

Document downloaded from:

<http://hdl.handle.net/10251/56205>

This paper must be cited as:

Blanes Zamora, S.; Ponsoda Miralles, E. (2012). Time-averaging and exponential integrators for non-homogeneous linear IVPs and BVPs. *Applied Numerical Mathematics*. 62(8):875-894. doi:10.1016/j.apnum.2012.02.001.



The final publication is available at

<http://dx.doi.org/10.1016/j.apnum.2012.02.001>

Copyright Elsevier

Additional Information

Exponential integrators for non-homogeneous linear IVPs and BVPs

S. Blanes* and E. Ponsoda

*Instituto de Matemática Multidisciplinar, Building 8G, second floor,
Universitat Politècnica de València. 46022 Valencia, Spain.*

Abstract

We consider exponential integrators based on the Magnus series expansion for the numerical integration of general linear non-homogeneous differential equations. The schemes can be considered as averaged methods which transform, for one time step, a non-autonomous problem into an autonomous one whose flows agree up to a given order of accuracy at the end of the time step. The problem is reformulated as a particular case of a matrix Riccati differential equation and the Möbius transformation is considered, leading to an homogeneous linear problem. The methods proposed can be used both for initial value problems (IVPs) as well as for two point boundary value problems (BVPs). In addition, they allow to use different approximations for different parts of the equation, e.g. the homogeneous and non-homogeneous parts, or to use adaptive time steps. The particular case of separated boundary conditions using the imbedding formulation is also considered. This formulation allows to transform an stiff and badly conditioned BVP into a set of IVPs which can be integrated using some of the previous methods. The performance of the methods is illustrated on some numerical examples.

Key words:

Exponential integrators, matrix Riccati differential equation, linear initial and boundary value problem, imbedding formulation

2010 MSC: 65L10, 34B05, 65D30

* Corresponding author.

Email addresses: serblaza@imm.upv.es (S. Blanes), eponsoda@imm.upv.es (E. Ponsoda).

1 Introduction

In this work we consider exponential integrators for the numerical integration of the general non-homogeneous linear matrix differential equation

$$Y'(t) = M(t)Y(t) + Y(t)N(t) + F(t), \quad t \in [t_0, T] \quad (1)$$

where $Y(t), F(t) \in \mathbb{C}^{p \times q}$, $M(t) \in \mathbb{C}^{p \times p}$ and $N(t) \in \mathbb{C}^{q \times q}$. We consider numerical methods which are valid both for the initial value problem (IVP) with initial conditions

$$Y(t_0) = Y_0 \quad (2)$$

as well as for the boundary value problem (BVP) with boundary conditions at two points

$$B_0Y(t_0) + B_1Y(T) = \gamma, \quad (3)$$

with $B_0, B_1 \in \mathbb{C}^{p \times p}$, $\gamma \in \mathbb{C}^{p \times q}$ (or its extension to conditions on m different points, $B_1Y(t_1) + \dots + B_mY(t_m) = \gamma$). It is important to remark that some methods for solving nonlinear BVPs consider a linearization of the nonlinear differential equation and then one has to solve a number of non autonomous linear BVPs with appropriate boundary conditions [44,47].

The formal solution of (1)-(2) is given by

$$Y(t) = \Phi_M(t, t_0)Y_0\Phi_N^*(t, t_0) + \Psi(t, t_0) \quad (4)$$

with

$$\Psi(t, t_0) = \int_{t_0}^t \Phi_M(t, s)F(s)\Phi_N^*(t, s)ds \quad (5)$$

where $\Phi_M(t, t_0)$ and $\Phi_N^*(t, t_0)$ are the fundamental solution matrices of the associated homogeneous equations

$$\Phi_M'(t, t_0) = M(t)\Phi_M(t, t_0), \quad \Phi_M(t_0, t_0) = I_p,$$

$$\Phi_N^{*'}(t, t_0) = \Phi_N^*(t, t_0)N(t), \quad \Phi_N^*(t_0, t_0) = I_q.$$

If M is a constant matrix then $\Phi_M(t, t_0) = e^{(t-t_0)M}$, and similarly for N . However, the non-homogeneous term in (5) can not be solved, in general, in closed form and has to be integrated numerically. The accurate numerical integration of this non-homogeneous part turns into a much more involved task when the matrices $M(t)$ and $N(t)$ are explicitly time-dependent where, in addition, Φ_M and Φ_N^* can not be written, in general, in a closed form. Notice that to approximate $\Psi(t, t_0)$ by using a quadrature rule requires the

computation of $\Phi_M(t, s)$ and $\Phi_N^*(t, s)$ at the values of s given by the quadrature nodes.

Problem (1) can be reformulated in a simpler way as follows

$$\mathbf{y}'(t) = M(t)\mathbf{y}(t) + \mathbf{h}(t), \quad t_0 \leq t \leq T, \quad (6)$$

with $\mathbf{y}(t), \mathbf{h}(t) \in \mathbb{C}^{pq}$, $M(t) \in \mathbb{C}^{pq \times pq}$, with appropriate initial or boundary conditions. This formulation is the most frequently used in the literature because formally simplifies the problem, but one has to deal with larger dimensional and sparse matrices.

For the IVP, one considers a mesh $t_0 < t_1 < t_2 < \dots < t_L = T$ where $t_{i+1} = t_i + h_i$. For simplicity in the presentation, we can consider at this point an equidistant time grid $t_i = t_0 + ih$, $0 \leq i \leq L$, with constant time step $h = (T - t_0)/L$. Then, our goal is to look for numerical approximations to the solution at the mesh points, $Y_i \equiv Y(t_i)$. In particular, in this work we look for efficient numerical methods to approximate the fundamental solution matrices, Φ_M, Φ_N^* , as well as $\Psi(t_i, t_0)$, i.e. to obtain an approximate solution in the form

$$Y_i = \Phi_{M,i} Y_0 \Phi_{N,i}^* + \Psi_i, \quad (7)$$

where

$$\Phi_{M,i+1} = \Phi_M(t_{i+1}, t_i) \Phi_{M,i}$$

$$\Phi_{N,i+1}^* = \Phi_N^*(t_{i+1}, t_i)$$

$$\Psi_{i+1} = \Psi_i + \Psi(t_{i+1}, t_i),$$

with $\Phi_{M,0} = I_p$, $\Phi_{N,0}^* = I_q$, $\Psi_0 = 0_{p \times q}$.

For the BVP (1)-(3), the situation is slightly more involved. For the simplest shooting method one can compute the end point, $Y_L = \Phi_{M,L} Y_0 \Phi_{N,L}^* + \Psi_L$, to be included in the boundary conditions,

$$B_0 Y_0 + B_1 (\Phi_{M,L} Y_0 \Phi_{N,L}^* + \Psi_L) = \gamma, \quad (8)$$

and then we obtain simple algebraic equations which allows (if the problem is well conditioned and has unique solution) to find the missing values for the initial conditions, Y_0 . Alternatively, if the algebraic equations are not well conditioned, we can consider $Y_0 = \Phi_{M,L}^{-1} (Y_L - \Psi_L) \Phi_{N,L}^{*-1}$ and to look for the missing values on the final conditions Y_L (or at any intermediate point). If L is not too large which makes it feasible to store $\Phi_{M,i}, \Phi_{N,i}^*, \Psi_i$, $i = 1, \dots, L$, then one can use (7) to obtain the approximated solution on the mesh, otherwise one solves (8) and proceeds as for the IVP. The shooting method is frequently used with variable step and variable order explicit methods, showing a high efficiency for most problems. However, if the problem is stiff or very sensitive to the initial conditions, explicit methods frequently suffers from stability problems and

the schemes can be affected by unacceptable error accumulations which makes the numerical solution useless. This drawback can be considerably diminished with the multiple shooting technique. This technique is presented in detail in, e.g. [3,37,52] for the problem (6), and basically requires to compute and store $\Phi_{M,i}, \Phi_{N,i}^*, \Psi_i, i = 1, \dots, L^*$ in a reduced mesh $t_0 < \tau_1 < \tau_2 < \dots < \tau_{L^*} = T$ with $L^* < L$. Depending on the problem, one can use the *standard* multiple shooting or the more elaborated marching techniques for multiple shooting.

For stiff problems, implicit methods are usually preferred. Finite difference methods are perhaps the most popular methods. They require to solve relatively large systems of algebraic equations which in most cases is equivalent to use implicit Runge-Kutta methods whose implicit equations are solved all together at the end of the integrations. An attempt to use explicit methods for stiff problems corresponds to the imbedding formulation or Riccati method, see [3,20,21,36], which intends to transform the system of ODEs into IVPs which are more stable. This technique involves the transformation of the equations and requires the numerical integration forward and backward of coupled equations, leading to relatively involved numerical algorithms.

In this work, we analyse the numerical solution of the IVP or BVP (1)-(3) using Magnus integrators as a representative of Lie group methods developed for the homogeneous linear problems and in particular, we focus on the exponential integrators. Magnus integrators have also been developed for nonlinear differential equations [14] (see also [7] and references therein) but we consider the techniques used for linear problems which will allow us to obtain simpler and more efficient algorithms. Previous attempts to use high order exponential integrators for linear problems are mainly addressed to the basic homogeneous problem ($N = 0, F = 0$) (see [7,31] and references therein) and occasionally to the more general homogeneous case ($F = 0$), see [28]. We consider this problem as a particular case of the matrix Riccati equation when it is solved using the Möbius transformation (Radon's Lemma) which transforms the problem to an homogeneous system, see [1,51]. Exponential integrators are explicit methods which can be used with variable time step and variable order, and have similar stability to the implicit methods. For this reason, they are methods suitable for stiff and oscillatory problems.

1.1 The Riccati differential equation

Let us review the close relation between the linear systems and the matrix Riccati differential equation (RDE)

$$X'(t) = B(t) + A(t)X(t) - X(t)D(t) - X(t)C(t)X(t), \quad X(t_0) = X_0, \quad (9)$$

where $X(t) \in \mathbb{C}^{p \times q}$, $A(t) \in \mathbb{C}^{p \times p}$, $B(t) \in \mathbb{C}^{p \times q}$, $C(t) \in \mathbb{C}^{q \times p}$, $D(t) \in \mathbb{C}^{q \times q}$.

Initial value problems for Riccati differential systems arise frequently also in important applications to classical control theory, e.g. [5,49], and in non-cooperative control theory appearing in economic or military problems (see Refs. [4,17,40] and the references therein). The case $p \neq q$ appears, for instance, in differential games when Nash strategies in non-cooperative control problems are tackled, see [18,54] and in general linear two point boundary value problems with separated boundary conditions, e.g. [3,20,37].

This equation has received considerable attention in the literature, both focused to its theoretical aspect, e.g [1,5,49] and its numerical treatment, see [11,19,33–35,38,51]. Frequently, the RDE appears coupled with other differential equations and the numerical schemes for the RDE have to be adapted properly, being this part the bottle neck to build efficient algorithms. The computational cost and the accuracy achieved strongly depend on the algorithm used for the numerical integration of this part.

In this work we consider exponential integrators for the numerical integration of the RDE by using Möbius schemes, see [1,51]. With appropriate change of coordinates the system transforms into an equivalent larger linear system which allows to use numerical integrators tailored for homogeneous linear systems which have shown a high performance.

Let us consider the Riccati equation (9) and apply the transformation

$$X(t) = V(t)W^{-1}(t), \quad (10)$$

with $V(t) \in \mathbb{C}^{p \times q}$, $W(t) \in \mathbb{C}^{q \times q}$ and $V(t_0) = X_0, W(t_0) = I_q$, in the region where $W(t)$ is invertible. Then eq. (9) is equivalent to solve the linear system

$$Y'(t) = S(t)Y(t), \quad Y(t_0) = \begin{bmatrix} X_0 \\ I_q \end{bmatrix}, \quad (11)$$

with

$$Y(t) = \begin{bmatrix} V(t) \\ W(t) \end{bmatrix}, \quad S(t) = \begin{bmatrix} A(t) & B(t) \\ C(t) & D(t) \end{bmatrix}. \quad (12)$$

If we denote $Y(t) = \Phi(t, t_0)Y_0$ with Φ the fundamental matrix solution given by

$$\Phi(t, t_0) = \begin{bmatrix} \alpha(t) & \beta(t) \\ \gamma(t) & \delta(t) \end{bmatrix}$$

where $\alpha(t) \in \mathbb{C}^{p \times p}$, $\beta(t) \in \mathbb{C}^{p \times q}$, $\gamma(t) \in \mathbb{C}^{q \times p}$ and $\delta(t) \in \mathbb{C}^{q \times q}$, then from the initial condition given by (11), it follows that $V(t) = \alpha(t)X_0 + \beta(t)$,

$W(t) = \gamma(t)X_0 + \delta(t)$, and from (10) we obtain

$$X(t) = [\alpha(t)X_0 + \beta(t)][\gamma(t)X_0 + \delta(t)]^{-1}. \quad (13)$$

Notice that $X(t)$ is not well defined for those values of t where the matrix $[\gamma(t)X_0 + \delta(t)]$ has no inverse. On the other hand, since the linear equation (11) can be solved at any time (if $S(t)$, given by (12), has no singularities) then we can obtain accurate numerical approximations to $X(t)$ when it evolves close to a singularity, see [51]. One can expect large errors around the singularities which then disappear once we cross them. If the RDE is coupled with other equations, a large error for $X(t)$ at a given instant will cause large errors in the remaining equations and then one should expect unacceptable errors on the numerical solutions of the whole problem.

In this work we consider the equation (1) as a particular case of (9) and then, as a particular case of (11)-(12). This way allows to use the highly efficient geometric integrators recently developed in the literature, and which we briefly present.

In sections 2 and 3 we present Lie group and Magnus integrators which can be used, independently, for solving IVPs as well as BVPs. Section 4 focuses in two point BVPs and how to transform the problem into linear IVPs which are not badly conditioned, and section 5 presents some numerical examples.

2 Lie-group solvers for linear homogeneous equations

Lie group methods and, in particular, exponential integrators for linear initial value problems have experienced a renewed interest during the last years (see [26,31] and references therein). These methods have shown to be superior (both qualitatively and quantitatively) to standard methods for solving linear IVPs for an important number of cases. This is frequently the case, e.g. for oscillatory or stiff problems, where the exponential representation is more accurate than polynomial or rational approximations as it is the case for most standard methods like, say, Runge–Kutta methods. It is then important to know if this performance still remains when applied to linear BVPs.

Exponential methods are usually computationally expensive methods. However, by using appropriate algorithms, this extra cost can be minimized and can be compensated with more accurate results. They are explicit methods which can be considered as generalizations of implicit methods (see Appendix A). This fact allows to use relatively large time steps.

Let us consider the homogeneous linear equation

$$Z'(t) = A(t)Z(t), \quad Z(t_0) = Z_0.$$

The fundamental matrix solution satisfies the homogeneous matrix differential equation

$$\Phi'_A(t, t_0) = A(t) \Phi_A(t, t_0), \quad \Phi_A(t_0, t_0) = I,$$

and it is well known that the matrix solution Φ_A can be written, in a neighborhood of t_0 as follows (for simplicity we denote $\Phi(t_n) \equiv \Phi_A(t_n, t_0)$)

$$\Phi(t_{n+1}) = e^{\Omega(h)} \Phi(t_n) \tag{14}$$

where Ω can be written as the Magnus series expansion [39]. References about Magnus expansion (14) can be found, e.g., in [7,9,12,32]. Alternatively, one can use the Fer expansion [10,24,48], or its symmetric version [10], or the Cayley method [22,29].

2.1 Application to the non-homogeneous linear equation

If we consider the non-homogeneous linear problem (1) as a particular case of the Riccati equation (9) with the Möbius transformation, then it is equivalent to solve the linear system

$$\begin{bmatrix} V(t) \\ W(t) \end{bmatrix}' = \begin{bmatrix} M(t) & F(t) \\ 0 & -N(t) \end{bmatrix} \begin{bmatrix} V(t) \\ W(t) \end{bmatrix}, \quad \begin{bmatrix} V(0) \\ W(0) \end{bmatrix} = \begin{bmatrix} X_0 \\ I_q \end{bmatrix}, \tag{15}$$

with

$$\Phi(t) = \begin{bmatrix} \Phi_M(t) & \beta(t) \\ 0 & \Phi_{-N}(t) \end{bmatrix}$$

and we obtain from (11)-(13)

$$Y(t) = [\Phi_M(t)Y_0 + \beta(t)]\Phi_{-N}(t)^{-1},$$

where it is well known that $\Phi_{-N}(t)^{-1} = \Phi_N^*(t)$. It is then clear that $\beta(t) = \Psi(t)\Phi_{-N}(t)$, since it must agree with the solution (4). Notice that, contrarily to the general case for the solution of the RDE (13), $Y(t)$ is well defined for all values of t .

If $M(t)$, $N(t)$ and $F(t)$ are periodic with period T then the Floquet theory can be trivially applied if we reformulate the problem as in (15). We have that

$$\Phi(t) = P(t) \exp(tQ)$$

with Q a constant matrix and $P(t + T) = P(t)$, both with the following structures

$$P(t) = \begin{bmatrix} P_M(t) & p_\beta(t) \\ 0 & P_N(t) \end{bmatrix}, \quad Q = \begin{bmatrix} Q_M & q_\beta \\ 0 & Q_N \end{bmatrix}.$$

In the following section we present Magnus integrators which in this case would allow to approximate each block matrix on Q and P with different orders of accuracy.

3 Magnus approximations

As a representative for the Lie group methods, we consider methods based on the Magnus series expansion (14), which can be considered as a class of exponential integrators (similar schemes can be obtained with other schemes (see [10])). The Magnus series is given by $\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t)$ where the first two terms are given by

$$\Omega_1(t) = \int_{t_0}^t A(t_1) dt_1, \quad \Omega_2(t) = \frac{1}{2} \int_{t_0}^t \int_{t_0}^{t_1} [A(t_1), A(t_2)] dt_2 dt_1,$$

and $[A, B] \equiv AB - BA$ is the matrix commutator of A and B . It is known that the Magnus series is absolutely convergent for a general complex matrix, $A(t)$, for $t_0 \leq t < T$, with

$$T = \max \left\{ t \geq t_0 : \int_{t_0}^t \|A(s)\|_2 ds < 1.08686869 \dots \right\}, \quad (16)$$

where $\|\cdot\|_2$ denotes the two-norm of a matrix. This is a sufficient condition and, under certain conditions, this bound can be enlarged (see [7,13,41] and references therein). The following result can be useful if one applies the Magnus expansion to the non homogeneous equation (6) rewritten as

$$\begin{bmatrix} \mathbf{y}(t) \\ 1 \end{bmatrix}' = \begin{bmatrix} M(t) & \mathbf{h}(t) \\ \mathbf{0}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y}(t) \\ 1 \end{bmatrix} \quad (17)$$

or shortly

$$\mathbf{x}'(t) = A(t)\mathbf{x}(t). \quad (18)$$

Theorem 3.1 *The convergence of the Magnus expansion applied to the non homogeneous equation (6) rewritten as (17)-(18), with a bounded non homogeneous term, is independent of the nonhomogeneous term.*

Proof 3.2 *Let us denote*

$$(Z, \mathbf{z}) \equiv \begin{bmatrix} Z & \mathbf{z} \\ \mathbf{0}^T & 0 \end{bmatrix}$$

and consider the matrices and vectors $A, \mathbf{a}, B, \mathbf{b}$ such that for the two-norm of matrices and vectors we have $\|A\| \leq C_A$, $\|B\| \leq C_B$ and exist a constant α such that $\|\mathbf{a}\| \leq \alpha C_A$, $\|\mathbf{b}\| \leq \alpha C_B$. Consider now

$$(C, \mathbf{c}) = (A, \mathbf{a}) + (B, \mathbf{b}), \quad (D, \mathbf{d}) = [(A, \mathbf{a}), (B, \mathbf{b})]$$

and the constants $C_C = C_A + C_B$, $C_D = 2C_A C_B$, then it is immediate to check that for the two-norm

$$\|C\| \leq C_C, \quad \|\mathbf{c}\| \leq \alpha C_C, \quad \text{and} \quad \|D\| \leq C_D, \quad \|\mathbf{d}\| \leq \alpha C_D,$$

with the same value of α . Given the equation (17) and a positive definite function $K(t)$ such that $\|M(t)\| \leq K(t)$, if $\mathbf{h}(t)$ is a bounded function then it exists a constant, α such that $\|\mathbf{h}(t)\| \leq \alpha K(t)$. Under the convergent condition (16) for the Magnus series expansion we have that

$$\Omega_A = (\Omega_M, \omega_{\mathbf{h}}).$$

Since the Magnus series only involves operations in the algebra (linear combinations and commutators) for the two-norm we have that inside the convergence domain it exists a function, $Q(t)$, such that $\|\Omega_M(t)\| \leq Q(t)$ (see [7]) and then

$$\|\omega_{\mathbf{h}}\| \leq \alpha Q(t)$$

which is also a bounded function, and then the convergence does not depend on the non-homogeneous vector function $\mathbf{h}(t)$. \square

Remark 3.3 *From the equivalence between norms, there exists positive constants k_1 and k_2 such that*

$$k_1 \|B\| \leq |||B||| \leq k_2 \|B\|,$$

where $\|\cdot\|$ and $|||\cdot|||$ are two different matrix norms and B is a square matrix. In this way, from [46, p. 25], given any $\epsilon > 0$, there exists a norm $\|\cdot\|_\epsilon$ such that

$$\|A(t)\|_\epsilon = \left\| \begin{bmatrix} M(t) & \mathbf{h}(t) \\ \mathbf{0}^T & 0 \end{bmatrix} \right\|_\epsilon \leq \rho(A(t)) + \epsilon,$$

where $\rho(B)$ denotes the spectral radius of the square matrix B . If we denote by $\sigma(B)$ the set of all eigenvalues of the square matrix B , note that $\sigma(A(t)) = \{0\} \cup \sigma(M(t))$, and then $\rho(A(t)) = \rho(M(t))$. So, the result of theorem 3.1 is very clear.

The convergence of the Magnus series expansion applied to (17) depends only on $M(t)$. However, for problems where $\|\mathbf{h}(t)\| \gg \|M(t)\|$ the convergence can be slow and the error originated from the approximated solutions obtained when truncating the series can be significant. For this reason, it is convenient to build new tailored methods for this problem. We first review general methods and next we show how they can be adapted to this problem.

In order to obtain methods which can be easily used with any quadrature rule we introduce the averaged (or generalized momentum) matrices

$$A^{(i)}(h) \equiv \frac{1}{h^i} \int_{t_n}^{t_n+h} (t - t_{1/2})^i A(t) dt = \frac{1}{h^i} \int_{-h/2}^{h/2} t^i A(t + t_{1/2}) dt \quad (19)$$

for $i = 0, \dots, s-1$, and with $t_{1/2} = t_n + h/2$. It is known that the matrices $A^{(i)}$, $i = 0, \dots, s-1$ suffice to approximate all previous expansions up to order $p = 2s$ in the time step, i.e. the local error is $\mathcal{O}(h^{p+1})$. For instance, to build a second order method in the Magnus expansion we can consider that

$$\Omega^{[2]} = A^{(0)}, \quad (20)$$

to build a fourth order method we have that

$$\Omega^{[4]} = A^{(0)} + [A^{(1)}, A^{(0)}], \quad (21)$$

and to build a sixth order method

$$\begin{aligned} \alpha_1 &= \frac{3}{4}(3A^{(0)} - 20A^{(2)}), & \alpha_2 &= 12A^{(1)}, & \alpha_3 &= -15(A^{(0)} - 12A^{(2)}), \\ C_1 &= [\alpha_1, \alpha_2], & C_2 &= -\frac{1}{60}[\alpha_1, 2\alpha_3 + C_1], & C_3 &= \frac{1}{240}[-20\alpha_1 - \alpha_3 + C_1, \alpha_2 + C_2] \\ \Omega^{[6]} &\equiv A^{(0)} + C_3, \end{aligned} \quad (22)$$

(see [10,9] for eight-order methods). They satisfy $\Omega^{[p]} = \Omega + \mathcal{O}(h^{p+1})$, so

$$\exp(\Omega^{[p]}(t_{i+1}, t_i)) = \Phi(t_{i+1}, t_i) + \mathcal{O}(h^{p+1}).$$

If the exact evaluation of the unidimensional integrals in (19) is not possible or is computationally expensive, any numerical quadrature of the same order as the method may be used instead. Then, it suffices to approximate the matrices $A^{(0)}$ up to second order, or $A^{(0)}, A^{(1)}$ up to order four or $A^{(0)}, A^{(1)}, A^{(2)}$ up to order six to get methods of order two, four or six, respectively. Suppose that b_i, c_i , ($i = 1, \dots, k$), are the weights and nodes of a particular quadrature rule of order p , then it is possible to approximate all the integrals $A^{(i)}$ (up to the required order) just by using only the evaluations $A_i \equiv A(t_n + c_i h)$ at the nodes c_i of the quadrature rule, i.e.

$$\tilde{A}^{(i)} = h \sum_{j=1}^k b_j \left(c_j - \frac{1}{2}\right)^i A_j. \quad i = 0, \dots, s-1,$$

and substituting $\tilde{A}^{(i)}$ by $A^{(i)}$ in (20)-(22), they still retain the order of accuracy in $\Omega^{[p]}$.

Let us consider the particular case (15) which concerns us in this work. If we denote by $\tilde{\Omega}^{[p]}$ an approximation to the Magnus series expansion accurate up to order h^p , then we have that from (20)

$$\tilde{\Omega}^{[2]} = \begin{bmatrix} M^{(0)} & F^{(0)} \\ 0 & -N^{(0)} \end{bmatrix},$$

and from (21)

$$\tilde{\Omega}^{[4]} = \begin{bmatrix} M^{[4]} & F^{[4]} \\ 0 & -N^{[4]} \end{bmatrix},$$

where

$$M^{[4]} = M^{(0)} + [M^{(1)}, M^{(0)}], \quad N^{[4]} = N^{(0)} + [N^{(1)}, N^{(0)}]$$

$$F^{[4]} = F^{(0)} + M^{(1)}F^{(0)} - M^{(0)}F^{(1)} + F^{(0)}N^{(1)} - F^{(1)}N^{(0)}.$$

Here we have used the notation for the averaged momentum matrices (19). If M, N are constant matrices but F is time-dependent, it is immediate to see that then

$$M^{(0)} = hM, \quad N^{(0)} = hN, \quad M^{(1)} = 0, \quad N^{(1)} = 0,$$

and the fourth-order method (as well as higher-order methods) is just a small modification of the second order one where only $F^{[p]}$ changes with p .

By exponentiation we have, for the integration on the interval $[t_{i+1}, t_i]$

$$\exp\left(\Omega_i^{[p]}\right) = \begin{bmatrix} e^{\Omega_{M,i}} & g_i^{[p]} \\ 0 & e^{\Omega_{-N,i}} \end{bmatrix},$$

leading to a recursive scheme to obtain the necessary fundamental matrix solution at each interval.

$$\begin{aligned} \begin{bmatrix} \Phi_{M,i} & G_i \\ 0 & \Phi_{-N,i} \end{bmatrix} &= \exp\left(\Omega_i^{[p]}\right) \cdot \begin{bmatrix} \Phi_{M,i-1} & G_{i-1} \\ 0 & \Phi_{-N,i-1} \end{bmatrix} \\ &= \begin{bmatrix} e^{\Omega_{M,i}}\Phi_{M,i-1} & e^{\Omega_{M,i}}G_{i-1} + g_i^{[p]}\Phi_{-N,i-1} \\ 0 & e^{\Omega_{-N,i}}\Phi_{-N,i-1} \end{bmatrix} \end{aligned}$$

which allows to approximate the solution (4) at the mesh for different initial or boundary conditions.

3.1 Methods of order (p, n) .

In order to illustrate the versatility of the proposed Magnus integrators, let us consider the case where the matrices $M(t), N(t)$ evolve adiabatically. Then, in a fourth-order method $M^{(1)}$ and $N^{(1)}$ can be neglected while $M^{(0)}$ and $N^{(0)}$ can be approximated by a second order rule (e.g. the midpoint rule or the trapezoidal rule). In this case, a second order method, which would correspond to a fourth-order method in the limit where $M(t), N(t)$ are constant matrices, is given by

$$\tilde{\Omega}^{[2,4]} = \begin{bmatrix} M^{(0)} & F^{(0)} - M^{(0)}F^{(1)} - F^{(1)}N^{(0)} \\ 0 & -N^{(0)} \end{bmatrix}. \quad (23)$$

This scheme has, basically, the cost and stability of a second order scheme (in the case where the cost to compute the matrices $M(t), N(t)$ is considerably higher than $F(t)$) but the accuracy of a fourth order one in the limit when the matrices $M(t), N(t)$ evolve adiabatically.

If we use the midpoint rule to approximate $M^{(0)}, N^{(0)}$ and the fourth-order Gauss-Legendre quadrature rule for $F^{(0)}, F^{(1)}$, we have

$$M^{(0)} \simeq hM\left(t_n + \frac{h}{2}\right), \quad N^{(0)} \simeq hN\left(t_n + \frac{h}{2}\right), \quad F^{(0)} \simeq \frac{h}{2}(F_1 + F_2), \quad F^{(1)} \simeq \frac{h\sqrt{3}}{12}(F_2 - F_1),$$

where $F_i = F(t_n + c_i h)$, $i = 1, 2$ with $c_1 = \frac{1}{2} - \frac{\sqrt{3}}{6}$, $c_2 = \frac{1}{2} + \frac{\sqrt{3}}{6}$. Alternatively, if for example $N(t), F(t)$ evolve adiabatically a $(4,2)$ method could be built

with

$$\tilde{\Omega}^{[2,4]} = \begin{bmatrix} M^{[4]} & F^{(0)} + M^{(1)}F^{(0)} \\ 0 & -N^{(0)} \end{bmatrix}.$$

This can be extended to higher orders and to different families of problems (e.g. if $F(t)$ is highly oscillatory) being this an important subject under investigation at this moment. It is important to remark that, by using this formulation, it is possible to use in a relatively simple way preliminary transformations like the interaction picture or adiabatic picture (see [7] and references therein) to improve the performance of the integration of (1)-(3).

The schemes proposed require an efficient algorithm to compute the exponential or its action on a vector (see for example [2,15,26,42,53] and references therein). In addition, variable time-steps can be easily used (see Appendix B).

3.2 Solving the autonomous problem

The map (14) can be considered as the h -flow for the autonomous problem $\Phi' = \frac{1}{h}\Omega(h)\Phi$ so, the scheme has to be implemented with an efficient algorithm to approximate the exponential or its action on a vector. We review some results for the following autonomous problem

$$\mathbf{y}' = A\mathbf{y} + \mathbf{g},$$

with $\mathbf{y}(t) \in \mathbb{C}^p$ and the solution, for one time step, h , is given by

$$\mathbf{y}(t_n + h) = e^{hA}\mathbf{y}(t_n) + \int_{t_n}^{t_n+h} e^{(t_n+h-s)A}\mathbf{g}ds = e^{hA}\mathbf{y}(t_n) + h\varphi(hA)\mathbf{g},$$

where $\varphi(z) = (e^z - 1)/z$. The algorithm requires again to compute the exponential of a matrix or its action on a vector (see [15,26,42,43,53] and references therein). For large systems like in discretisations of PDEs it is usual to consider approximations to the products $e^D\mathbf{v}$ or $\varphi(D)\mathbf{v}$ (for a given matrix D and a vector \mathbf{v}) by considering Krylov subspaces, see [50] and references therein. For this reason, some authors recommend to rewrite the solution as (see [53])

$$\begin{aligned} \mathbf{y}(t_n + h) &= e^{hA}\mathbf{y}(t_n) + h\varphi(hA)\mathbf{g} \\ &= (hA\varphi(hA) + I)\mathbf{y}(t_n) + h\varphi(hA)\mathbf{g} \\ &= \mathbf{y}(t_n) + h\varphi(hA)(A\mathbf{y}(t_n) + \mathbf{g}), \end{aligned}$$

which only involves one product and one Krylov approximation to the function $\varphi(hA)$. Another possibility, which has not been much explored, is to consider

an approximation to the exponential

$$\begin{bmatrix} \mathbf{y}(t) \\ 1 \end{bmatrix} = \exp \begin{bmatrix} tA & tg \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y}(0) \\ 1 \end{bmatrix},$$

whose computational cost is very similar to the cost of e^{tA} since it only includes one column and a zero row. In [2] it is suggested to consider the approximation to this exponential because the algorithms for the φ functions are much less well developed. There exists, however, recent works to improve the algorithms to compute the actions of the φ functions (see [45] and references therein).

The situation is slightly more involved for the autonomous problem (with A , B , y , g of appropriate dimensions)

$$y' = Ay + yB + g,$$

because one has to approximate the solution

$$\exp \begin{bmatrix} tA & tg \\ 0 & -tB \end{bmatrix} = \begin{bmatrix} e^{tA} & G_t \\ 0 & e^{-tB} \end{bmatrix},$$

where

$$G_t = \sum_{n=0}^{\infty} \frac{t^{n+1}}{(n+1)!} \sum_{k=0}^n (-1)^{n-k} A^k g B^{n-k}.$$

The solution is given by

$$y(t) = H_t(A, B, g, y_0) \equiv (e^{tA} y_0 + G_t) e^{tB},$$

where, depending on the problem, one has to look for different numerical algorithms to approximate the map H_t acting on the initial conditions. The methods based on the Magnus series expansion give the constant matrix to be exponentiated or, equivalently, provide us the constant matrices to be used in the map H_t at each time step.

4 Linear BVP: The imbedding formulation

Let us now consider the linear two-point boundary value problem of the form ($t_0 = 0$)

$$\mathbf{y}'(t) = M(t)\mathbf{y}(t) + \mathbf{h}(t), \quad 0 \leq t \leq T, \quad (24)$$

with separated boundary conditions

$$B_0 \mathbf{y}(0) = \gamma_1, \quad B_1 \mathbf{y}(T) = \gamma_2. \quad (25)$$

Here $\mathbf{y}(t), \mathbf{h}(t) \in \mathbb{C}^n$, $M(t) \in \mathbb{C}^{n \times n}$, $B_0 \in \mathbb{C}^{p \times n}$, $B_1 \in \mathbb{C}^{q \times n}$, $\gamma_1 \in \mathbb{C}^p$, $\gamma_2 \in \mathbb{C}^q$, with $p + q = n$, and we assume that $\text{rank}(B_0) = p$ and $\text{rank}(B_1) = q$.

The goal of the imbedding formulation is to transform a stiff BVP into a well conditioned nonstiff IVP, allowing to use efficient methods for IVPs. We analyse this technique to find how the exponential integrators can be used in this new problem, in order to build efficient numerical schemes. It is convenient to consider separately the cases $q \leq p$ and $p < q$ in order to build methods which only require to solve $r = \min\{q, p\}$ initial value problems (IVPs) for the homogeneous problem ($\mathbf{h} = \mathbf{0}$) and two IVPs of non-homogeneous linear equations of dimensions p and q , respectively. Notice that the cases $p = 0$ or $q = 0$ correspond to initial and final value problems respectively.

We adapt the exponential integrators for the integration of (24)-(25), but in the imbedding formulation where the two-point boundary value problem is replaced by a set of initial value problems, and it is closely related to the shooting method, but it is adapted for stiff problems. This procedure requires to numerically solve a matrix Riccati differential equation. However, in this problem, the RDE appears coupled with other equations and then, the previous methods can not be used in a straightforward way so, an efficient algorithm for solving the whole system requires an special care for the numerical integration of this RDE.

Let us write the two-point boundary value problem (24)-(25) in the following form

$$\mathbf{y}' \equiv \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}' = \begin{bmatrix} A(t) & B(t) \\ C(t) & D(t) \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{f}_1(t) \\ \mathbf{f}_2(t) \end{bmatrix}, \quad 0 \leq t \leq T, \quad (26)$$

with separated boundary conditions

$$\begin{bmatrix} K_{11} & K_{12} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1(0) \\ \mathbf{y}_2(0) \end{bmatrix} = \gamma_1, \quad \begin{bmatrix} K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1(T) \\ \mathbf{y}_2(T) \end{bmatrix} = \gamma_2. \quad (27)$$

Here $A(t) \in \mathbb{C}^{p \times p}$, $B(t) \in \mathbb{C}^{p \times q}$, $C(t) \in \mathbb{C}^{q \times p}$, $D(t) \in \mathbb{C}^{q \times q}$, whereas $\mathbf{y}_1(t)$, $\mathbf{f}_1(t)$, $\gamma_1 \in \mathbb{C}^p$, $\mathbf{y}_2(t)$, $\mathbf{f}_2(t)$, $\gamma_2 \in \mathbb{C}^q$ and $K_{11} \in \mathbb{C}^{p \times p}$, $K_{12} \in \mathbb{C}^{p \times q}$, $K_{21} \in \mathbb{C}^{q \times p}$, $K_{22} \in \mathbb{C}^{q \times q}$.

We consider the case where K_{11} and/or K_{22} are non-singular. If this was not the case, since $\text{rank}(B_0) = p$ and $\text{rank}(B_1) = q$ in (25), with appropriate reordering of the equations (accomplished by the application of permutation matrices) we can transform (26) with new matrices, say $\tilde{A}, \tilde{B}, \dots$, such that the new matrices \tilde{K}_{11} or \tilde{K}_{22} are non-singular.

Let us consider the problem (26)-(27) with $q \leq p$. We next introduce the time

dependent change of variables $\mathbf{y} = Z(t)\mathbf{w}$ with $\mathbf{w} = [\mathbf{w}_1 \ \mathbf{w}_2]^T$ and

$$Z(t) = \begin{bmatrix} I_p & X(t) \\ 0 & I_q \end{bmatrix}, \quad Z^{-1}(t) = \begin{bmatrix} I_p & -X(t) \\ 0 & I_q \end{bmatrix}, \quad (28)$$

where $X(t) \in \mathbb{C}^{p \times q}$. It is immediate to see that the original boundary value problem can be solved as follows [36]:

(I) solve, from $t = 0$ to $t = T$, the initial value problems

$$X'(t) = B(t) + A(t)X(t) - X(t)D(t) - X(t)C(t)X(t), \quad X(0) = -K_{11}^{-1}K_{12}. \quad (29)$$

$$\mathbf{w}_1'(t) = [A(t) - X(t)C(t)]\mathbf{w}_1(t) - X(t)\mathbf{f}_2(t) + \mathbf{f}_1(t), \quad \mathbf{w}_1(0) = K_{11}^{-1}\gamma_1. \quad (30)$$

(II) To solve from $t = T$ to $t = 0$, the equation

$$\mathbf{w}_2'(t) = [D(t) + C(t)X(t)]\mathbf{w}_2(t) + C(t)\mathbf{w}_1(t) + \mathbf{f}_2(t), \quad (31)$$

with the starter final condition obtained from

$$[K_{21}X(T) + K_{22}]\mathbf{w}_2(T) + K_{21}\mathbf{w}_1(T) = \gamma_2.$$

(III) To recover $\mathbf{y}(t) = Z(t)\mathbf{w}(t)$, with $Z(t)$ given by (28).

The case $p < q$ is similar to the previous case. We introduce the time dependent change of variables $\mathbf{y} = \tilde{Z}\tilde{\mathbf{w}}$ where

$$\tilde{Z}(t) = \begin{bmatrix} I_p & 0 \\ \tilde{X}(t) & I_q \end{bmatrix}, \quad \tilde{Z}^{-1}(t) = \begin{bmatrix} I_p & 0 \\ -\tilde{X}(t) & I_q \end{bmatrix},$$

with $\tilde{X}(t) \in \mathbb{C}^{q \times p}$, and we proceed in the same way, obtaining a similar RDE for \tilde{Z} , but with final conditions.

From the analytical point of view, this problem is clear and well understood, see [3,20,36]. However, to build efficient numerical integrators is a more involved task, see [21]. Remember that $X(t)$ is not well defined for those values of t where the matrix $[\gamma(t)X_0 + \delta(t)]$ appearing in (13) has no inverse, and one can expect large errors around the singularities which then disappear once we cross them. However, because the RDE is coupled with the equations for \mathbf{w}_1 and \mathbf{w}_2 , a large error for $X(t)$ at a given instant will cause large errors in \mathbf{w}_1 and \mathbf{w}_2 which are then propagated, leading unacceptable errors for the solutions \mathbf{y}_1 and \mathbf{y}_2 . This problem can be solved by covering the interval $[0, T]$ by a finite set of intervals, $t_0 < \tau_1 < \tau_2 < \dots < \tau_L = T$, where the problem can be reformulated on each time interval, $t \in [\tau_k, \tau_{k+1}]$, with appropriate permutation matrices (see [36] for more details). This implies that at each time

interval one has to solve a permuted linear differential equation such that the inverse of the matrix appearing in (13) is far from being singular.

Notice that to solve the matrix differential equation is equivalent to solve the homogeneous equation associated to (26) q times with different initial conditions. These IVPs are solved in tandem with a non-homogeneous equation of dimension p for \mathbf{w}_1 and a non-homogeneous equation of dimension q for \mathbf{w}_2 . If $p < q$ the problem to solve is similar but the RDE has to be integrated backward in time using final conditions. The equations for $\tilde{\mathbf{w}}_1$ and $\tilde{\mathbf{w}}_2$ have also similar complexity. Then, a numerical method for the RDE has to take into account some of the following computational aspects.

- First, to solve eq. (29) we have to compute $A(t), B(t), C(t), D(t)$ at a given mesh. Some of the computed matrices can also be used in the numerical solutions of eqs. (30) and (31), so one has to decide if it is convenient to store them or it is more appropriate to compute them again when required.
- We obtain numerical approximations to $X(t)$ in a mesh, X_0, X_1, \dots, X_N , which have to be stored for its use in eqs. (30) and (31).
- Next, we have to solve the linear non-homogeneous equation (31) making use of the approximate values X_i at a given mesh and $\mathbf{w}_{1,j}$ at a usually different mesh. To use high order methods for solving this equation requires, in general, to consider approximations to $X(t)$ and $\mathbf{w}_1(t)$ from X_i and $\mathbf{w}_{1,j}$ by interpolation, with a loss of accuracy.
- Finally, we have to obtain approximations to the solution $\mathbf{y}(t) = Z(t) \mathbf{w}(t)$ from $X_i, \mathbf{w}_{1,j}, \mathbf{w}_{2,k}$ obtained at possibly different mesh points.

For these reasons and from the computational point of view, we found that high order methods for the numerical integration of the equations for X, \mathbf{w}_1 and \mathbf{w}_2 can lead into very costly and involved schemes. Then, we consider more convenient to build just a symmetric second order exponential integrator instead of, say, higher order Magnus integrators. Notice that variable time steps can be used as shown in Appendix B. If we consider approximations to the solutions $\mathbf{y}_1(t), \mathbf{y}_2(t)$ using different time steps, the extrapolation technique can be easily adapted, allowing to get higher order and very efficient methods.

4.1 Second order exponential integrators for linear systems

Let us consider the non-homogeneous problem

$$Y' = M(t)Y + F(t), \quad Y(0) = I_n$$

with formal solution for one time step

$$Y(t+h) = \Phi(t+h, t)Y(t) + \int_t^{t+h} \Phi(t+h, \tau)F(\tau) d\tau.$$

We can reformulate this problem as in (15) with $N = 0$, then to use the second order approximation given by $\tilde{\Omega}^{[2]}$ where $M^{(0)}$ and $F^{(0)}$ are approximated by a symmetric second order quadrature rule, and to exponentiate this matrix to obtain a second order symmetric method. Alternatively, we can also proceed as follows which avoids to use the φ -function. If we consider the trapezoidal rule to approximate the integral, we approximate Φ by $\exp(M^{(0)})$ and use the trapezoidal rule to approximate $M^{(0)}$, we obtain the following symmetric second order exponential integrator

$$Y_{n+1} = \exp\left(\frac{h}{2}(M_{n+1} + M_n)\right) \left(Y_n + \frac{h}{2}F_n\right) + \frac{h}{2}F_{n+1}, \quad (32)$$

where $F_i = F(t_i)$, $M_i = M(t_i)$, $Y_i \simeq Y(t_i)$. To prove that the time-symmetry property is preserved, we simply need to solve for Y_n in (32), i.e.

$$Y_n = \exp\left(-\frac{h}{2}(M_{n+1} + M_n)\right) \left(Y_{n+1} - \frac{h}{2}F_{n+1}\right) - \frac{h}{2}F_n,$$

so, it is equivalent to change in (32) t_{n+1}, Y_{n+1} by t_n, Y_n and h by $-h$.

Notice that if the midpoint rule is used then, one has to compute an additional fundamental matrix, $\Phi(t_n + h/2, t_n)$, whose approximation up to the required order while preserving time-symmetry, can increase considerably the computational cost of the algorithm. This can be a serious drawback when higher order methods or higher order quadrature rules are considered, see [6].

We can apply this scheme for the numerical integration of X, \mathbf{w}_1 and \mathbf{w}_2 , obtaining symmetric second order approximations, $\mathbf{y}_{1,n} = \mathbf{y}_1(t_n) + \mathcal{O}(h^3)$, $\mathbf{y}_{2,n} = \mathbf{y}_2(t_n) + \mathcal{O}(h^3)$. Since the scheme is time-symmetric, we consider more convenient to rise the orders of accuracy by extrapolation, applied to the approximated solutions $\mathbf{y}_{1,n}, \mathbf{y}_{2,n}$.

REMARK: As previously mentioned, exponential integrators preserve several qualitative properties. In general, the extrapolation technique does not preserve the group properties, see [8,16] (some qualitative properties are not exactly preserved). This drawback can be diminished if the extrapolation is only used for the output.

5 Numerical examples

To illustrate the performance of the new exponential integrators as well as their implementation, we consider several simple examples which intend to show the good properties of the methods and under which circumstances they can show a high performance.

Example 1. Let us consider the non-homogeneous linear equation

$$\mathbf{y}'(t) = M(t)\mathbf{y}(t) + \mathbf{h}(t), \quad (33)$$

with $\mathbf{y}(t), \mathbf{h}(t) \in \mathbb{C}^n$, $M(t) \in \mathbb{C}^{n \times n}$. We consider the matrix $M(t)$ to be skew-symmetric (the fundamental matrix solution associated to the homogeneous equations is orthogonal for all t). In particular, we choose for the upper triangular elements of $M(t)$:

$$M_{ij} = \log \left(1 + t \frac{j-i}{j+i} \right), \quad 1 \leq i < j \leq n$$

and for the non-homogeneous term we take

$$\mathbf{h}_i = \frac{i \alpha}{i + \alpha t^2}, \quad 1 \leq i \leq n.$$

We have considered this problem for different values of the parameter α , the dimension n and time integrations, i.e., $t \in [0, T]$ for different values of T .

We analyze the performance of the second, fourth and sixth-order Magnus integrators (20)-(22) as well as the hybrid method (23) when the integrals are computed using the Gaussian quadrature rules of the same order of the method. We compare the results with the implicit Gauss-Legendre RK methods of the same order which require to compute $M(t)$ and $\mathbf{h}(t)$ at the same nodes (unless the scheme (23) which requires some few additional evaluations of $\mathbf{h}(t)$).

In particular, we take $n = 5$ and integrate for the interval $t \in [0, 10]$ and we compute the fundamental matrix solution associated to the reformulated homogeneous problem (17) which should indicate the performance of the methods when used either for an IVP as well as for a BVP. The exact solution is computed numerically to round off accuracy using a high-order method and a sufficiently small time step.

In the first experiment, we compare the second order Magnus integrators given by (20) and (23) with the implicit midpoint RK method (the second-order RKGL method). In order to analyze the contribution of the non-homogeneous term, we repeat the experiments for the values $\alpha = 1$ and $\alpha = 100$. The results are shown in Fig. 1, where the corresponding relative error norm for the fundamental matrix is computed for different values of the time step h , $\|\Phi(T, 0) - \Phi_h(T, 0)\|/\|\Phi(T, 0)\|$, where Φ_h is the approximation to the fundamental matrix solution obtained by the numerical schemes when using a time step, h . The results are plotted in a double logarithmic scale. We observe a significant superiority of the exponential integrators which deal very accurately with the oscillatory terms originated by the matrix $M(t)$ (more

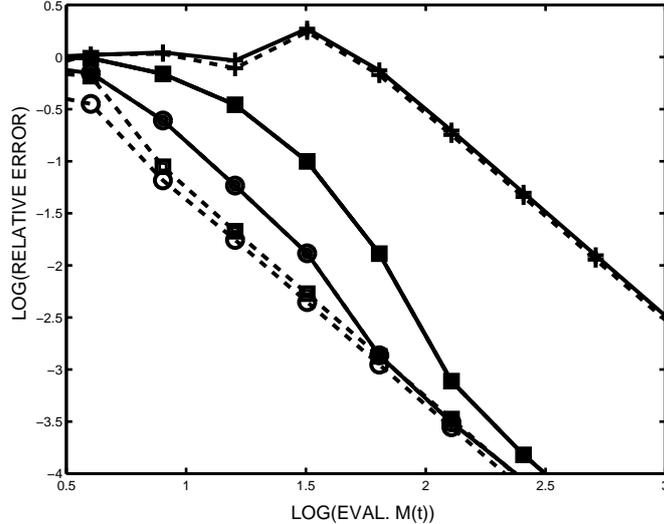


Fig. 1. Relative error in the fundamental matrix solution associated to the homogeneous equation (33) versus the number of evaluations of the matrix $M(t)$ for $n = 5$ and final time $T = 10$, obtained using different time steps, in double logarithmic scale. The following schemes are considered: the second order Magnus integrators (20) (squares) and (23) (circles), and the second-order RKGL method (crosses) for $\alpha = 1$ (dashed lines) and $\alpha = 100$ (solid lines).

than ten times faster). However, this performance deteriorates when the non-homogeneous term increases, and this can be solved using the scheme (23) which incorporates additional terms for this non-homogeneous part (notice that higher order terms for this non-homogeneous part can also be considered with not a significant increase of the computational cost).

Next, we take $\alpha = 1$ and compare the performance of the schemes (20)-(22). The results are shown in Fig. 2. The superiority of the exponential integrators is clearly manifest. Additional numerical experiments, not reported in this work, show that this superiority increases both by increasing the final time, T , as well as by increasing the dimension n , for this problem. The numerical experiments are carried out with constant time steps because the implicit RKGL methods are not appropriately designed to be used with variable time step, contrarily to the exponential integrators, as reported in Appendix B.

Example 2. To analyze the performance of the exponential integrators in the imbedding formulation we consider the simple equation

$$x'' - 2x' + x = t(e^t - 1), \quad 0 \leq t \leq 4 \quad (34)$$

with boundary conditions $x(0) = 0$, $x(4) = 6(e^4 - 1)$. The solution of this problem is given by $x(t) = \left(\frac{1}{6}t^3 - \frac{5}{3}t + 2\right)e^t - t - 2$. To measure the performance of the numerical methods, we consider the average error of the approximations

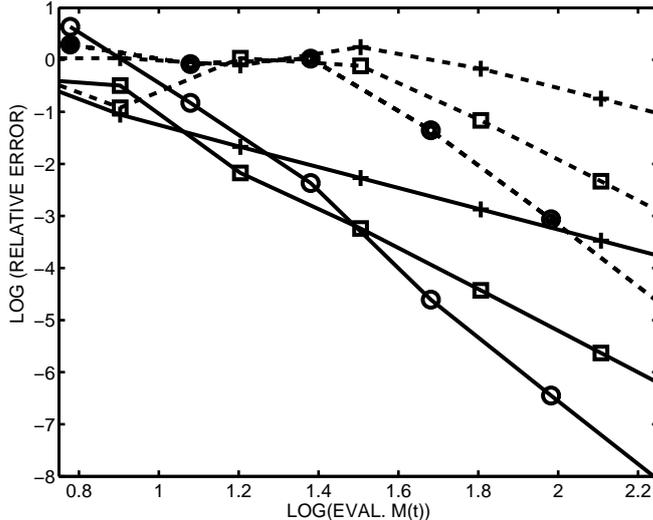


Fig. 2. Same as Fig. 1, for $\alpha = 1$. The following schemes are considered: the Magnus integrators (20)-(22) (solid lines) versus the implicit RKGL methods (dashed lines) for schemes of order two (crosses), of order four (squares), and of order six (circles).

obtained by the methods along the whole interval

$$\mathcal{E} = \left(\frac{1}{N} \sum_{i=1}^N (x(t_i) - x_i)^2 \right)^{1/2}$$

where $h = T/N$, for different values of N . We compare the second order exponential integrator with the second order finite difference method where both require only one new evaluations of the time-dependent functions per step, and we measure the cost of the method by the number of evaluations, N .

Next, we get fourth-order methods by using standard extrapolation. We compute $x_i^{[1]}$ using a time step, $h^{[1]} = T/N$ and we repeat the same integration but using as the time-step, $h^{[2]} = h^{[1]}/2$, to obtain $x_i^{[2]}$. Then, a fourth-order approximation is obtained by $x_i^{[e]} = (4x_{2i}^{[2]} - x_i^{[1]})/3$, $i = 1, \dots, N$. Each step requires three evaluations and then the number of evaluations is $3N$. In Fig. 3 we present the results obtained, in double logarithmic scale, by the second order finite difference method, FD2, the second order exponential integrator, EI2, as well as the fourth-order methods obtained by extrapolation, FD4 and EI4. The slope of the curves indicate the order of the methods. We observe that for this mildly stiff problem the imbedding formulation using exponential methods clearly outperform the finite difference methods.

Example 3. Let us now consider the following problem

$$x'' + \frac{2}{t}x' - \frac{2}{t^2}x = \frac{\sin(\ln(t))}{t^2}, \quad t_0 \leq t \leq t_f \quad (35)$$

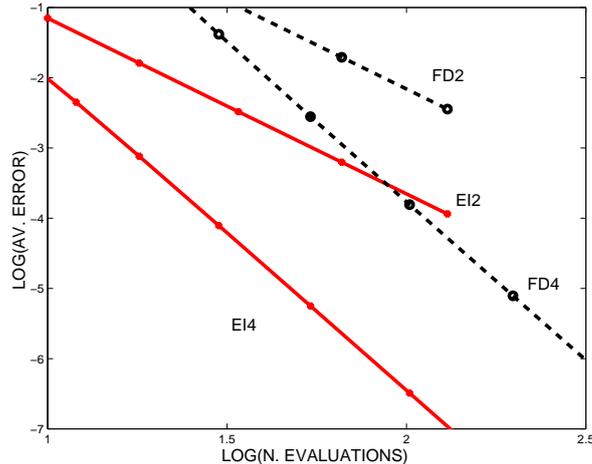


Fig. 3. Averaged error in the solution, $x(t)$ for the equation (34), obtained using different time steps, in double logarithmic scale. FD2 corresponds to the second order finite difference method and EI2 to the second order exponential integrator in the imbedding formulation. FD4 and EI4 are obtained by extrapolation from FD2 and EI2, respectively (they require three evaluations per step).

with boundary conditions $x(t_0) = 1$, $x(t_f) = 1$. The solution of this problem is given by $x(t) = c_1 t + \frac{c_2}{t^2} - \frac{3}{10} \sin(\ln(t)) - \frac{1}{10} \cos(\ln(t))$, with appropriate values for the constants c_1, c_2 . If we take $t_f = 1$ and $0 < t_0 \ll 1$ the problem is stiff. We show that, for this problem, the simple exponential methods, which compute the missing initial values with (8), provide similar or higher performance as the explicit embedding scheme. This performance can be further improved with the simple variable step strategy shown in Appendix B. Fig. 4 shows the maximum error for different fourth-order methods versus the number of function evaluations, in double logarithmic scale, for $t_0 = e^{-\pi}$. In Fig. 5 we show the solution, which is stiff near the initial time, and the time steps chosen by the variable time step scheme along the integration for a safety constant factor $\alpha = 0.95$ and a tolerance $T_I = 10^{-5}$.

6 Conclusions

We have considered exponential integrators based on the Magnus series expansion for the numerical integration of a general linear initial value problem or a two point boundary value problem. Exponential integrators have recently been developed and analyzed extensively and have shown to be very efficient for many problems like stiff or highly oscillatory systems. The exponential integrators, usually designed for the homogeneous linear problem, are used after the problem is reformulated as a particular case of a matrix Riccati differential equation and the Möbius transformation is considered, and new hybrid methods are obtained.

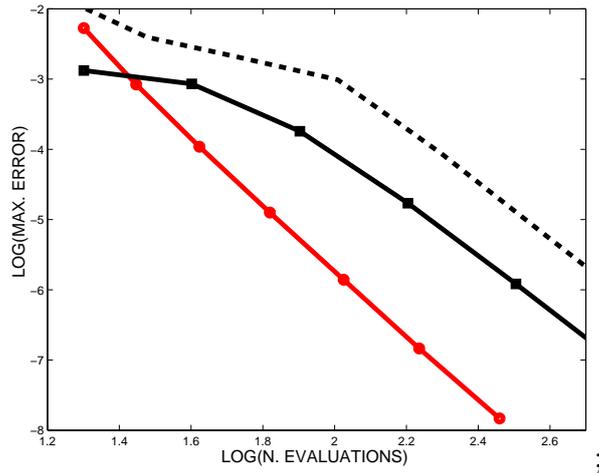


Fig. 4. Averaged error in the solution, $x(t)$ for the equation (35) with $t_0 = e^{-\pi}$ for the following fourth-order exponential methods: the imbedding method with extrapolation (dashed line), the Magnus integrator (21) (solid line with squares) and the Magnus integrator (21) with variable time steps according to Appendix B (solid line with circles).

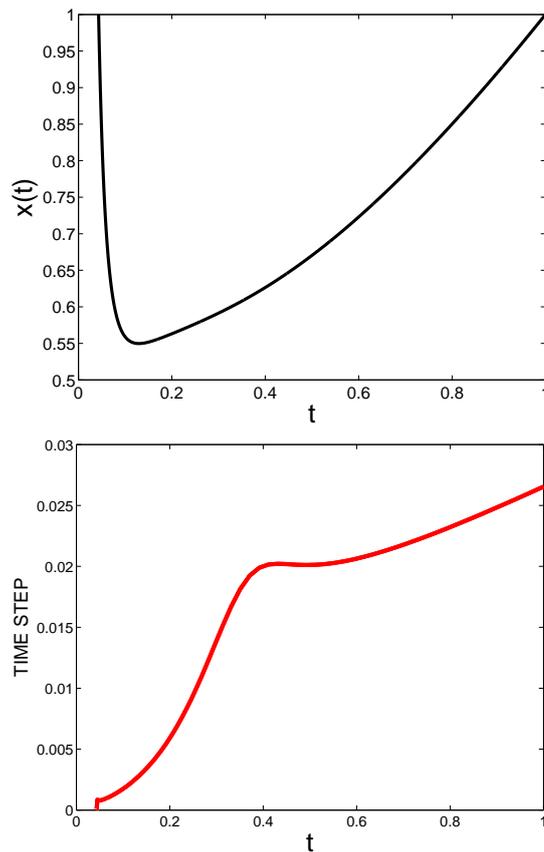


Fig. 5. Above: Solution, $x(t)$, for the equation (35) with $t_0 = e^{-\pi}$, and Below: Time steps used by the variable time-step fourth-order exponential integrator, for solving the equation (35) with $t_0 = e^{-\pi}$ and a tolerance $T_I = 10^{-5}$.

To approximate the exponentials by, say, Padé approximations can be considered equivalent to implicit methods, and then exponential methods share the same advantages as implicit methods. Then, exponential methods are explicit methods which can be used for stiff problems. We have presented in detail the Magnus integrators, showing they can be easily implemented with variable time steps. In addition, they can be easily adapted and optimized to the structure of the problem, depending if some of the time-dependent functions evolve smoothly or not.

Finally, the imbedding formulation or Riccati method has also been considered for the BVP with separated boundary conditions. It requires to numerically integrate forward and backward three IVPs: a matrix Riccati differential equation and two non-homogeneous equations. Since the matrix RDE can be reformulated as a linear equation, we can use exponential integrators for linear problems. We have also presented a symmetric second order exponential integrator which approximates the whole set of equations. This method can be easily used with extrapolation to get higher order methods.

It is also important to remark that in this work we have presented a procedure to numerically integrate Riccati differential equations which are coupled with other linear equations, where standard methods for the RDE are not efficient. Riccati differential equations coupled with other linear equations appear, for instance, in linear quadratic optimal control, e.g. [27], as well as in differential games, see [23], and the algorithm presented in this work can be easily adapted to numerically solve these problems.

We have shown that for some stiff problems, exponential methods used as shooting methods to compute the missing initial or final conditions can lead to more efficient and simple algorithms than in the imbedding formulation. There are other problems, however, where this situation is just the reverse. This is the case, for example, for

$$x'' - (1 + t^2)x = 0, \quad x(0) = 1, \quad x(t_f) = 0.$$

The solution of this problem is given by $x(t) = e^{t^2/2}(1 - \operatorname{erf}(t)/\operatorname{erf}(t_f))$. The solution is relatively smooth but, for $t_f > 8$ the equations to obtain the missing initial conditions in (8) are badly conditioned. The imbedding formulation can be used, however, without problem for larger values of t_f . There exist the question if, with a previous transformation (e.g. the interaction picture or the adiabatic picture) the badly conditioned algebraic equations transform into well conditioned equations and then one can use the exponential integrators in the transformed system. This analysis could be of great interest for solving Sturm-Liouville problems with exponential integrators or to compute certain special functions which are solutions of non-autonomous linear second scalar differential equations [25]. It is also of interest the analysis of the most appropriate exponential methods for solving non-linear IVPs or BVPs using the

quasi linearization technique. The nonlinear equation

$$y' = f(t, y)$$

with $y(t) \in \mathbb{R}^n$ can be solved iteratively by solving the following linear non-homogeneous equations

$$y'_{n+1} + A(t)y_{n+1} = g(t)$$

with

$$A(t) = - \left(\frac{\partial f}{\partial y} \right)_{y_n}, \quad g(t) = f(t, y_n) - \left(\frac{\partial f}{\partial y} \right)_{y_n} y_n$$

which has to be started with an initial guess, $y_0(t)$ satisfying the initial or boundary conditions. For boundary value problems, the boundary conditions can be easily adjusted (see Refs. [3,44]). This iteration converges to a solution $\tilde{y}(t)$ which is an approximation to the exact solution, $y(t)$, of second order.

Acknowledgements

The authors acknowledge the support of the Generalitat Valenciana through the project GV/2009/032. The work of SB has also been partially supported by Ministerio de Ciencia e Innovación (Spain) under the coordinated project MTM2010-18246-C03 (co-financed by the ERDF of the European Union) and the work of EP has also been partially supported by Ministerio de Ciencia e Innovación of Spain, by the project MTM2009-08587.

A Relation between explicit exponential integrators and implicit methods

Implicit methods are frequently used for problems with rapid evolutions or stiff problems. They are usually more costly than explicit methods but due to their stability properties allow for large time steps and can compensate this extra cost. This is an important property because the storage requirements can play an important role in this problem.

Let us consider Runge–Kutta (RK) methods for solving the linear equation (6) which can be rewritten using the Möbius scheme as (17), or shortly as in (18). The application of a s -stage RK method characterized by the real numbers a_{ij} , b_i ($i, j = 1, \dots, s$) and $c_i = \sum_{j=1}^s a_{ij}$, for the integration step

$t_n \mapsto t_{n+1} = t_n + h$ is given by

$$\mathbf{x}_{n+1} = \left(I_d + h \begin{pmatrix} b_1 A_1 & \cdots & b_s A_s \end{pmatrix} (I_{sd} - h\tilde{A})^{-1} \mathbb{I}_{sd \times d} \right) \mathbf{x}_n,$$

where $A_i = A(t_n + c_i h)$, \tilde{A} is a block-matrix with $\tilde{A}_{ij} = a_{ij} A_j$, $\mathbb{I}_{sd \times d} = (I_d, I_d, \dots, I_d)^T$ and I_d is the $d \times d$ identity matrix. If the method is implicit, this approximation can be considered as a rational approximation to the formal solution given by, say, the Magnus series expansion (14), as we will see in more detail later in a simple example.

REMARK: It is important to notice that the algorithms provided by the RK methods remain unaltered when they are used either to solve the equation (6) or to solve (18) (RK methods are linear methods) but this is not the case for the Lie group integrators that we are considering in this work. This property gives to the Lie group methods an important versatility which has not been previously explored.

Let us consider, for instance, the following symmetric implicit second order RK method

$$\frac{1}{2h}(\mathbf{x}_{n+1} - \mathbf{x}_{n-1}) = A_n \frac{1}{2}(\mathbf{x}_{n+1} + \mathbf{x}_{n-1}) \Rightarrow \mathbf{x}_{n+1} = (I + hA_n)(I - hA_n)^{-1} \mathbf{x}_{n-1},$$

or equivalently

$$\mathbf{x}_{n+1} = \left(I + \frac{h}{2} A_{n+1/2} \right) \left(I - \frac{h}{2} A_{n+1/2} \right)^{-1} \mathbf{x}_n,$$

where $A_{n+1/2} = A(t_n + h/2)$. It is immediate to identify that

$$\exp(hA_{n+1/2}) = \left(I + \frac{h}{2} A_{n+1/2} \right) \left(I - \frac{h}{2} A_{n+1/2} \right)^{-1} + \mathcal{O}(h^3),$$

i.e. this is just a second order symmetric diagonal Padé approximation to the exponential, where $hA_{n+1/2}$ corresponds to the midpoint rule which is also a second order symmetric approximation to the integral, say, $A^{(0)}$. For instance, a fourth-order approximation to the exponential by a diagonal Padé approximation to the exponential is given by

$$\begin{aligned} & \exp(hA_{n+1/2}) \\ &= \left(I + \frac{h}{2} A_{n+1/2} + \frac{h^2}{12} A_{n+1/2}^2 \right) \left(I - \frac{h}{2} A_{n+1/2} + \frac{h^2}{12} A_{n+1/2}^2 \right)^{-1} + \mathcal{O}(h^5), \end{aligned}$$

which can be considered as a slightly more elaborated implicit second order method. For some problems, this approximation provides similar results as

the fourth-order implicit Runge-Kutta-Gauss-Legendre method, while requiring only one evaluation of $A(t)$ per step. Among all Padé approximants to the exponential, it is well known that, in general, only the diagonal ones preserve the structure of J -orthogonal Lie group. By increasing slightly the computational cost in the approximation to the exponential, more accurate results can be obtained in many cases. This can allow to use large time steps (with a reduced number of evaluations of the matrix $A(t)$) and to store a reduced number of intermediate stages. This is one of the main reasons exponential methods can lead to very accurate results and can also be used with relatively large time steps, like implicit methods.

B Variable step size implementation.

The previous Magnus integrators can be easily adapted with a variable step size implementation in different ways. For instance, Magnus methods implemented with variable step size integration algorithms based on a local error estimate can be found in [30]. Alternatively, the local extrapolation procedure can be easily implemented into Magnus based integration schemes. As is well known, in this technique one computes two numerical approximations to the solution, Y_h and \hat{Y}_h , with Y_h being of lower order than \hat{Y}_h . Then the difference $Y_h - \hat{Y}_h$ can be used for the purpose of step size selection when the integration is continued with the higher order approximation. We illustrate the procedure in the context of Magnus with the 2-nd and 4-th order methods. Let us consider the following approximations of order 2 and 4 to the exact solution given by (20) and (21)

$$Y_h = \exp(A^{(0)})Y_0, \quad \hat{Y}_h = \exp(A^{(0)} + [A^{(1)}, A^{(0)}])Y_0.$$

Then an estimate of the difference can be done with the Baker-Campbell-Hausdorff formula:

$$\begin{aligned} \hat{Y}_h - Y_h &= \left(I - \exp(A^{(0)}) \exp(-A^{(0)} - [A^{(1)}, A^{(0)}]) \right) \hat{Y}_h \\ &= \left(I - \exp([A^{(0)}, A^{(1)}] - \frac{1}{2}[A^{(0)}, [A^{(1)}, A^{(0)}]] + O(h^6)) \right) \hat{Y}_h \\ &= ([A^{(1)}, A^{(0)}] - \frac{1}{2}[A^{(0)}, [A^{(1)}, A^{(0)}]]) \hat{Y}_h + O(h^6), \end{aligned}$$

so that

$$E_r \equiv \|\hat{Y}_h - Y_h\| \sim \left\| \left([A^{(0)}, A^{(1)}] + \frac{1}{2}[A^{(0)}, A^{(1)}] \right) \cdot \hat{Y}_h \right\|.$$

Similar error terms can be obtained when using sixth- or eighth-order methods,

see [9]. Then, the step size to be employed for a method of order p is given by

$$h_{i+1} = \alpha \left(\frac{T_I}{E_r} \right)^{1/(p-1)} h_i,$$

where T_I is the prescribed tolerance and α is a safety constant factor.

When solving a differential equation with some particular structure or with stability difficulties, it is convenient to make a preliminary transformation previous to the use of the numerical schemes. This transformation can, however, change the structure of the equations and then can restrict the numerical methods which can be used. This is clearly manifest for the linear two point boundary value problem with separated boundary conditions in the imbedding formulation which we now present in some detail as a representative case.

References

- [1] ABOU-KANDIL, H., FREILING, G., IONESCU, V. & JANK, G. (2003) *Matrix Riccati equations in control and systems theory*. Virkāuser Verlag, Basel.
- [2] AL-MOHY, A. H. & HIGHAM, N. J. (2011) *Computing the Action of the Matrix Exponential, with an Application to Exponential Integrators*. *SIAM J. Sci. Comp.* **33**, 488–511.
- [3] ASCHER, U. M., MATTHEIJ, R. M. & RUSSELL, R. D. (1988) *Numerical solutions of boundary value problems for ordinary differential equations*. Prentice-Hall. Englewood Cliffs, NJ.
- [4] BAGHI, A. (1984) *Stackelberg differential games in economic models*. Springer Lectures Notes in Control and Information Sciences. Springer. Berlin.
- [5] BITTANTI, S., LAUB A. J. & WILLEMS, J. C. (1991) *The Riccati equation*. Springer.
- [6] BLANES, S. & CASAS, F. (2003) Optimization of Lie group methods for differential equations. *Future Gen. Comput. Sys.*, **19**, 331–339.
- [7] BLANES, S., CASAS, F., OTEO, J. A. & ROS, J. (2009) The Magnus expansion and some of its applications. *Physics Reports*, **470**, 151–238.
- [8] BLANES, S., CASAS, F. & ROS, J. (1999) Extrapolation of symplectic integrators. *Celest. Mech. & Dyn. Astr.*, **75**, 149–161.
- [9] BLANES, S., CASAS, F. & ROS, J. (2000) Improved high order integrators based on the Magnus expansion. *BIT*, **40**, 434–450.
- [10] BLANES, S., CASAS, F. & ROS, J. (2002) High order optimized geometric integrators for linear differential equations. *BIT*, **42**, 262–284.

- [11] BLANES, S., JÓDAR, L. & PONSODA, E. (2000) Approximate solutions with a priori error bounds for continuous coefficient matrix Riccati equations. *Math. Comp. Mod.*, **31**, 1–15.
- [12] BLANES, S. & MOAN, P. C. (2006) Fourth- and sixth-order commutator-free Magnus integrators for linear and non-linear dynamical systems. *Appl. Numer. Math.*, **56**, 1519–1537.
- [13] CASAS, F. (2007) Sufficient conditions for the convergence of the Magnus expansion. *J. Phys. A: Math. Theor.*, **40**, 15001–15017.
- [14] CASAS, F. & ISERLES, A. (2006) Explicit Magnus expansions for nonlinear equations. *J. Phys. A: Math. Gen.*, **39**, 5445–5461
- [15] CELLEDONI, E. & ISERLES, A. (2000) Approximating the exponential from a Lie algebra to a Lie group. *Math. Comp.*, **69**, 1457–1480.
- [16] CHAN, R. P. K. & MURUA, A. (2000) Extrapolation of symplectic methods for hamiltonian problems. *Appl. Numer. Math.*, **34**, 189–205.
- [17] CLEMHOUT, S. & WAN JR., H. Y. (1979) Interactive economic dynamics and differential games. *J. Optim. Theory Appl.*, **27**, 7–30.
- [18] CRUZ, J. B. & CHEN C. I. (1971) Series Nash solution of two person nonzero sum linear quadratic games. *J. Optim. Theory Appl.*, **7**, 240–257.
- [19] DIECI, L. (1992) Numerical integration of the differential Riccati equation and some related issues. *SIAM J. Numer. Anal.*, **29**, 781–815.
- [20] DIECI, L., OSBORNE, M. R. & RUSSELL, R. D. (1988a) A Riccati transformation method for solving linear BVP I: Theoretical aspects. *SIAM J. Numer. Anal.*, **25**, 1055–1073.
- [21] DIECI, L., OSBORNE, M. R. & RUSSELL, R. D. (1988b) A Riccati transformation method for solving linear BVP II: Computational aspects. *SIAM J. Numer. Anal.*, **25**, 1074–1092.
- [22] DIELE, F., LOPEZ, L. & PELUSO, R. (1998) A Riccati transformation method for solving linear BVP II: Computational aspects. *Adv. Comp. Math.*, **8**, 317–334.
- [23] ENGWERDA, J. (2005) *LQ Dynamic optimization and differential games*. Wiley.
- [24] FER, F. (1958) Résolution de l'équation matricielle $\dot{U} = pU$ par produit infini d'exponentielles matricielles. *Bull. Classe Sci. Acad. Roy. Bel.*, **44**, 818–829.
- [25] GIL, A., SEGURA, J. & TEMME, N. M. (2007) Numerical Methods for Special Functions. *SIAM*.
- [26] HOCHBRUCK, M. & OSTERMANN, A. (2010) Exponential integrators. *Acta Numerica*, **19**, 209–286.
- [27] HU, G. D. (2007) Symplectic Runge-Kutta methods for the linear quadratic regulator problem. *J. Math. Anal.*, **1**, 293–304.

- [28] ISERLES, A. (2001a) A Magnus expansion for the equation $Y' = AY - YB$. *J. Comp. Math.*, **19**, 15–26.
- [29] ISERLES, A. (2001b) On Cayley-transform methods for the discretization of Lie-group equations. *Found. Comp. Math.*, **1**, 129–160.
- [30] ISERLES, A., MARTHINSEN, A. & NØRSETT, S. P. (1999) On the implementation of the method of Magnus series for linear differential equations. *BIT*, **39**, 281–304.
- [31] ISERLES, A., MUNTHE-KAAS, H. Z., NØRSETT, S. P. & ZANNA, A. (2000) Lie group methods. *Acta Numerica*, **9**, 215–365.
- [32] ISERLES, A. & NØRSETT, S. P. (1999) On the solution of linear differential equations in Lie groups. *Philosophical Trans. Royal Soc.*, **A 357**, 983–1019.
- [33] JÓDAR, L. & ABOU-KANDIL, H. (1988) A resolution method for Riccati differential systems coupled in their quadratic terms. *SIAM J. Math. Anal.*, **19**, 1225–1230.
- [34] JÓDAR, L. & CORTÉS, J. C. (1998) Rational matrix approximation with a priori error bounds for non-symmetric matrix Riccati equations with analytic coefficients. *IMA J. Numer. Anal.*, **18**, 545–561.
- [35] JÓDAR, L. & PONSODA, E. (1995) Non-autonomous Riccati-type matrix differential equations: existence interval, construction of continuous numerical solutions and error bounds. *IMA J. Numer. Anal.*, **15**, 61–74.
- [36] KELLER, H. B. & LENTINI, M. (1982) Invariant imbedding, the box scheme and an equivalence between them. *SIAM J. Numer. Anal.*, **19**, 942–962.
- [37] KELLER, H. B. (1976) *Numerical solution of two point boundary value problems*. CBMS-NSF Regional Conference Series in Applied Mathematics, **24**, SIAM, Philadelphia.
- [38] KENNEY, S. S. & LEIPNIK, R. B. (1985) Numerical integration of the differential matrix Riccati equation. *IEEE Trans. Aut. Control*, **AC-30**, 962–970.
- [39] MAGNUS, W. (1954) On the exponential solution of differential equations for a linear operator. *Commun. Pure Appl. Math.*, **7**, 649–673.
- [40] MARITON, M. (1990) *Linear systems in automatic control*. New York. Marcel Dekker.
- [41] MOAN, P.C. & NIESEN, J. (2008) Convergence of the Magnus series. *Found. Comput. Math.*, **8**, 291–301.
- [42] MOLER, C. B. & VAN LOAN, C. F. (1978) Nineteen dubious ways to compute the exponential of a matrix *SIAM Review*, **20**, 801–836.
- [43] MOLER, C. B. & VAN LOAN, C. F. (2003) Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, **45**, 3–49.

- [44] NA, T.Y. (1979) *Computational methods in engineering boundary value problems*. Mathematics in science and engineering, **145**, Accademic Press, New York.
- [45] NIESEN, J. & WRIGHT, W. M. (2009) A Krylov subspace algorithm for evaluating the φ -functions appearing in exponential integrators. *Technical Report arXiv:0907.4631*, 20 pp.
- [46] ORTEGA, J. M. (1972) *Mathematical Analysis. A second course*. Academic Press. New York.
- [47] PONSODA, E., BLANES, S. & BADER, P. (2011) New efficient numerical methods to describe the heat transfer in a solid medium. *Math. Comput. Mod.*. In press. doi:10.1016/j.mcm.2010.11.067
- [48] PONSODA, E. & CASTAÑO, J. I. (1999) Approximate solutions and error bounds for matrix differential equations using Fer's factorization. *J. Appl. Sci. Comp.*, **5**, 69–79.
- [49] REID, W. T. (1972) *Riccati differential equations*. New York. Academic.
- [50] SAAD, Y. (1992) Analysis of some Krylov subspace approximations to the matrix exponential operator. *SIAM J. Num. Anal.*, **29**, 209–228.
- [51] SCHIFF, J. & SCHNIDER, S. (1999) A natural approach to the numerical integration of Riccati differential equations. *SIAM J. Sci. Comp.*, **36**, 1392–1413.
- [52] SCOTT, M. R. & WATTS, H. A. (1977) Computational solution of linear 2- point boundary values problems via orthonormalization. *SIAM J. Numer. Anal.*, **14**, 40–70.
- [53] SIDJE, R. B. (1998) Expokit: a software package for computing matrix exponentials. *ACM Trans. Math. Software*, **24**, 130–156.
- [54] STARR, A. W. & HO, Y. C. (1969) Non-zero sum differential games. *J. Optim. Theory Appl.*, **24**, 130–156.