F(s)ds. \end{aligned} \tag{2.13}$$

By imposing that $X(t)$ given by (2.11) satisfies the initial conditions $X(0) = C_0$, $X^{(1)}(0) = C_1$, it follows that the matrices $C(0), D(0)$ must satisfy

$$\begin{aligned} X(0) &= C_0 = C(0), \\ X^{(1)}(0) &= C_1 = X_0C(0) + D(0). \end{aligned} \tag{2.14}$$

Solving (2.14), one gets that

$$\begin{bmatrix} C(0) \\ D(0) \end{bmatrix} = \begin{bmatrix} I & 0 \\ X_0 & I \end{bmatrix}^{-1} \begin{bmatrix} C_0 \\ C_1 \end{bmatrix} = \begin{bmatrix} C_0 \\ C_1 - X_0C_0 \end{bmatrix}.$$

Hence the result is established.

Now we are interested in finding an error bound for the approximate solution $X_n(t)$, given by (2.1)–(2.2), in terms of the approximation error $\|Z_n - X_0\|$, of the approximation Z_n of the solution X_0 of equation (1.4).

From (2.2) and (2.4), it follows that

$$\begin{aligned} C_n(t) - C(t) = \int_0^t \int_0^s &\{\exp(-uX_0)\exp((-u+s)(X_0+A_1)) \\ &- \exp(-uZ_n)\exp((-u+s)(Z_n+A_1))\}F(s)duds. \end{aligned} \tag{2.15}$$

Let b be a fixed positive number and let f be the constant defined by

$$f = \sup\{\|F(t)\|;\quad 0 \le t \le b\} \tag{2.16}$$

and let us consider the decomposition

$$\begin{aligned} &\exp(-uX_0)\exp((-u+s)(X_0+A_1)) - \exp(-uZ_n)\exp((-u+s)(Z_n+A_1)) \\ &\quad = \exp(-uX_0)\{\exp((-u+s)(X_0+A_1)) - \exp((-u+s)(Z_n+A_1))\} \\ &\quad +(\exp(-uX_0) - \exp(-uZ_n))\exp((-u+s)(Z_n+A_1)). \end{aligned} \tag{2.17}$$

If a is a positive number satisfying

$$\|Z_n\| \le a, \quad n \ge 1 \tag{2.18}$$

from (1.5) and (2.15)–(2.18), for $0 \le s \le t \le b$, it follows that

$$\|C_n(t) - C(t)\| \le 2ft^2\exp((2a + \|A_1\|)t)\exp((t\|Z_n - X_0\|) - 1). \tag{2.19}$$

From (2.2) and (2.4), it follows that

$$D_n(t) - D(t) = (X_0 - Z_n)C_0 + \int_0^t \{\exp(s(Z_n + A_1)) - \exp(s(X_0 + A_1))\}F(s)ds. \quad (2.20)$$

From (1.5), (2.18) and (2.20), it follows that

$$\|D_n(t) - D(t)\| \le \|C_0\| \, \|Z_n - X_0\| + tf\exp(t(a + \|A_1\|))(\exp(t\|Z_n - X_0\|) - 1). \quad (2.21)$$

From (2.1)–(2.4), one gets

$$\begin{aligned} X_n(t) - X(t) &= \exp(tZ_n)C_n(t) - \exp(tX_0)C(t) + \Big(\int_0^t \exp((t-s)Z_n)\exp(-s(Z_n \\ &\quad + A_1))ds\Big)D_n(t) - \Big(\int_0^t \exp((t-s)X_0)\exp(-s(X_0 + A_1))ds\Big)D(t) \\ &= \exp(tZ_n)(C_n(t) - C(t)) + (\exp(tZ_n) - \exp(tX_0))C(t) \\ &\quad + \Big(\int_0^t \exp((t-s)Z_n)\exp(-s(Z_n + A_1))ds\Big)(D_n(t) - D(t)) \\ &\quad + \Big(\int_0^t \{(\exp((t-s)Z_n)\exp(-s(Z_n + A_1)) \\ &\quad - \exp((t-s)X_0)\exp(-s(X_0 + A_1))\}ds\Big)D(t). \end{aligned} \quad (2.22)$$

From (2.2), (2.4), (2.16) and (2.18), it follows that $\|D_n(t)\|$ and $\|D(t)\|$ are bounded by

$$\|C_1\| + a\|C_0\| + tf\exp(t(a + \|A_1\|)).$$

From (2.4), (2.16), one gets

$$\|C(t)\| \le \|C_0\| + t^2 f\exp((2a + \|A_1\|)t). \quad (2.23)$$

Hence and from (1.5) it follows that

$$\begin{aligned} \|(\exp(tZ_n) - \exp(tX_0))C(t)\| &\le (\exp(t\|Z_n - X_0\|) - 1)\exp(at) \\ &\quad \times (\|C_0\| + t^2 f\exp((2a + \|A_1\|)t). \end{aligned}$$

From (2.18) and (2.19), we have

$$\|\exp(tZ_n)(C_n(t) - C(t))\| \le 2ft^2\exp((3a + \|A_1\|)t)(\exp(t\|Z_n - X_0\|) - 1), \quad (2.24)$$

$$\Big\|\int_0^t \exp((t-s)Z_n)\exp(-s(Z_n + A_1))ds\Big\| \le t\exp(t(2a + \|A_1\|)). \quad (2.25)$$

Taking into account the decomposition

$$\begin{aligned} &\exp((t-s)Z_n)\exp(-s(Z_n + A_1)) - \exp((t-s)X_0)\exp(-s(X_0 + A_1)) \\ &= \exp((t-s)Z_n)(\exp(-s(Z_n + A_1)) - \exp(-s(X_0 + A_1))) + (\exp((t-s)Z_n) \\ &\quad - \exp((t-s)X_0))\exp(-s(X_0 + A_1)) \end{aligned} \quad (2.26)$$

and the expressions (1.5), (2.18), it follows that

$$\begin{aligned} &\Big\|\int_0^t \{\exp((t-s)Z_n)\exp(-s(Z_n + A_1)) - \exp((t-s)X_0)\exp(-s(X_0 + A_1))\}ds\Big\| \\ &\le 2t\exp(t(2a + \|A_1\|))(\exp(t\|Z_n - X_0\|) - 1). \end{aligned}$$

Hence and from (2.15)–(2.26), it follows that

$$\|X_n(t) - X(t)\| \le K_1(t)\|Z_n - X_0\| + K_2(t)(\exp(t\|Z_n - X_0\| - 1) \tag{2.27}$$

where

$$K_1(t) = t\|C_0\| \exp(t(2a + \|A_1\|)),$$

$$K_2(t) = 6ft^2 \exp((3a + \|A_1\|)t) + 2t(\|C_1\| + a\|C_0\|) \exp(t(2a + \|A_1\|)) + \|C_0\| \exp(at). \tag{2.28}$$

Note that the scalar functions $K_i(t)$, for $i = 1, 2$, defined by (2.28), are continuous and thus they are bounded on any finite interval $[0, b]$. Hence and from the previous comments the following result has been established:

Theorem 2. *Let us suppose that X_0 is a solution of equation (1.4), and let $\{Z_n\}_{n\ge1}$ be a sequence of matrices norm convergent to X_0 and let $X_n(t)$ be defined by (2.1)–(2.2).*

(i) *The sequence of matrix functions $X_n(t)$ is pointwise convergent, on the real line, to the unique solution $X(t)$ of problem (1.1). For fixed real number t one has $\|X_n(t) - X(t)\| = O(\|Z_n - X_0\|)$, when $n \to \infty$.*

(ii) *On any finite interval $[0, b]$, the sequence of matrix functions $X_n(t)$ is uniformly convergent to the solution $X(t)$ of problem (1.1). Also, if $t \in [0, b]$, it follows that $\|X_n(t) - X(t)\| = O(\|Z_n - X_0\|)$, when $n \to \infty$, uniformly for $t \in [0, b]$.*

Remark 1. It is interesting to remark that the method provided by Theorems 1, 2 is useful independently of the procedure for obtaining the approximate solutions of equation (1.4). Theorem 2 means that the approximation error for the approximate solutions $X_n(t)$, of the differential problem (1.1), has the same order as that of the approximate solutions of the algebraic problem (1.4). Also, our method only is applicable when a solution of equation (1.4) is available. Unlike the scalar case, equation (1.4) may be unsolvable; for instance, if A_0 and A_1 commute, and the minimal polynomial of the matrix $A_1^2 - 4A_0$ has a double root at the origin, then equation (1.4) has no root[7]. In order to compute the approximate solutions $X_n(t)$, given by Theorem 1, it is interesting to recall that in [15] useful methods for computing the integrals appearing in (2.1)–(2.2) are given. Also, a numerically useful method for computing the exponential $\exp(tZ_n)$, avoiding the computation of the eigenvalues of Z_n, is given in [13].

Example 1. Let us consider the problem

$$X^{(2)} + A_1X^{(1)} + A_0X = F(t); \quad X(0) = C_0, \quad X^{(1)}(0) = C_1, \quad F \text{ continuous} \tag{2.29}$$

where A_1 is an invertible matrix such that

$$d = (1 - 4\|A_1^{-1}\|\ \|A_1^{-1}A_0\|)^{1/2} > 0$$

and let a, b be defined by the expressions

$$a = (1 - d)/(2\|A_1^{-1}\|), \quad b = (2\|A_1^{-1}\|)^{-1}.$$

Then starting from any matrix Z_1, satisfying $\|Z_1\| \le \alpha$, where $a \le \alpha < b$, the sequence of matrices $\{Z_n\}_{n\ge1}$ defined by

$$Z_{n+1} = F(Z_n), \quad F(Z_n) = -A_1^{-1}Z_n^2 - A_1^{-1}A_0, \quad n \ge 1 \tag{2.30}$$

converges to a solution X_0 of equation (1.4), such that

$$\|Z_n - X_0\| \le \{(f'(\alpha))^n/(1 - f'(\alpha))\}\|Z_2 - Z_1\|, \quad n \ge 2$$

where $f(s) = \|A_1^{-1}\|s^2 + \|A_1^{-1}A_0\|$, and one has $0 \leq f'(\alpha) < 1$; see [2, p.476] for details. Hence and from Theorem 2, on any bounded interval, the sequence of approximations $X_n(t)$, defined by Theorem 1, where Z_n are defined by (2.30), satisfies $\|X_n(t) - X(t)\| = O((f'(\alpha))^n)$, when $n \to \infty$, where $X(t)$ represents the exact solution of problem (2.29).

References

[1] Jr.J.E. Dennis, J.F. Traub and R.P. Weber, The algebraic theory of matrix polynomials, *SIAM J. Numer. Anal.,* **6** (1976), 831–845.

[2] J. Eisenfeld, Operator equations and nonlinear eigenparameter problems, *J. Funct. Anal.,* **4** (1973), 475–490.

[3] G. Golub and C.P. Van Loan, Matrix Computations, John Hopkins Univ. Press, Baltimore, M.D., 1983.

[4] L. Jódar, Boundary value problems for second order operator differential equation, *Linear Algebra and Its Appls.,* **83** (1986), 29–38.

[5] L. Jódar, Explicit expressions for Sturm-Liouville operator problems, Proc. of The Edinburgh Math. Soc., **30** (1987), 301–309.

[6] P. Lancaster, Lambda Matrices and Vibrating Systems, Pergamon Elmsford, New York, 1966.

[7] L. Lovass-Nagy and D.L. Powers, A note on block diagonalization of some partitioned matrices, *Linear Algebra and Its Appls.,* **5** (1972), 339–346.

[8] C.B. Moler, and C.P. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, *SIAM Review,* **20** (1978), 801–836.

[9] M. Nashed, On general iterative methods for the solution of a class of nonlinear operator equations, *Math. Comp.,* **19** (1965), 14–24.

[10] J.M. Ortega and W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[11] W. Petryshyn, Remarks on fixed point theorems and their extensions, *Trans. Amer. Math. Soc.,* **126** (1967), 43–54.

[12] L. Rall, Computational Solution of Nonlinear Operator Equations, John Wiley, New York, 1969.

[13] H.J. Runckell and P. Pittelkow, Practical computation of matrix functions, *Linear Algebra and Its Appls.,* **49** (1983), 161–178.

[14] J. Traub, Iterative Methods for the Solution of Equations, Prentice Hall, Englewood Cliffs, New Jersey, 1964.

[15] C.P. Van Loan, Computing integrals involving the matrix exponential, *IEEE Trans. Autom. Control,* **23** (1974), 395–404.