Physics Reports 470 (2009) 151–238

Contents lists available at ScienceDirect

Physics Reports

journal homepage: www.elsevier.com/locate/physrep

The Magnus expansion and some of its applications

S. Blanes [a], F. Casas [b], J.A. Oteo [c], J. Ros [c,d,*]

[a] Instituto de Matemática Multidisciplinar, Universidad Politécnica de Valencia, E-46022 Valencia, Spain
[b] Departament de Matemàtiques, Universitat Jaume I, E-12071 Castellón, Spain
[c] Departament de Física Teòrica, Universitat de València, E-46100 Burjassot, Valencia, Spain
[d] IFIC, Centre Mixt Universitat de València-CSIC, E-46100 Burjassot, Valencia, Spain

ARTICLE INFO

Article history:
Accepted 13 November 2008
Available online 25 November 2008
editor: D.L. Mills

PACS:
03.65
02.30.Hq
02.60.Cb
02.20.Rt

ABSTRACT

Approximate resolution of linear systems of differential equations with varying coefficients is a recurrent problem, shared by a number of scientific and engineering areas, ranging from Quantum Mechanics to Control Theory. When formulated in operator or matrix form, the *Magnus expansion* furnishes an elegant setting to build up approximate exponential representations of the solution of the system. It provides a power series expansion for the corresponding exponent and is sometimes referred to as *Time-Dependent Exponential Perturbation Theory*. Every Magnus approximant corresponds in Perturbation Theory to a partial re-summation of infinite terms with the important additional property of preserving, at any order, certain symmetries of the exact solution.

The goal of this review is threefold. First, to collect a number of developments scattered through half a century of scientific literature on Magnus expansion. They concern the methods for the generation of terms in the expansion, estimates of the radius of convergence of the series, generalizations and related non-perturbative expansions. Second, to provide a bridge with its implementation as generator of *especial purpose numerical integration methods*, a field of intense activity during the last decade. Third, to illustrate with examples the kind of results one can expect from Magnus expansion, in comparison with those from both perturbative schemes and standard numerical integrators. We buttress this issue with a revision of the wide range of physical applications found by Magnus expansion in the literature.

Contents

* Corresponding author at: Departament de Física Teòrica, Universitat de València, E-46100 Burjassot, Valencia, Spain.
E-mail address: Jose.Ros@uv.es (J. Ros).

0370-1573/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.physrep.2008.11.001

1. Introduction

1.1. Motivation, overview and history

The outstanding mathematician Wilhelm Magnus (1907–1990) made important contributions to a wide variety of fields in mathematics and mathematical physics [1]. Among them, one can mention combinatorial group theory [2] and his collaboration in the Bateman project on higher transcendental functions and integral transforms [3]. In this report we review another of his long-lasting constructions: the so-called Magnus expansion (hereafter referred to as ME). ME was introduced as a tool to solve non-autonomous linear differential equations for linear operators. It is interesting to observe that, in his seminal paper of 1954, [4], although it is essentially mathematical in nature, Magnus recognizes that his work was stimulated by results of K.O. Friedrichs on the theory of linear operators in Quantum Mechanics [5]. Furthermore, as the first antecedent of his proposal, he quotes a paper by R.P. Feynman in the Physical Review [6]. We stress these facts to show that already, at its very beginning, ME was strongly related to Physics, and has been ever since, and there is no reason to doubt that it will continue to be. This is the first motivation to offer here a review as the present one.

Magnus' proposal has the very attractive property of leading to approximate solutions which exhibit, at any order of approximation, some qualitative or physical characteristics, which the first principles guarantee for the exact (but unknown) solution of the problem. Important physical examples are the symplectic or unitary character of the evolution operator when dealing with classical or quantum mechanical problems, respectively. This is at variance with most standard perturbation theories, and is apparent when formulated in a correct algebraic setting: Lie algebras and Lie groups. But this great advantage has sometimes been tainted in the past, by the difficulties both in constructing explicitly higher order terms, and in assuring existence and convergence of the expansion.

In our opinion, recent years have witnessed great improvement in this situation. Concerning general questions of existence and convergence, new results have appeared. From the point of view of applications, some new approaches in old fields have been published, while completely new and promising avenues have been opened by the use of the Magnus expansion in Numerical Analysis. It seems reasonable to expect fruitful cross fertilization between these new developments and the most conventional perturbative approach to ME and, from it, further applications and new calculations.

This new scenario makes it desirable for the Physics community in different areas (and scientists and engineers in general) to have access, in as unified a way as possible, to all the information concerning ME which, so far, has been treated in very different settings, and has appeared scattered through very different bibliographic sources.

As implied by the preceding paragraphs, this report is mainly addressed to a Physics audience, or close neighbors, and consequently we shall keep the treatment of its mathematical aspects within reasonable limits and refer the reader to more detailed literature where necessary. By the same token, the applications presented will be limited to examples from Physics, or from the closely related field of Physical Chemistry. We shall also emphasize its instrumental character for numerically solving physical problems.

In the present section, as an introduction, we present a brief overview and sketch a history of more than 50 years of ME. To start with, let us consider the initial value problem associated with the linear ordinary differential equation

$$Y'(t) = A(t)Y(t), \qquad Y(t_0) = Y_0, \tag{1}$$

where, as usual, the prime denotes the derivative with respect to the real independent variable which we take as time t, although much of what will be said also applies to a complex independent variable. In order of increasing complexity, we may consider the equation above in different contexts:

(a) $Y : \mathbb{R} \longrightarrow \mathbb{C}, A : \mathbb{R} \longrightarrow \mathbb{C}$. This means that the unknown Y and the given A are complex scalar valued functions of one real variable. In this case, there is no problem at all: the solution reduces to a quadrature and an ordinary exponential evaluation:

$$Y(t) = \exp\left(\int_{t_0}^{t} A(s)\mathrm{d}s\right) Y_0. \tag{2}$$

(b) $Y : \mathbb{R} \longrightarrow \mathbb{C}^n, A : \mathbb{R} \longrightarrow \mathbb{M}_n(\mathbb{C})$, where $\mathbb{M}_n(\mathbb{C})$ is the set of $n \times n$ complex matrices. Now Y is a complex vector valued function, and A a complex $n \times n$ matrix valued function. At variance with the previous case, only in very special cases is the solution easy to state: when for any pair of values of t, t_1 and t_2, one has $A(t_1)A(t_2) = A(t_2)A(t_1)$, which is certainly the case if A is constant. Then the solution reduces to a quadrature (trivial or not) and a matrix exponential. With the obvious changes in the meaning of the symbols, Eq. (2) still applies. In the general case, however, there is no compact expression for the solution and (2) is no longer the solution.

(c) $Y : \mathbb{R} \longrightarrow \mathbb{M}_n(\mathbb{C}), A : \mathbb{R} \longrightarrow \mathbb{M}_n(\mathbb{C})$. Now both Y and A are complex matrix valued functions. A particular case, but still general enough to encompass the most interesting physical and mathematical applications, corresponds to $Y(t) \in \mathcal{G}$, $A(t) \in \mathfrak{g}$, where $\mathcal{G}$ and $\mathfrak{g}$ are, respectively, a matrix Lie group and its corresponding Lie algebra. Why this is of interest, is easy to grasp: the key reason for the failure of (2) is the non-commutativity of matrices in general. So one can expect that the (in general non-vanishing) commutators play an important role. But when commutators enter the play, one immediately thinks in Lie structures. Furthermore, plainly speaking, the way from a Lie algebra to its Lie group is covered by the exponential operation — a fact that will be of no surprise in this context. The same comments of the previous case are valid here. In this report, we shall mostly deal with this matrix case.

(d) The most general situation one can think of corresponds to $Y(t)$ and $A(t)$ being operators in some space, e.g., Hilbert space in Quantum Mechanics. Perhaps the most paradigmatic example of (1) in this setting is the time-dependent Schrödinger equation.

Observe that case (b) above can be reduced to case (c). This is easily seen if one introduces what, in mathematical literature, is called the matrizant, a concept dating back at least to the beginning of the 20th century in the work of Baker [7]. It is the $n \times n$ matrix $U(t, t_0)$ defined through

$$Y(t) = U(t, t_0)Y_0. \tag{3}$$

Without loss of generality, we will take $t_0 = 0$ unless otherwise explicitly stated, for the sake of simplicity. When no confusion may arise, we write only one argument in U and denote $U(t, 0) \equiv U(t)$, which then satisfies the differential equation and initial condition

$$U'(t) = A(t)U(t), \qquad U(0) = I, \tag{4}$$

where I stands for the n-dimensional identity matrix. The reader will have recognized $U(t)$ as what, in physical terms, is known as the time evolution operator.

We are now ready to state Magnus' proposal: a solution to (4) which is a true matrix exponential

$$U(t) = \exp \Omega(t), \qquad \Omega(0) = O, \tag{5}$$

and a series expansion for the matrix in the exponent

$$\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t), \tag{6}$$

which is what we call the *Magnus expansion*. The mathematical elaborations explained in the next section determine $\Omega_k(t)$. Here we just write down the three first terms of that series:

$$\begin{aligned}
\Omega_1(t) &= \int_0^t A(t_1)\mathrm{d}t_1, \\
\Omega_2(t) &= \frac{1}{2}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\, [A(t_1), A(t_2)] \\
\Omega_3(t) &= \frac{1}{6}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3\, ([A(t_1), [A(t_2), A(t_3)]] + [A(t_3), [A(t_2), A(t_1)]])
\end{aligned} \tag{7}$$

where $[A, B] \equiv AB - BA$ is the matrix commutator of A and B.

The interpretation of these equations seems clear: $\Omega_1(t)$ coincides exactly with the exponent in (2). But this equation cannot give the whole solution, as has already been said. So, if one insists on having an exponential solution, the exponent has to be corrected. The rest of the ME in (6) gives that correction, necessary to keep the exponential form of the solution.

The terms appearing in (7) already suggest the most appealing characteristic of ME. Remember that a matrix Lie algebra is a linear space in which one has defined the commutator as the second internal composition law. If, as we suppose, $A(t)$ belongs to a Lie algebra $\mathfrak{g}$ for all t, so does any sum of multiple integrals of nested commutators. Then, if all terms in ME have a structure similar to that of the ones shown before, the whole $\Omega(t)$ and any approximation to it obtained by truncation of ME, will also belong to the same Lie algebra. In the next section, it will be shown that this turns out to be the case and, *a fortiori*, its exponential will be in the corresponding Lie group.

Why is this so important for physical applications? Just because many of the properties of evolution operators derived from first principles are linked to the fact that they belong to a certain Lie group: e.g. unitary group in Quantum Mechanics, symplectic group in Classical Mechanics. In that way, use of (truncated) ME leads to approximations which share, with the exact solution of Eq. (4), important qualitative (very often, geometric) properties. For instance, in Quantum Mechanics every approximant preserves probability conservation.

From the present point of view, we could say that the last paragraphs summarize, in a nut shell, the main contents of the famous paper of Magnus of 1954. With no exaggeration, its appearance can be considered a turning point in the treatment of the initial value problem defined by (4).

But important as it certainly was, Magnus' paper left some problems, at least partially, open:

- First, for what values of t and for what operators A does Eq. (4) admit a true exponential solution? This, we call the existence problem.
- Second, for what values of t and for what linear operators A does the series in Eq. (6) converge? This we describe as the convergence problem. We want to emphasize that, although related, these are two different problems. To see why, think of the scalar equation $y' = y^2$ with $y(0) = 1$. Its solution $y(t) = (1-t)^{-1}$ exists for $t \neq 1$, but its form in power series $y(t) = \sum_0^\infty t^n$ converges only for $|t| < 1$.
- Third, how does one construct higher order terms $\Omega_k(t)$, $k \geq 3$, in the series? Moreover, is there a closed-form expression for $\Omega_k(t)$?
- Fourth, how does one calculate in an efficient way $\exp \Omega^{[N]}$, where $\Omega^{[N]} \equiv \sum_{k=1}^N \Omega_k(t)$ is a truncation of the ME?

All these questions, and many others, will be dealt with in the rest of the paper. But before entering that analysis, we think it is interesting to present a view, however brief, from the historical perspective of the half-century of developments on the Magnus series. Needless to say, we by no means try to present a detailed and exhaustive chronological account of the many approaches followed by authors from very different disciplines. To minimize duplication with later sections, we simply mention some representative samples, so that the reader can understand the evolution of the field.

Including some precedents, and with a (as undeniable as unavoidable) dose of arbitrariness, we may distinguish four periods in the history of our topic:

(1) Before 1953. The problem which ME solves has a centennial history dating back at least to the work of Peano, by the end of 19th century, and Baker, at the beginning of the 20th (for references to the original papers see e.g. [8]). They combine the theory of differential equations with an algebraic formulation. Intimately related to these treatments from the very beginning, is the study of the so called Baker–Campbell–Hausdorff (or BCH formula for short) [9–11] which gives C in terms of A, B and their multiple nested commutators when expressing $\exp(A)\exp(B)$ as $\exp(C)$. This topic has a long history in its own right, and will also be discussed in Section 2. As one of its early hallmarks, we quote [12]. In Physics literature, the interest in the problem posed by Eq. (4) was very much revived with the advent of Quantum Electrodynamics (QED). The works of Dyson [13] and, in particular, Feynman [6] in the late forties and early fifties, are worth mentioning here.

(2) 1953–1970. We have quoted as the 'birth certificate' of ME, the paper [4] by Magnus in 1954. This is not strictly true: there is a Research Report [14] dated June 1953 which differs from the published paper in the title and in a few minor details, and which should in fact be taken as a preliminary draft of it. In both publications, the result summarized above on ME appears, in almost identical words. The work of Pechukas and Light [15] gave, for the first time, a more specific analysis of the problem of convergence than the rather vague considerations in Magnus' paper. Wei and Norman [16, 17] did the same for the existence problem. Robinson, to the best of our knowledge, seems to have been the first to apply ME to a physical problem [18]. Special mention in this period should be given to a paper by Wilcox [19], in which useful mathematical tools are given, and ME is presented together with other algebraic treatments of Eq. (4), in particular Fer's infinite product expansion [20]. Also worth mentioning here, is the first application of ME as a numerical tool for integrating the time-independent Schrödinger equation for potential scattering, by Chang and Light [21].

(3) 1971–1990. During these years, ME was consolidated on different fronts. It was successfully applied to a wide spectrum of fields in Physics and Chemistry: from atomic [22] and molecular [23] Physics to Nuclear Magnetic Resonance (NMR) [24,25] to Quantum Electrodynamics [26] and elementary particle Physics [27]. A number of case studies also helped to clarify its mathematical structure, see for example [28]. The construction of higher order terms was approached from different angles. The intrinsic and growing complexity of ME allows for different schemes. One which has shown itself very useful in tackling other questions like the convergence problem, was the recurrent scheme by Klarsfeld and Oteo [29].

(4) Since 1991. The last decade of the 20th century witnessed a renewed interest in ME, which still continues now. It has followed different lines. Concerning the basic problems of existence and convergence of ME, there has been definite progress [30–33]. ME has also been adapted for specific types of equations: Floquet theory when $A(t)$ is a periodic function [34], stochastic differential equations [35] or equations of the form $Z' = AZ - ZB$ [36]. Special mention should be made of the new field, open in this most recent period, that uses Magnus' scheme to build novel algorithms [37] for the numerical integration of differential equations within the widest field of geometric integration [38]. After optimization [39,40], these integrators have proved to be highly competitive.

As proof of the persistent impact the 1954 paper by Magnus has had in scientific literature, we present in Figs. 1 and 2, the number of citations per year and the cumulative number of citations, respectively, as December 2007 with data taken from ISI Web of Science. The original paper appears about 750 times of which, roughly, 50, 320 and 380 correspond, respectively, to each of the last three periods we have considered. The enduring interest in that seminal paper is clear from the figures.

The presentation of this report is organized as follows. In the remainder of this section, we include some mathematical tools and notations that will be used time and again in our treatment. In Section 2 we formally introduce the Magnus expansion, study its main features and thoroughly analyze the convergence issue. Next, in Section 3 several generalizations of the Magnus expansion are reviewed, with special emphasis in its application to general nonlinear differential equations. In order to illustrate the main properties of ME, in Section 4 we consider simple examples for which the computations

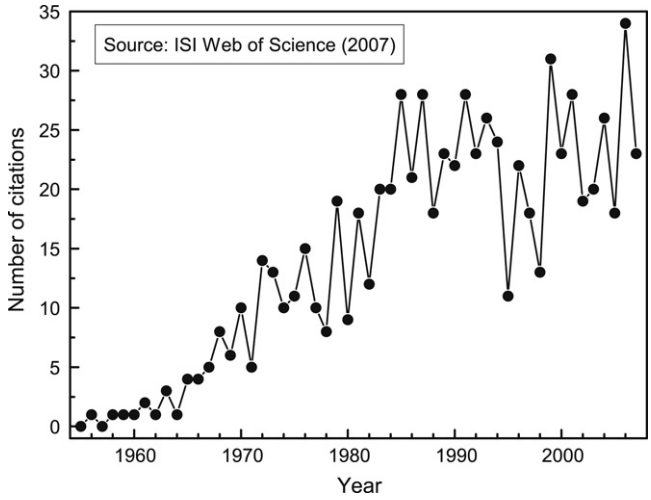


Fig. 1. Persistency of Magnus' original paper: number of citations per year.

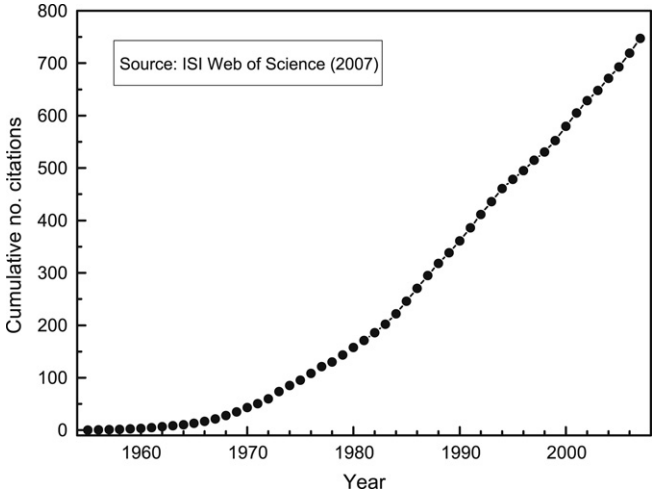


Fig. 2. Persistency of Magnus' original paper: cumulative number of citations.

required are relatively straightforward. Section 5 is devoted to an aspect that has been most recently studied in this setting: the design of new algorithms for the numerical integration of differential equations based on the Magnus expansion. There, after a brief characterization of numerical integrators, we present several methods that are particularly efficient, as shown by the examples considered. Given the relevance of the new numerical schemes, we briefly review in Section 6 some of its applications in different contexts, ranging from boundary-value problems to stochastic differential equations. In Section 7, on the other hand, applications of the ME to significant physical problems are considered. Finally, the paper ends with some concluding remarks.

1.2. Mathematical preliminaries and notations

Here we collect, for the reader's convenience, some mathematical expressions, terminology and notations which appear most frequently in the text. Needless to say, we have made no attempt at being completely rigorous. We just try to facilitate the casual reading of isolated sections.

As already mentioned, the natural mathematical habitat for most of the objects we will deal with in this report is a Lie group, or its associated Lie algebra. Although most of the results discussed in these pages are valid in a more general setting, we will essentially consider only matrix Lie groups and algebras.

By a Lie group $\mathcal{G}$ we understand a set which combines an algebraic structure with a topological one. At the algebraic level, every two elements of $\mathcal{G}$ can be combined by an internal composition law to produce a third element, also in $\mathcal{G}$. The law is required to be associative, to have an identity element, and every element must have an inverse. The ordinary product and the inverse of the invertible matrix, play that role in the cases we are more interested in. The topological exigence forces the composition law and the association of an inverse to be sufficiently smooth functions.

A Lie algebra $\mathfrak{g}$ is a vector space whose elements can be combined by a second law, the Lie bracket, which we represent by $[A, B] = C$, with A, B, C elements of $\mathfrak{g}$, in such a way that the law is bilinear, skew-symmetric and satisfies the well known Jacobi identity,

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \tag{8}$$

When dealing with matrices we take, as Lie bracket, the familiar commutator:

$$[A, B] = AB - BA, \quad A \in \mathfrak{g}, B \in \mathfrak{g}, \tag{9}$$

where AB stands for the usual matrix product. If we consider a finite-dimensional Lie algebra with dimension d and denote by A_i, $i = 1, \ldots, d$, the vectors of one of its bases, then the fundamental brackets one has to know are

$$[A_i, A_j] = c_{ij}^k A_k, \tag{10}$$

where sum over repeated indexes is understood. The coefficients c_{ij}^k are the so-called structure constants of the algebra.

Associated with any $A \in \mathfrak{g}$ we can define a linear operator $\mathrm{ad}_A : \mathfrak{g} \to \mathfrak{g}$ which acts according to

$$\mathrm{ad}_A B = [A, B], \qquad \mathrm{ad}_A^j B = [A, \mathrm{ad}_A^{j-1} B], \qquad \mathrm{ad}_A^0 B = B, \quad j \in \mathbb{N}, B \in \mathfrak{g}. \tag{11}$$

Also of interest, is the exponential of this ad_A operator,

$$\mathrm{Ad}_A = \exp(\mathrm{ad}_A), \tag{12}$$

whose action on $\mathfrak{g}$ is given by

$$\mathrm{Ad}_A(B) = \exp(A) B \exp(-A) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{ad}_A^k B, \quad B \in \mathfrak{g}. \tag{13}$$

The type of matrices we will handle more frequently are orthogonal, unitary and symplectic. Here are their characterization and the notation we shall use for their group and algebra.

The special orthogonal group, $\mathrm{SO}(n)$, is the set of all $n \times n$ real matrices with unit determinant satisfying $A^{\mathrm{T}} A = A A^{\mathrm{T}} = I$, where A^{T} is the transpose of A and I denotes the identity matrix. The corresponding algebra $\mathfrak{so}(n)$ consists of the skew-symmetric matrices.

A $n \times n$ complex matrix A is called unitary if $A^\dagger A = A A^\dagger = I$, where $A^\dagger$ is the conjugate transpose or Hermitian adjoint of A. The special unitary group, $\mathrm{SU}(n)$, is the set of all $n \times n$ unitary matrices with unit determinant. The corresponding algebra $\mathfrak{su}(n)$ consists of the skew-Hermitian traceless matrices. Special relevance in some quantum mechanical problems we discuss will have the case $n = 2$. In this case a convenient basis for $\mathfrak{su}(2)$ is made up by the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{14}$$

They satisfy the identity

$$\sigma_j \sigma_k = \delta_{jk} + \mathrm{i} \epsilon_{jkl} \sigma_l, \tag{15}$$

and correspondingly

$$[\sigma_j, \sigma_k] = 2\mathrm{i} \epsilon_{jkl} \sigma_l, \tag{16}$$

which directly give the structure constants for SU(2). The following identities will prove useful for $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^3$:

$$(\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) = \mathbf{a} \cdot \mathbf{b}\, I + \mathrm{i}(\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}, \qquad [\mathbf{a} \cdot \boldsymbol{\sigma}, \mathbf{b} \cdot \boldsymbol{\sigma}] = 2\mathrm{i}(\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}, \tag{17}$$

where we have denoted $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$. Any $U \in \mathrm{SU}(2)$ can be written as

$$U = \exp(\mathrm{i}\mathbf{a} \cdot \boldsymbol{\sigma}) = \cos(a) I + \mathrm{i} \frac{\sin(a)}{a} \mathbf{a} \cdot \boldsymbol{\sigma}, \tag{18}$$

where $a = \|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$. A more elaborate expression which we shall make use of, in later sections, is (with $a = 1$)

$$\exp(\mathrm{i}\mathbf{a} \cdot \boldsymbol{\sigma} t)(\mathbf{b} \cdot \boldsymbol{\sigma}) \exp(-\mathrm{i}\mathbf{a} \cdot \boldsymbol{\sigma} t) = \mathbf{b} \cdot \boldsymbol{\sigma} + \sin 2t\, (\mathbf{b} \times \mathbf{a}) \cdot \boldsymbol{\sigma} + \sin^2 t (\mathbf{a} \times (\mathbf{b} \times \mathbf{a})) \cdot \boldsymbol{\sigma}. \tag{19}$$

In Hamiltonian problems, the symplectic group $\mathrm{Sp}(n)$ plays a fundamental role. It is the group of $2n \times 2n$ real matrices satisfying

$$A^{\mathrm{T}} J A = J, \quad \text{with } J = \begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix} \tag{20}$$

and I_n denotes the n-dimensional identity matrix. Its corresponding Lie algebra $\mathfrak{sp}(n)$ consists of matrices verifying $B^{\mathrm{T}} J + J B = O_{2n}$. In fact, these can be considered particular instances of the so-called J-orthogonal group, defined as [41]

$$\mathrm{O}_J(n) = \{A \in \mathrm{GL}(n) : A^{\mathrm{T}} J A = J\}, \tag{21}$$

where GL(n) is the group of all $n \times n$ nonsingular real matrices and J is some constant matrix in GL(n). Thus, one recovers the orthogonal group when $J = I$, the symplectic group Sp(n) when J is the basic symplectic matrix given in (20), and the Lorentz group SO(3, 1) when $J = \mathrm{diag}(1, -1, -1, -1)$. The corresponding Lie algebra is the set

$$\mathfrak{o}_J(n) = \{B \in \mathfrak{gl}_n(\mathbb{R}) : B^{\mathrm{T}}J + JB = O\}, \tag{22}$$

where $\mathfrak{gl}_n(\mathbb{R})$ is the Lie algebra of all $n \times n$ real matrices. If $B \in \mathfrak{o}_J(n)$, then its Cayley transform

$$A = (I - \alpha B)^{-1}(I + \alpha B) \tag{23}$$

is J-orthogonal.

Another important matrix Lie group not included in the previous characterization, is the special linear group SL(n), formed by all $n \times n$ real matrices with unit determinant. The corresponding Lie algebra $\mathfrak{sl}(n)$ comprises all traceless matrices. For real 2×2 matrices in $\mathfrak{sl}(2)$ one has

$$\exp\begin{pmatrix} a & b \\ c & -a \end{pmatrix} = \begin{pmatrix} \cosh(\eta) + \dfrac{a}{\eta}\sinh(\eta) & \dfrac{b}{\eta}\sinh(\eta) \\ \dfrac{c}{\eta}\sinh(\eta) & \cosh(\eta) - \dfrac{a}{\eta}\sinh(\eta) \end{pmatrix} \tag{24}$$

with $\eta = \sqrt{a^2 + bc}$.

When dealing with convergence problems it is necessary to use some type of norm for a matrix. By such we mean a non-negative real number $\|A\|$ associated with each matrix $A \in \mathbb{C}^{n\times n}$ and satisfying

(a) $\|A\| \geq 0$ for all A and $\|A\| = 0$ iff $A = O_n$.
(b) $\|\alpha A\| = |\alpha|\,\|A\|$, for all scalars α.
(c) $\|A + B\| \leq \|A\| + \|B\|$.

Quite often, one adds the sub-multiplicative property

$$\|AB\| \leq \|A\|\,\|B\|, \tag{25}$$

but not all matrix norms satisfy this condition [42].

There exist different families of matrix norms. Among the more popular ones, we have the p-norm $\|A\|_p$ and the Frobenius norm $\|A\|_F$. For a matrix A with elements a_{ij}, $i, j = 1 \ldots n$, they are defined as

$$\|A\|_p = \max_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p \tag{26}$$

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{\mathrm{tr}(A^\dagger A)}, \tag{27}$$

respectively, where $\|\mathbf{x}\|_p = (\sum_{j=1}^{n} |x_j|^p)^{\frac{1}{p}}$ and $\mathrm{tr}(A)$ is the trace of the matrix A. Although both verify (25), the p-norms have the important property that for every matrix A and $\mathbf{x} \in \mathbb{R}^n$ one has $\|A\mathbf{x}\|_p \leq \|A\|_p\,\|\mathbf{x}\|_p$. The most used p-norms correspond to $p = 1$, $p = 2$ and $p = \infty$.

Of paramount importance in numerical linear algebra is the case $p = 2$. The resulting 2-norm of a vector is nothing but the Euclidean norm, whereas in the matrix case it is also called the spectral norm of A and can be characterized as the square root of the largest eigenvalue of $A^\dagger A$. A frequently used inequality relating to Frobenius and spectral norms is the following:

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\,\|A\|_2. \tag{28}$$

In fact, this last inequality can be made more stringent [43]:

$$\|A\|_F \leq \sqrt{\mathrm{rank}(A)}\,\|A\|_2. \tag{29}$$

Considering in a matrix Lie algebra $\mathfrak{g}$ a norm satisfying property (25), it is clear that $\|[A, B]\| \leq 2\|A\|\|B\|$, and the ad operator defined by (11) is bounded, since

$$\|\mathrm{ad}_A\| \leq 2\|A\|$$

for any matrix A.

A matrix norm is said to be unitarily invariant if $\|UAV\| = \|A\|$ whenever U, V are unitary matrices. Frobenius and p-norms are both unitarily invariant [44].

In some of the most basic formulas for the Magnus expansion, there will appear the so-called Bernoulli numbers B_n, which are defined through the generating function [45]

$$\frac{t\mathrm{e}^{zt}}{\mathrm{e}^t - 1} = \sum_{n=0}^{\infty} B_n(z)\frac{t^n}{n!}, \quad |t| < 2\pi$$

as $B_n = B_n(0)$. Equivalently,

$$\frac{x}{\mathrm{e}^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n,$$

whereas the formula

$$\frac{\mathrm{e}^x - 1}{x} = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} x^n$$

will be also useful in the sequel. The first few nonzero Bernoulli numbers are $B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}, B_4 = -\frac{1}{30}$. In general one has $B_{2m+1} = 0$ for $m \geq 1$.

2. The Magnus expansion (ME)

Magnus' proposal with respect to the linear evolution equation

$$Y'(t) = A(t)Y(t) \tag{30}$$

with initial condition $Y(0) = I$, was to express the solution as the exponential of a certain function,

$$Y(t) = \exp \Omega(t). \tag{31}$$

This is in contrast to the representation

$$Y(t) = \mathcal{T}\left(\exp \int_0^t A(s)\mathrm{d}s\right)$$

in terms of the *time-ordering operator* $\mathcal{T}$ introduced by Dyson [13].

It turns out that $\Omega(t)$ in (31) can be obtained explicitly in a number of ways. The crucial point is to derive a differential equation for the operator Ω that replaces (30). Here we reproduce the result first established by Magnus as Theorem III in [4]:

Theorem 1 (*Magnus 1954*). *Let $A(t)$ be a known function of t (in general, in an associative ring), and let $Y(t)$ be an unknown function satisfying (30) with $Y(0) = I$. Then, if certain unspecified conditions of convergence are satisfied, $Y(t)$ can be written in the form*

$$Y(t) = \exp \Omega(t),$$

where

$$\frac{\mathrm{d}\Omega}{\mathrm{d}t} = \sum_{n=0}^{\infty} \frac{B_n}{n!} \mathrm{ad}_{\Omega}^n A, \tag{32}$$

and B_n are the Bernoulli numbers. Integration of (32), by iteration, leads to an infinite series for Ω, the first terms of which are

$$\Omega(t) = \int_0^t A(t_1)\mathrm{d}t_1 - \frac{1}{2}\int_0^t \left[\int_0^{t_1} A(t_2)\mathrm{d}t_2, A(t_1)\right]\mathrm{d}t_1 + \cdots.$$

2.1. A proof of Magnus' Theorem

The proof of this theorem is largely based on the derivative of the matrix exponential map, which we discuss next. Given a scalar function $\omega(t) \in \mathbb{R}$, the derivative of the exponential is given by $\mathrm{d}\exp(\omega(t))/\mathrm{d}t = \omega'(t)\exp(\omega(t))$. One could think of a similar formula for a matrix $\Omega(t)$. However, this is not the case, since in general $[\Omega, \Omega'] \neq 0$. Instead, one has the following result.

Lemma 2. *The derivative of a matrix exponential can be written alternatively as*

$$\text{(a)} \quad \frac{\mathrm{d}}{\mathrm{d}t}\exp(\Omega(t)) = \mathrm{d}\exp_{\Omega(t)}(\Omega'(t))\exp(\Omega(t)), \tag{33}$$

$$\text{(b)} \quad \frac{\mathrm{d}}{\mathrm{d}t}\exp(\Omega(t)) = \exp(\Omega(t))\mathrm{d}\exp_{-\Omega(t)}(\Omega'(t)), \tag{34}$$

$$\text{(c)} \quad \frac{\mathrm{d}}{\mathrm{d}t}\exp(\Omega(t)) = \int_0^1 \mathrm{e}^{x\Omega(t)}\,\Omega'(t)\mathrm{e}^{(1-x)\Omega(t)}\mathrm{d}x, \tag{35}$$

where $\mathrm{d}\exp_\Omega(C)$ *is defined by its (everywhere convergent) power series*

$$\mathrm{d}\exp_\Omega(C) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\,\mathrm{ad}_\Omega^k(C) \equiv \frac{\exp(\mathrm{ad}_\Omega) - I}{\mathrm{ad}_\Omega}(C). \tag{36}$$

Proof. Let $\Omega(t)$ be a matrix-valued differentiable function and set

$$Y(\sigma, t) \equiv \frac{\partial}{\partial t}\left(\exp(\sigma\Omega(t))\right)\exp(-\sigma\Omega(t))$$

for $\sigma, t \in \mathbb{R}$. Differentiating with respect to σ,

$$\begin{aligned}\frac{\partial Y}{\partial \sigma} &= \frac{\partial}{\partial t}\left(\exp(\sigma\Omega)\Omega\right)\exp(-\sigma\Omega) + \frac{\partial}{\partial t}\left(\exp(\sigma\Omega)\right)(-\Omega)\exp(-\sigma\Omega)\\ &= \left(\exp(\sigma\Omega)\Omega' + \frac{\partial}{\partial t}\left(\exp(\sigma\Omega)\right)\Omega\right)\exp(-\sigma\Omega) - \frac{\partial}{\partial t}\left(\exp(\sigma\Omega)\right)\Omega\exp(-\sigma\Omega) = \exp(\sigma\Omega)\Omega'\exp(-\sigma\Omega)\\ &= \exp(\mathrm{ad}_{\sigma\Omega})(\Omega') = \sum_{k=0}^{\infty}\frac{\sigma^k}{k!}\mathrm{ad}_\Omega^k(\Omega'),\end{aligned}$$

where the first equality in the last line follows readily from (12) and (13). On the other hand

$$\frac{\mathrm{d}}{\mathrm{d}t}(\exp\Omega)\exp(-\Omega) = Y(1,t) = \int_0^1 \frac{\partial}{\partial\sigma}Y(\sigma,t)\mathrm{d}\sigma \tag{37}$$

since $Y(0,t) = 0$, and

$$\int_0^1 \frac{\partial}{\partial\sigma}Y(\sigma,t)\mathrm{d}\sigma = \int_0^1 \sum_{k=0}^{\infty}\frac{\sigma^k}{k!}\mathrm{ad}_\Omega^k(\Omega')\mathrm{d}\sigma = \sum_{k=0}^{\infty}\frac{1}{(k+1)!}\mathrm{ad}_\Omega^k(\Omega'),$$

from which formula (33) follows. The convergence of the power series (36) is a consequence of the boundedness of the ad operator: $\|\mathrm{ad}_\Omega\| \le 2\|\Omega\|$.

Multiplying both sides of (33) by $\exp(-\Omega)$, we have

$$\mathrm{e}^{-\Omega}\frac{\mathrm{d}\mathrm{e}^{\Omega}}{\mathrm{d}t} = \mathrm{e}^{-\Omega}\mathrm{d}\exp_\Omega(\Omega')\mathrm{e}^{\Omega} = \mathrm{e}^{\mathrm{ad}_{-\Omega}}\mathrm{d}\exp_\Omega(\Omega') = \frac{\mathrm{e}^{\mathrm{ad}_{-\Omega}} - I}{\mathrm{ad}_{-\Omega}}\Omega' = \mathrm{d}\exp_{-\Omega}(\Omega')$$

from which (34) follows readily. Finally, Eq. (35) is obtained by taking

$$\int_0^1 \frac{\partial}{\partial\sigma}Y(\sigma,t)\mathrm{d}\sigma = \int_0^1 \exp(\sigma\Omega)\Omega'\exp(-\sigma\Omega)\mathrm{d}\sigma$$

in (37). ■

According to Rossmann [46] and Sternberg [47], formula (33) was first proved by F. Schur in 1890 [48] and was taken up later from a different point of view by Poincaré (1899), whereas the integral formulation (35) has been derived a number of times in the physics literature [19].

As a consequence of the Inverse Function Theorem, the exponential map has a local inverse in the vicinity of a point Ω at which $\mathrm{d}\exp_\Omega = (\exp(\mathrm{ad}_\Omega) - I)/\mathrm{ad}_\Omega$ is invertible. The following lemma establishes when this takes place.

Lemma 3 (*Baker 1905*). *If the eigenvalues of the linear operator* ad_Ω *are different from* $2m\pi\mathrm{i}$ *with* $m \in \{\pm1, \pm2, \ldots\}$, *then* $\mathrm{d}\exp_\Omega$ *is invertible. Furthermore,*

$$\mathrm{d}\exp_\Omega^{-1}(C) = \frac{\mathrm{ad}_\Omega}{\mathrm{e}^{\mathrm{ad}_\Omega} - I}C = \sum_{k=0}^{\infty}\frac{B_k}{k!}\mathrm{ad}_\Omega^k(C) \tag{38}$$

and the convergence of the $\mathrm{d}\exp_\Omega^{-1}$ *expansion is certainly assured if* $\|\Omega\| < \pi$.

Proof. The eigenvalues of $\mathrm{d}\exp_\Omega$ are of the form

$$\mu = \sum_{k\ge0}\frac{\nu^k}{(k+1)!} = \frac{\mathrm{e}^{\nu} - 1}{\nu},$$

where ν is an eigenvalue of ad_Ω. By assumption, the values of μ are non-zero, so that $\mathrm{d}\exp_\Omega$ is invertible. By definition of the Bernoulli numbers, the composition of (38) with (36) gives the identity. Convergence for $\|\Omega\| < \pi$ follows from $\|\mathrm{ad}_\Omega\| \le 2\|\Omega\|$ and from the fact that the radius of convergence of the series expansion for $x/(\mathrm{e}^x - 1)$ is 2π. ■

It remains to determine the eigenvalues of the operator ad_Ω. In fact, it is not difficult to show that if Ω has n eigenvalues $\{\lambda_j,\ j = 1, 2, \ldots, n\}$, then ad_Ω has n^2 eigenvalues $\{\lambda_j - \lambda_k,\ j, k = 1, 2, \ldots, n\}$.

As a consequence of the previous discussion, Theorem 1 can be rephrased more precisely in the following terms.

Theorem 4. *The solution of the differential equation $Y' = A(t)Y$ with initial condition $Y(0) = Y_0$ can be written as $Y(t) = \exp(\Omega(t))Y_0$ with $\Omega(t)$ defined by*

$$\Omega' = \mathrm{d}\exp_\Omega^{-1}(A(t)), \quad \Omega(0) = O, \tag{39}$$

where

$$\mathrm{d}\exp_\Omega^{-1}(A) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathrm{ad}_\Omega^k(A).$$

Proof. Comparing the derivative of $Y(t) = \exp(\Omega(t))Y_0$,

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\exp(\Omega(t))\right) Y_0 = \mathrm{d}\exp_\Omega(\Omega') \exp(\Omega(t))Y_0$$

with $Y' = A(t)Y$, we obtain $A(t) = \mathrm{d}\exp_\Omega(\Omega')$. Applying the inverse operator $\mathrm{d}\exp_\Omega^{-1}$ to this relation yields the differential equation (39) for $\Omega(t)$. ■

Taking into account the numerical values of the first few Bernoulli numbers, the differential equation (39) therefore becomes

$$\Omega' = A(t) - \frac{1}{2}[\Omega, A(t)] + \frac{1}{12}[\Omega, [\Omega, A(t)]] + \cdots,$$

which is nonlinear in Ω. By defining

$$\Omega^{[0]} = O, \qquad \Omega^{[1]} = \int_0^t A(t_1)\mathrm{d}t_1,$$

and applying Picard fixed point iteration, one gets

$$\Omega^{[n]} = \int_0^t \left(A - \frac{1}{2}[\Omega^{[n-1]}, A] + \frac{1}{12}[\Omega^{[n-1]}, [\Omega^{[n-1]}, A]] + \cdots\right) \mathrm{d}t_1$$

and $\lim_{n\to\infty} \Omega^{[n]}(t) = \Omega(t)$ in a suitably small neighborhood of the origin.

2.2. Formulae for the first terms in Magnus expansion

Suppose now that A is of first order in some parameter ε and try a solution in the form of a series

$$\Omega(t) = \sum_{n=1}^{\infty} \Omega_n(t), \tag{40}$$

where Ω_n is supposed to be of order ε^n. Equivalently, we replace A by εA in (30) and determine the successive terms of

$$\Omega(t) = \sum_{n=1}^{\infty} \varepsilon^n \Omega_n(t). \tag{41}$$

This can be done explicitly, at least for the first terms, by substituting the series (41) in (39) and equating powers of ε. Obviously, the Magnus series (40) is recovered by taking $\varepsilon = 1$. Thus, using the notation $A(t_i) \equiv A_i$, the first four orders read

(1) $\Omega_1' = A$, so that

$$\Omega_1(t) = \int_0^t \mathrm{d}t_1 A_1. \tag{42}$$

(2) $\Omega_2' = -\frac{1}{2}[\Omega_1, A]$. Thus

$$\Omega_2(t) = \frac{1}{2}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 [A_1, A_2]. \tag{43}$$

(3) $\Omega_3' = -\frac{1}{2}[\Omega_2, A] + \frac{1}{12}[\Omega_1, [\Omega_1, A]]$. After some work, and using the formula

$$\int_0^\alpha \mathrm{d}x \int_0^x f(x, y)\mathrm{d}y = \int_0^\alpha \mathrm{d}y \int_y^\alpha f(x, y)\mathrm{d}x \tag{44}$$

we obtain

$$\Omega_3(t) = \frac{1}{6}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3 \{[A_1,[A_2,A_3]] + [[A_1,A_2],A_3]\}. \tag{45}$$

(4) $\Omega_4' = -\frac{1}{2}[\Omega_3, A] + \frac{1}{12}[\Omega_2,[\Omega_1,A]] + \frac{1}{12}[\Omega_1,[\Omega_2,A]]$, which yields

$$\begin{aligned}\Omega_4(t) = \frac{1}{12}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_4 \{&[[[A_1,A_2],A_3]A_4] \\ &+ [A_1,[[A_2,A_3],A_4]] + [A_1,[A_2,[A_3,A_4]]] + [A_2,[A_3,[A_4,A_1]]]\}.\end{aligned} \tag{46}$$

The apparent symmetry in the formulae above is deceptive. High orders require repeated use of (44) and become unwieldy. Prato and Lamberti [49] give explicitly the fifth order using an algorithmic point of view. One can also find in the literature, quite involved explicit expressions for an arbitrary order [50–54]. In the next subsection we describe a recursive procedure to generate the terms in the expansion.

2.3. Magnus expansion generator

The above procedure can, indeed, provide a recursive procedure to generate all the terms in the Magnus series (40). Thus, by substituting $\Omega(t) = \sum_{n=1}^{\infty} \Omega_n$ into Eq. (39) and equating terms of the same order, one gets in general

$$\begin{aligned}\Omega_1' &= A \\ \Omega_n' &= \sum_{j=1}^{n-1} \frac{B_j}{j!} S_n^{(j)}, \quad n \geq 2,\end{aligned} \tag{47}$$

where

$$S_n^{(k)} = \sum [\Omega_{i_1}, [\ldots[\Omega_{i_k}, A]\ldots]] \quad (i_1 + \cdots + i_k = n-1). \tag{48}$$

Notice that, in the last equation, the order in A has been explicitly reckoned, whereas k represents the number of Ω's. The newly defined operators $S_n^{(k)}$ can again be calculated recursively. The recurrence relations are now given by

$$\begin{aligned}S_n^{(j)} &= \sum_{m=1}^{n-j} \left[\Omega_m, S_{n-m}^{(j-1)}\right], \quad 2 \leq j \leq n-1 \\ S_n^{(1)} &= [\Omega_{n-1}, A], \qquad S_n^{(n-1)} = \mathrm{ad}_{\Omega_1}^{n-1}(A).\end{aligned} \tag{49}$$

After integration we reach the final result in the form

$$\begin{aligned}\Omega_1 &= \int_0^t A(\tau)\mathrm{d}\tau \\ \Omega_n &= \sum_{j=1}^{n-1} \frac{B_j}{j!} \int_0^t S_n^{(j)}(\tau)\mathrm{d}\tau, \quad n \geq 2.\end{aligned} \tag{50}$$

Alternatively, the expression of $S_n^{(k)}$ given by (48) can be inserted into (50), thus arriving at

$$\Omega_n(t) = \sum_{j=1}^{n-1} \frac{B_j}{j!} \sum_{\substack{k_1+\cdots+k_j=n-1 \\ k_1\geq 1,\ldots,k_j\geq 1}} \int_0^t \mathrm{ad}_{\Omega_{k_1}(s)} \mathrm{ad}_{\Omega_{k_2}(s)} \cdots \mathrm{ad}_{\Omega_{k_j}(s)} A(s)\mathrm{d}s \quad n \geq 2. \tag{51}$$

Notice that each term $\Omega_n(t)$ in the Magnus series is a multiple integral of combinations of $n-1$ nested commutators containing n operators $A(t)$. If, in particular, $A(t)$ belongs to some Lie algebra $\mathfrak{g}$, then it is clear that $\Omega(t)$ (and in fact any truncation of the Magnus series) also stays in $\mathfrak{g}$ and therefore $\exp(\Omega) \in \mathcal{G}$, where $\mathcal{G}$ denotes the Lie group whose corresponding Lie algebra (the tangent space at the identity of $\mathcal{G}$) is $\mathfrak{g}$.

2.4. Magnus expansion and time-dependent perturbation theory

It is not difficult to establish a connection between Magnus series and Dyson perturbative series [13]. The latter gives the solution of (30) as

$$Y(t) = I + \sum_{n=1}^{\infty} P_n(t), \tag{52}$$

where P_n are time-ordered products

$$P_n(t) = \int_0^t \mathrm{d}t_1 \ldots \int_0^{t_{n-1}} \mathrm{d}t_n\, A_1 A_2 \ldots A_n,$$

where $A_i \equiv A(t_i)$. Then

$$\sum_{j=1}^{\infty} \Omega_j(t) = \log\left(I + \sum_{j=1}^{\infty} P_j(t) \right).$$

As stated by Salzman [55],

$$\Omega_n = P_n - \sum_{j=2}^{n} \frac{(-1)^n}{j} R_n^{(j)}, \quad n \geq 2, \tag{53}$$

where

$$R_n^{(k)} = \sum P_{i_1} P_{i_2} \ldots P_{i_k} \quad (i_1 + \cdots + i_k = n)$$

obeys the quadratic recursion formula

$$R_n^{(j)} = \sum_{m=1}^{n-j+1} R_m^{(1)} R_{n-m}^{(j-1)}, \tag{54}$$
$$R_n^{(1)} = P_n, \qquad R_n^{(n)} = P_1^n.$$

Eq. (54) represents the Magnus expansion generator in Salzman's approach. It may be useful to write down the first few equations provided by this formalism:

$$\begin{aligned} \Omega_1 &= P_1 \\ \Omega_2 &= P_2 - \frac{1}{2} P_1^2 \\ \Omega_3 &= P_3 - \frac{1}{2}(P_1 P_2 + P_2 P_1) + \frac{1}{3} P_1^3. \end{aligned} \tag{55}$$

A similar set of equations was developed by Burum [56], thus providing

$$\begin{aligned} P_1 &= \Omega_1, \\ P_2 &= \Omega_2 + \frac{1}{2!}\Omega_1^2, \\ P_3 &= \Omega_3 + \frac{1}{2!}(\Omega_1 \Omega_2 + \Omega_2 \Omega_1) + \frac{1}{3!}\Omega_1^3 \end{aligned} \tag{56}$$

and so on. The general term reads

$$\Omega_n = P_n - \sum_{j=2}^{n} \frac{1}{j} Q_n^{(j)}, \quad n \geq 2, \tag{57}$$

where

$$Q_n^{(k)} = \sum \Omega_{i_1} \ldots \Omega_{i_k}, \quad (i_1 + \cdots + i_k = n). \tag{58}$$

As before, subscripts indicate the order with respect to the parameter ε, while superscripts represent the number of factors in each product. Thus, the summation in (58) extends over all possible products of k (in general non-commuting) operators Ω_i such that the overall order of each term is equal to n. By regrouping terms, one has

$$Q_n^{(k)} = \Omega_1 \sum_{i_2 + \cdots + i_k = n-1} \Omega_{i_2} \cdots \Omega_{i_k} + \Omega_2 \sum_{i_2 + \cdots + i_k = n-2} \Omega_{i_2} \cdots \Omega_{i_k} + \cdots + \Omega_{n-k+1} \sum_{i_2 + \cdots + i_k = k-1} \Omega_{i_2} \cdots \Omega_{i_k}, \tag{59}$$

where $Q_n^{(j)}$ may also be obtained recursively from

$$Q_n^{(j)} = \sum_{m=1}^{n-j+1} Q_m^{(1)} Q_{n-m}^{(j-1)}, \tag{60}$$
$$Q_n^{(1)} = \Omega_n, \qquad Q_n^{(n)} = \Omega_1^n.$$

By working out this recurrence, one gets the same expressions as (54) for the first terms. Further aspects of the relationship between Magnus, Dyson series and time-ordered products can be found in [57,58].

2.5. Graph theoretical analysis of Magnus expansion

The previous recursions allow us, in principle, to express any Ω_k in the Magnus series in terms of $\Omega_1, \ldots, \Omega_{k-1}$. In fact, this procedure has some advantages from a computational point of view. On the other hand, as we have mentioned before, when the recursions are solved explicitly, Ω_k can be expanded as a linear combination of terms that are composed from integrals and commutators acting iteratively on A. The actual expression, however, becomes increasingly complex with k, as it should be evident from the first terms (42)–(46). An alternative form of the Magnus expansion, amenable also for recursive derivation by using graphical tools, can be obtained by associating each term in the expansion with a *binary rooted tree*, an approach worked out by Iserles and Nørsett [37]. For completeness, in the sequel, we show the equivalence of the recurrence (49)–(50) with this graph theoretical approach.

In essence, the idea of Iserles and Nørsett is to associate each term in Ω_k with a rooted tree, according to the following prescription.

Let $\mathcal{T}_0$ be the set consisting of the single rooted tree with one vertex, then $\mathcal{T}_0 = \{\bullet\}$, establish the relationship between this tree and A through the map

$$\bullet \rightsquigarrow A(t)$$

and define recursively

$$\mathcal{T}_m = \left\{ \begin{array}{c} \tau_1 \\ \diagdown\!\diagup \\ \tau_2 \end{array} : \tau_1 \in \mathcal{T}_{k_1},\ \tau_2 \in \mathcal{T}_{k_2}, k_1 + k_2 = m - 1 \right\}.$$

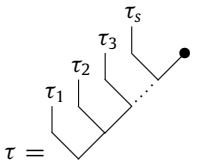
Next, given two expansion terms H_{τ_1} and H_{τ_2}, which have been associated previously with $\tau_1 \in \mathcal{T}_{k_1}$ and $\tau_2 \in \mathcal{T}_{k_2}$, respectively $(k_1 + k_2 = m - 1)$, we associate

$$H_\tau(t) = \left[\int_0^t H_{\tau_1}(\xi) \mathrm{d}\xi, H_{\tau_2}(t) \right] \quad \text{with } \tau = \begin{array}{c} \tau_1 \\ \diagdown\!\diagup \\ \tau_2 \end{array}.$$

Thus, each H_τ for $\tau \in \mathcal{T}_m$ involves exactly m integrals and m commutators.

These composition rules establish a one-to-one relationship between a rooted tree $\tau \in \mathcal{T} \equiv \cup_{m\geq 0} \mathcal{T}_m$, and a matrix function $H_\tau(t)$ involving A, multivariate integrals and commutators.

From here it is easy to deduce that every $\tau \in \mathcal{T}_m$, $m \geq 1$, can be written in a unique way as

$\tau =$ 

or $\tau \equiv a(\tau_1, \tau_2, \ldots, \tau_s)$. Then the Magnus expansion can be expressed in the form [37,59]

$$\Omega(t) = \sum_{m=0}^{\infty} \sum_{\tau \in \mathcal{T}_m} \alpha(\tau) \int_0^t H_\tau(\xi) \mathrm{d}\xi, \tag{61}$$

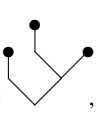
with the scalar $\alpha(\bullet) = 1$ and, in general,

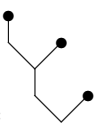
$$\alpha(\tau) = \frac{B_s}{s!} \prod_{l=1}^{s} \alpha(\tau_l).$$

Let us illustrate this procedure by writing down, explicitly, the first terms in the expansion in a tree formalism. In $\mathcal{T}_1$ we only have $k_1 = k_2 = 0$, so that a single tree is possible,

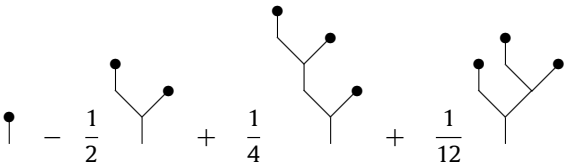
$\tau_1 = \bullet, \quad \tau_2 = \bullet, \quad \Rightarrow \quad \tau =$ 

with $\alpha(\tau) = -1/2$. In $\mathcal{T}_2$ there are two possibilities, namely $k_1 = 0, k_2 = 1$ and $k_1 = 1, k_2 = 0$, and thus one gets

$\tau_1 = \bullet, \quad \tau_2 =$  $\quad \Rightarrow \quad \tau =$  $\quad \alpha(\tau) = \frac{1}{12}$

$\tau_1 =$  $\quad \tau_2 = \bullet \quad \Rightarrow \quad \tau =$  $\quad \alpha(\tau) = \frac{1}{4}$

and the process can be repeated for any $\mathcal{T}_m$. The correspondence between trees and expansion terms should be clear from the previous graphs. For instance, the last tree is nothing but the integral of A, commuted with A, integrated and commuted with A. In that way, by truncating the expansion (61) at $m = 2$ we have

$\Omega(t) =$  $+\cdots,$ (62)

i.e., the explicit expressions collected in Section 2.2.

Finally, the relationship between the tree formalism and the recurrence (49)–(50) can be established as follows. From (61) we can write

$$\sum_{m=1}^{\infty}\sum_{\tau\in\mathcal{T}_m}\alpha(\tau)H_\tau(t)=\sum_{s=1}^{m}\frac{B_s}{s!}\sum_{\substack{k_1,\ldots,k_s\\ k_1+\cdots+k_s=m-s}}\sum_{\tau_i\in\mathcal{T}_{k_i}}\alpha(\tau_1)\cdots\alpha(\tau_s)H_{a(\tau_1,\ldots,\tau_s)}.$$

Thus, by comparing (50) and (61) we have

$$\Omega_m(t)=\sum_{\tau\in\mathcal{T}_{m-1}}\alpha(\tau)\int_0^t H_\tau(\xi)\mathrm{d}\xi=\sum_{j=1}^{m-1}\frac{B_j}{j!}\int_0^t S_m^{(j)}(\xi)\mathrm{d}\xi$$

so that

$$S_m^{(j)}=\sum_{\substack{k_1,\ldots,k_j\\ k_1+\cdots+k_j=m-1-j}}\sum_{\tau_i\in\mathcal{T}_{k_i}}\alpha(\tau_1)\cdots\alpha(\tau_j)H_{a(\tau_1,\ldots,\tau_j)}.$$

In other words, each term $S_n^{(j)}$ in the recurrence (49) carries on a complete set of binary trees. Although both procedures are equivalent, the use of (49) and (50) can be particularly well suited when high orders of the expansion are considered, for two reasons: (i) the enormous number of trees involved for large values of m and (ii) in (61) many terms are redundant, and a careful graph theoretical analysis is needed to deduce which terms have to be discarded [37].

Recently, an ME-type formalism has been developed in the more abstract setting of dendriform algebras. This generalized expansion incorporates the usual one as a limit, but is formulated more in line with (non-commutative) Butcher series. In this context, the use of planar rooted trees to represent the expansion and the so-called pre-Lie product, allows one to reduce the number of terms at each order in comparison with expression (61) [60].

2.6. Time-symmetry of the expansion

The map $\varphi^t : Y(t_0) \longrightarrow Y(t)$ corresponding to the linear differential equation (30) with $Y(t_0) = Y_0$ is time symmetric, $\varphi^{-t}\circ\varphi^t = \mathrm{Id}$, since integrating (30) from $t = t_0$ to $t = t_f$ for every $t_f \geq t_0$ and back to t_0 leads us to the original initial value $Y(t_0) = Y_0$. Observe that, according to (3), the map φ^t can be expressed in terms of the fundamental matrix (or evolution operator) $U(t, t_0)$ as $\varphi^{t_f}(Y_0) = U(t_f, t_0)Y_0$. Then time-symmetry establishes that

$$U(t_0, t_f) = U^{-1}(t_f, t_0)$$

or, in terms of the Magnus expansion,

$$\Omega(t_f, t_0) = -\Omega(t_0, t_f).$$

To take advantage of this feature, let us write the solution of (30) at the final time $t_f = t_0 + s$ as

$$Y\left(t_{1/2}+\frac{s}{2}\right)=\exp\left(\Omega\left(t_{1/2}+\frac{s}{2},t_{1/2}-\frac{s}{2}\right)\right)Y\left(t_{1/2}-\frac{s}{2}\right), \tag{63}$$

where $t_{1/2} = (t_0 + t_f)/2$. Then

$$Y\left(t_{1/2}-\frac{s}{2}\right)=\exp\left(-\Omega\left(t_{1/2}+\frac{s}{2},t_{1/2}-\frac{s}{2}\right)\right)Y\left(t_{1/2}+\frac{s}{2}\right). \tag{64}$$

On the other hand, the solution at t_0 can be written as

$$Y\left(t_{1/2}-\frac{s}{2}\right)=\exp\left(\Omega\left(t_{1/2}-\frac{s}{2},t_{1/2}+\frac{s}{2}\right)\right)Y\left(t_{1/2}+\frac{s}{2}\right), \tag{65}$$

so that, by comparing (64) and (65),

$$\Omega\left(t_{1/2}-\frac{s}{2},t_{1/2}+\frac{s}{2}\right)=-\Omega\left(t_{1/2}+\frac{s}{2},t_{1/2}-\frac{s}{2}\right) \tag{66}$$

and thus Ω does not contain even powers of s. If $A(t)$ is an analytic function and a Taylor series centered around $t_{1/2}$ is considered, then each term in Ω_k is an odd function of s and, in particular, $\Omega_{2i+1}(s)=\mathcal{O}(s^{2i+3})$. This fact has been noticed in [61,62] and will be fully exploited in Section 5.4 when analyzing the Magnus expansion as a numerical device for integrating differential equations.

2.7. Convergence of the Magnus expansion

As we pointed out in the introduction, from a mathematical point of view, there are at least two different issues of paramount importance at the very basis of the Magnus expansion:

(1) (*Existence*). For what values of t and for what operators A does Eq. (30) admit an exponential solution in the form $Y(t)=\exp(\Omega(t))$ for a certain $\Omega(t)$?
(2) (*Convergence*). Given a certain operator $A(t)$, for what values of t does the Magnus series (40) converge?

Of course, given the relevance of the expansion, both problems have been extensively treated in the literature since Magnus proposed this formalism in 1954. We next review some of the most relevant contributions available regarding both aspects, with special emphasis on the convergence of the Magnus series.

2.7.1. On the existence of $\Omega(t)$

In most cases, one is interested in the case where A belongs to a Lie algebra $\mathfrak{g}$ under the commutator product. In this general setting, the Magnus theorem can be formulated as four statements concerning the solution of $Y'=A(t)Y$, each one more stringent than the preceding [16]. Specifically,

(A) The differential equation $Y'=A(t)Y$ has a solution of the form $Y(t)=\exp\Omega(t)$.
(B) The exponent $\Omega(t)$ lies in the Lie algebra $\mathfrak{g}$.
(C) The exponent $\Omega(t)$ is a continuous differentiable function of $A(t)$ and t, satisfying the nonlinear differential equation $\Omega'=\mathrm{d}\exp_{\Omega}^{-1}(A(t))$.
(D) The operator $\Omega(t)$ can be computed by the Magnus series (40).

Let us now analyze in detail the conditions under which statements (A)–(D) hold.

(A) If $A(t)$ and $Y(t)$ are $n\times n$ matrices, from well-known general theorems on differential equations it is clear that the initial value problem defined by (30) and $Y(0)=I$ always has a uniquely determined solution $Y(t)$ which is continuous and has a continuous first derivative in any interval in which $A(t)$ is continuous [63]. Furthermore, the determinant of Y is always different from zero, since

$$\det Y(t)=\exp\left(\int_0^t \operatorname{tr}A(s)\mathrm{d}s\right).$$

On the other hand, it is well known that any matrix Y can be written in the form $\exp\Omega$ if and only if $\det Y\neq 0$ [64, p. 239], so that it is always possible to write $Y(t)=\exp\Omega(t)$.

In the general context of Lie groups and Lie algebras, it is indeed the regularity of the exponential map from the Lie algebra $\mathfrak{g}$ to the Lie group $\mathcal{G}$ that determines the global existence of an $\Omega(t)\in\mathfrak{g}$ [65,66]: the exponential map of a complex Lie algebra is globally one-to-one if and only if the algebra is nilpotent, i.e. there exists a finite n such that $\mathrm{ad}_{x_1}\mathrm{ad}_{x_2}\cdots\mathrm{ad}_{x_{n-1}}x_n=0$, where x_j are arbitrary elements from the Lie algebra. In general, however, the injectivity of the exponential map is only assured for $\xi\in\mathfrak{g}$ such that $\|\xi\|<\rho_{\mathcal{G}}$ for a real number $\rho_{\mathcal{G}}>0$ and some norm in $\mathfrak{g}$ [32,67].

(B) Although in principle $\rho_{\mathcal{G}}$ constitutes a sharp upper bound for the mere existence of the operator $\Omega\in\mathfrak{g}$, its practical value in the case of differential equations is less clear. As we have noticed, any nonsingular matrix has a logarithm, but this logarithm might be in $\mathfrak{gl}(n,\mathbb{C})$ even when the matrix is real. The logarithm of $Y(t)$ may be complex even for real $A(t)$ [16]. In such a situation, the solution of (30) cannot be written as the exponential of a matrix belonging to the Lie algebra over the field of real numbers. One might argue that this is indeed possible over the field of complex numbers, but (i) the element Ω cannot be computed by the Magnus series (40), since it contains only real rational coefficients, and (ii) examples exist where the logarithm of a complex matrix does not lie in the corresponding Lie subalgebra [16].

It is therefore interesting to determine for which range of t a real $A(t)$ in (30) leads to a real logarithm. This issue has been tackled by Moan in [67] in the context of a complete normed (Banach) algebra, proving that if

$$\int_0^t \|A(s)\|_2\mathrm{d}s<\pi \tag{67}$$

then the solution of (30) can indeed be written as $Y(t)=\exp\Omega(t)$, where $\Omega(t)$ is in the Banach algebra.

(C) In his original paper [4], Magnus was well aware that if the function $\Omega(t)$ is assumed to be differentiable, it may not exist everywhere. In fact, he related the differentiability issue to the problem of solving $\mathrm{d}\exp_\Omega(\Omega') = A(t)$ with respect to Ω' and provided an implicit condition for an arbitrary A. More specifically, he proved the following result for the case of $n \times n$ matrices (Theorem V in [4]).

Theorem 5. *The equation $A(t) = \mathrm{d}\exp_\Omega(\Omega')$ can be solved by $\Omega' = \mathrm{d}\exp_\Omega^{-1} A(t)$ for an arbitrary $n \times n$ matrix A if and only if none of the differences between any two of the eigenvalues of Ω equals $2\pi\,\mathrm{i}m$, where $m = \pm 1, \pm 2, \ldots, (m \neq 0)$.*

This result can be considered, in fact, as a reformulation of Lemma 3, but, unfortunately, does not have very much practical application unless the eigenvalues of Ω can easily be determined from those of $A(t)$. One would instead like to have conditions based directly on A.

2.7.2. Convergence of the Magnus series

For dealing with the validity of statement (D) one has to analyze the convergence of the series $\sum_{k=1}^{\infty} \Omega_k$. Magnus also considered the question of when the series terminates at some finite index m, thus giving a globally valid $\Omega = \Omega_1 + \cdots + \Omega_m$. This will happen, for instance, if

$$\left[A(t), \int_0^t A(s)\mathrm{d}s\right] = 0$$

identically for all values of t, since then $\Omega_k = 0$ for $k > 1$. A sufficient (but not necessary) condition for the vanishing of all terms Ω_k with $k > n$ is that

$$[A(s_1), [A(s_2), [A(s_3), \ldots, [A(s_n), A(s_{n+1})]\ldots]]] = 0$$

for any choice of $s_1, \ldots, s_{n+1}$. In fact, the termination of the series cannot be established solely by consideration of the commutativity of $A(t)$ with itself, and Magnus considered an example illustrating this point.

In general, however, the Magnus series does not converge unless A is small in a suitable sense. Several bounds to the actual radius of convergence in terms of A have been obtained in the literature. Most of these results can be stated as follows. If $\Omega_m(t)$ denotes the homogeneous element with $m-1$ commutators in the Magnus series, as given by (51), then $\Omega(t) = \sum_{m=1}^{\infty} \Omega_m(t)$ is absolutely convergent for $0 \leq t < T$, with

$$T = \max\left\{t \geq 0 : \int_0^t \|A(s)\|_2 \mathrm{d}s < r_c\right\}. \tag{68}$$

Thus, both Pechukas and Light [15] and Karasev and Mosolova [68] obtained $r_c = \log 2 = 0.693147\ldots$, whereas Chacon and Fomenko [69] got $r_c = 0.57745\ldots$. In 1998, Blanes et al. [30] and Moan [70] obtained independently the improved bound

$$r_c = \frac{1}{2}\int_0^{2\pi} \frac{1}{2 + \frac{x}{2}(1 - \cot\frac{x}{2})}\mathrm{d}x \equiv \xi = 1.08686870\ldots \tag{69}$$

by analyzing the recurrence (49)–(50) and (51), respectively. Furthermore, Moan also obtained a bound on the individual terms Ω_m of the Magnus series [67] which is useful, in particular, for estimating errors when the series is truncated. Specifically, he showed that

$$\|\Omega_m(t)\| \leq \frac{f_m}{2}\left(2\int_0^t \|A(s)\|_2 \mathrm{d}s\right)^m \leq \pi\left(\frac{1}{\xi}\int_0^t \|A(s)\|_2\,\mathrm{d}s\right)^m,$$

where f_m are the coefficients of

$$G^{-1}(x) = \sum_{m\geq 1} f_m x^m = x + \frac{1}{4}x^2 + \frac{5}{72}x^3 + \frac{11}{576}x^4 + \frac{479}{86\,400}x^5 + \cdots,$$

the inverse function of

$$G(s) = \int_0^s \frac{1}{2 + \frac{x}{2}(1 - \cot\frac{x}{2})}\,\mathrm{d}x.$$

On the other hand, by analyzing some selected examples, Moan [67] concluded that, in order to get convergence *for all real matrices* $A(t)$, it is necessary that $r_c \leq \pi$ in (68), and more recently Moan and Niesen [71] have been able to prove that, indeed, $r_c = \pi$ provided only real matrices are involved.

In any case, it is important to remark that statement (D) is locally valid, but cannot be used to compute Ω in the large. However, as we have seen, the other statements need not depend on the validity of (D). In particular, if (B) and (C) are globally valid, one can still investigate many of the properties of Ω, even though one cannot compute it with the aid of (D).

2.7.3. An improved radius of convergence

The previous results on the convergence of the Magnus series have been established for $n \times n$ real matrices: if $A(t)$ is a real $n \times n$ matrix, then (67) gives a condition for $Y(t)$ to have a real logarithm. In fact, under the same condition, the Magnus series (40) converges precisely to this logarithm, i.e., its sum $\Omega(t)$ satisfies $\exp(\Omega(t)) = Y(t)$ [71].

One should have in mind, however, that the original expansion was conceived by requiring only that $A(t)$ be a linear operator depending on a real variable t in an associative ring (Theorem 1). The idea was to define, in terms of A, an operator $\Omega(t)$ such that the solution of the initial value problem $Y' = A(t)Y$, $Y(0) = I$, for a second operator Y is given as $Y = \exp \Omega$. The proposed expression for Ω is an infinite series, satisfying the condition that "its partial sums become Hermitian after multiplication by i if iA is a Hermitian operator" [4]. As this quotation illustrates, Magnus expansion was first derived in the context of quantum mechanics, and so one typically assumes that it is also valid when $A(t)$ is a linear operator in a Hilbert space. Therefore, it might be desirable to have conditions for the convergence of the Magnus series in this more general setting. In [31], by applying standard techniques of complex analysis and some elementary properties of the unit sphere, the bound $r_c = \pi$ has been shown to be also valid for *any* bounded normal operator $A(t)$ in a Hilbert space of arbitrary dimension. Next we review the main issues involved and refer the reader to [31] for a more detailed treatment.

Let us assume that $A(t)$ is a bounded operator in a Hilbert space $\mathcal{H}$, with $2 \leq \dim \mathcal{H} \leq \infty$. Then we introduce a new parameter $\varepsilon \in \mathbb{C}$ and denote by $Y(t;\varepsilon)$ the solution of the initial value problem

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \varepsilon A(t)Y, \qquad Y(0) = I, \tag{70}$$

where I now denotes the identity operator in $\mathcal{H}$. It is known that $Y(t;\varepsilon)$ is an analytic function of ε for a fixed value of t. Let us introduce the set $B_\gamma \subset \mathbb{C}$ characterized by the real parameter γ,

$$B_\gamma = \left\{ \varepsilon \in \mathbb{C} : |\varepsilon| \int_0^t \|A(s)\| \mathrm{d}s < \gamma \right\}.$$

Here $\|.\|$ stands for the norm defined by the inner product on $\mathcal{H}$, i.e., the 2-norm introduced in Section 1.2.

If t is fixed, the operator function $\varphi(\varepsilon) = \log Y(t;\varepsilon)$ is well defined in B_γ when γ is small enough, say $\gamma < \log 2$, as an analytic function of ε. As a matter of fact, this is a direct consequence of the results collected in Section 2.7.2: if, in particular, $|\varepsilon| \int_0^t \|A(s)\| \mathrm{d}s < \log 2$, the Magnus series corresponding to (70) converges and its sum $\Omega(t;\varepsilon)$ satisfies $\exp(\Omega(t;\varepsilon)) = Y(t;\varepsilon)$. In other words, the power series $\Omega(t;\varepsilon)$ coincides with $\varphi(\varepsilon)$ when $|\varepsilon| \int_0^t \|A(s)\| \mathrm{d}s < \log 2$, and so the Magnus series is the power series expansion of $\varphi(\varepsilon)$ around $\varepsilon = 0$.

Theorem 6. *The function $\varphi(\varepsilon) = \log Y(t;\varepsilon)$ is an analytic function of ε in the set B_π, with*

$$B_\pi = \left\{ \varepsilon \in \mathbb{C} : |\varepsilon| \int_0^t \|A(s)\| \mathrm{d}s < \pi \right\}.$$

If $\mathcal{H}$ is infinite dimensional, the statement holds true if Y is a normal operator.

In other words, $\gamma = \pi$. The proof of this theorem is based on some elementary properties of the unit sphere S^1 in a Hilbert space. Let us define the angle between any two vectors $x \neq 0$, $y \neq 0$ in $\mathcal{H}$, $\mathrm{Ang}\{x, y\} = \alpha$, $0 \leq \alpha \leq \pi$, from

$$\cos \alpha = \frac{\mathrm{Re}\langle x, y\rangle}{\|x\| \, \|y\|},$$

where $\langle \cdot, \cdot \rangle$ is the inner product on $\mathcal{H}$. This angle is a metric in S^1, i.e., the triangle inequality holds in S^1.

The first property we need is given by the next lemma [67].

Lemma 7. *For any $x \in \mathcal{H}$, $x \neq 0$, $\mathrm{Ang}\{Y(t;\varepsilon)x, x\} \leq |\varepsilon| \int_0^t \|A(s)\| \mathrm{d}s$.*

Observe that if Y is a normal operator in $\mathcal{H}$, i.e., $YY^\dagger = Y^\dagger Y$ (in particular, if Y is unitary), then $\|Y^\dagger x\| = \|Yx\|$ for all $x \in \mathcal{H}$ and therefore $\mathrm{Ang}\{Y^\dagger x, x\} = \mathrm{Ang}\{Yx, x\}$.

The second required property provides useful information on the location of the eigenvalues of a given bounded operator in $\mathcal{H}$ [72].

Lemma 8. *Let T be a (bounded) operator on $\mathcal{H}$. If $\mathrm{Ang}\{Tx, x\} \leq \gamma$ and $\mathrm{Ang}\{T^\dagger x, x\} \leq \gamma$ for any $x \neq 0$, $x \in \mathcal{H}$, where $T^\dagger$ denotes the adjoint operator of T, then the spectrum of T, $\sigma(T)$, is contained in the set*

$$\Delta_\gamma = \{z = |z| \mathrm{e}^{\mathrm{i}\omega} \in \mathbb{C} : |\omega| \leq \gamma\}.$$

Proof (*of* Theorem 6). Let us introduce the operator $T \equiv Y(t;\epsilon)$, with $\varepsilon \in B_\gamma$, $\gamma < \pi$. Then by Lemma 7, $\mathrm{Ang}\{Tx, x\} \leq \gamma$ for all $x \neq 0$, and thus, by Lemma 8,

$$\sigma(T) \subset \Delta_\gamma. \tag{71}$$

If $\dim \mathcal{H} = \infty$ and we assume that $Y(t;\epsilon)$ is a normal operator, then (71) also holds.

From Eq. (70) in integral form,

$$Y(t;\varepsilon) = I + \varepsilon \int_0^t A(s)Y \mathrm{d}s,$$

one gets $\|Y\| \leq 1 + |\varepsilon| \int_0^t \|A(s)\| \, \|Y\| \mathrm{d}s$, and application of Gronwall's lemma [73] leads to

$$\|Y(t;\varepsilon)\| \leq \exp\left(|\varepsilon| \int_0^t \|A(s)\| \mathrm{d}s \right).$$

An analogous reasoning for the inverse operator also proves that

$$\|Y^{-1}(t;\varepsilon)\| \leq \exp\left(|\varepsilon| \int_0^t \|A(s)\| \mathrm{d}s \right).$$

In consequence,

$$\|T\| \leq \mathrm{e}^{\gamma} \quad \text{and} \quad \|T^{-1}\| \leq \mathrm{e}^{\gamma}.$$

If $\lambda \neq 0 \in \sigma(T)$, then $|\lambda| \leq \|T\|$ [74] and therefore $|\lambda| \leq \mathrm{e}^{\gamma}$. In addition, $\frac{1}{\lambda} \in \sigma(T^{-1})$, so that $|\lambda| \geq \mathrm{e}^{-\gamma}$. Equivalently,

$$\sigma(T) \subset \{z \in \mathbb{C} : \mathrm{e}^{-\gamma} \leq |z| \leq \mathrm{e}^{\gamma}\} \equiv G_{\gamma}. \tag{72}$$

Putting together (71) and (72), one has

$$\sigma(T) \subset G_{\gamma} \cap \Delta_{\gamma} \equiv \Lambda_{\gamma}.$$

Now choose any value γ' such that $\gamma < \gamma' < \pi$ (e.g., $\gamma' = (\gamma + \pi)/2$) and consider the closed curve $\Gamma = \partial \Lambda_{\gamma'}$. Notice that the curve Γ encloses $\sigma(T)$ in its interior, so that it is possible to define [75] the function $\varphi(\varepsilon) = \log Y(t;\varepsilon)$ by the equation

$$\varphi(\epsilon) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} \log z \, (zI - Y(t;\epsilon))^{-1} \, \mathrm{d}z, \tag{73}$$

where the integration along Γ is performed in the counterclockwise direction. As is well known, (73) defines an analytic function of ε in $B_{\gamma'}$ [75] and the result of the theorem follows. ■

Theorem 9. *Let us consider the differential equation $Y' = A(t)Y$ defined in a Hilbert space $\mathcal{H}$ with $Y(0) = I$, and let $A(t)$ be a bounded operator in $\mathcal{H}$. Then, the Magnus series $\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t)$, with Ω_k given by (51) converges in the interval $t \in [0, T)$ such that*

$$\int_0^{\mathrm{T}} \|A(s)\| \mathrm{d}s < \pi$$

and the sum $\Omega(t)$ satisfies $\exp \Omega(t) = Y(t)$. The statement also holds when $\mathcal{H}$ is infinite-dimensional if Y is a normal operator (in particular, if Y is unitary).

Proof. Theorem 6 shows that $\log Y(t;\varepsilon) \equiv \varphi(\varepsilon)$ is a well defined and analytic function of ε for

$$|\varepsilon| \int_0^t \|A(s)\| \mathrm{d}s < \pi.$$

It has also been shown that the Magnus series $\Omega(t;\varepsilon) = \sum_{k=1}^{\infty} \varepsilon^k \Omega_k(t)$, with Ω_k given by (51), is absolutely convergent when $|\varepsilon| \int_0^t \|A(s)\| \mathrm{d}s < \xi = 1.0868\ldots$ and its sum satisfies $\exp \Omega(t;\varepsilon) = Y(t;\varepsilon)$. Hence, the Magnus series is the power series of the analytic function $\varphi(\varepsilon)$ in the disk $|\varepsilon| < \xi / \int_0^t \|A(s)\| \mathrm{d}s$. But $\varphi(\varepsilon)$ is analytic in $B_{\pi} \supset B_{\xi}$ and the power series has to be unique. In consequence, the power series of $\varphi(\varepsilon)$ in B_{π} has to be the same as the power series of $\varphi(\varepsilon)$ in B_{ξ}, which is precisely the Magnus series. Finally, by taking $\varepsilon = 1$ we get the desired result. ■

2.7.4. Further discussion

Theorem 9 provides sufficient conditions for the convergence of the Magnus series based on an estimate by the norm of the operator A. In particular, it guarantees that the operator $\Omega(t)$ in $Y(t) = \exp \Omega(t)$ can safely be obtained with the convergent series $\sum_{k \geq 1} \Omega_k(t)$ for $0 \leq t < T$ when the terms $\Omega_k(t)$ are computed with (51). A natural question at this stage is — what is the optimal convergence domain? In other words, is the bound estimate $r_c = \pi$ given by Theorem 9 sharp or is there still room for improvement? In order to clarify this issue, we next analyze two simple examples involving 2×2 matrices.

Example 1. Moan and Niesen [71] consider the coefficient matrix

$$A(t) = \begin{pmatrix} 2 & t \\ 0 & -1 \end{pmatrix}. \tag{74}$$

If we introduce, as before, the complex parameter ε in the problem, the corresponding exact solution $Y(t;\varepsilon)$ of (70) is given by

$$Y(t;\varepsilon) = \begin{pmatrix} \mathrm{e}^{2\varepsilon t} & \dfrac{1}{9\varepsilon}\mathrm{e}^{2\varepsilon t} - \left(\dfrac{1}{9\varepsilon} + \dfrac{1}{3}t\right)\mathrm{e}^{-\varepsilon t} \\ 0 & \mathrm{e}^{-\varepsilon t} \end{pmatrix} \tag{75}$$

and therefore

$$\log Y(t;\varepsilon) = \begin{pmatrix} 2t & g(t;\varepsilon) \\ 0 & -t \end{pmatrix}, \quad \text{with } g(t;\varepsilon) = \frac{t(1-\mathrm{e}^{3\varepsilon t}+3\varepsilon t)}{3(1-\mathrm{e}^{3\varepsilon t})}.$$

The Magnus series can be obtained by computing the Taylor expansion of $\log Y(t;\varepsilon)$ around $\varepsilon = 0$. Notice that the function g has a singularity when $\varepsilon t = \frac{2\pi}{3}\mathrm{i}$, and thus, by taking $\varepsilon = 1$, the Magnus series only converges up to $t = \frac{2}{3}\pi$. On the other hand, condition $\int_0^T \|A(s)\|\mathrm{d}s < \pi$ leads to $T \approx 1.43205 < \frac{2}{3}\pi$. In consequence, the actual convergence domain of the Magnus series is larger than the estimate provided by Theorem 9. □

Example 2. Let us introduce the matrices

$$X_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_3, \qquad X_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{76}$$

and define

$$A(t) = \begin{cases} \beta X_2 & 0 \le t \le 1 \\ \alpha X_1 & t > 1 \end{cases}$$

with α, β complex constants. Then, the solution of Eq. (30) at $t = 2$ is

$$Y(2) = \mathrm{e}^{\alpha X_1}\mathrm{e}^{\beta X_2} = \begin{pmatrix} \mathrm{e}^{\alpha} & \beta\mathrm{e}^{\alpha} \\ 0 & \mathrm{e}^{-\alpha} \end{pmatrix},$$

so that

$$\log Y(2) = \log(\mathrm{e}^{\alpha X_1}\mathrm{e}^{\beta X_2}) = \alpha X_1 + \frac{2\alpha\beta}{1-\mathrm{e}^{-2\alpha}} X_2, \tag{77}$$

an analytic function if $|\alpha| < \pi$ with first singularities at $\alpha = \pm\mathrm{i}\pi$. Therefore, the Magnus series cannot converge at $t = 2$ if $|\alpha| \ge \pi$, independently of $\beta \neq 0$, even when it is possible in this case to get a closed-form expression for the general term. Specifically, a straightforward computation with the recurrence (49)–(50) shows that

$$\sum_{n=1}^{\infty} \Omega_n(2) = \alpha X_1 + \beta X_2 + \sum_{n=2}^{\infty} (-1)^{n-1} \frac{2^{n-1}B_{n-1}}{(n-1)!}\, \alpha^{n-1}\beta X_2. \tag{78}$$

If we take the spectral norm, then $\|X_1\| = \|X_2\| = 1$ and

$$\int_0^{t=2} \|A(s)\|\mathrm{d}s = |\alpha| + |\beta|,$$

so that the convergence domain provided by Theorem 9 is $|\alpha| + |\beta| < \pi$ for this example. Notice that in the limit $|\beta| \to 0$ this domain is optimal. □

From the analysis of Examples 1 and 2 we can conclude the following. First, the convergence domain of the Magnus series provided by Theorem 9 is the best result one can get for a generic bounded operator $A(t)$ in a Hilbert space, in the sense that one may consider specific matrices $A(t)$, as in Example 2, where the series diverges for any time t such that $\int_0^t \|A(s)\|\mathrm{d}s > \pi$. Second, there are also situations (as in Example 1) where the bound estimate $r = \pi$ is still rather conservative: *indeed, the Magnus series converges for a larger time interval than that given by* Theorem 9. This is particularly evident if one takes $A(t)$ as a diagonal matrix: then, the exact solution $Y(t;\varepsilon)$ of (70) is a diagonal matrix whose elements are non-vanishing entire functions of ε, and obviously $\log Y(t;\varepsilon)$ is also an entire function of ε. In such circumstances, the convergence domain $|\varepsilon| \int_0^t \|A(s)\|\mathrm{d}s < \pi$ for the Magnus series does not make much sense. Thus, a natural question arises: is it possible to obtain a more precise criterion of convergence? In trying to answer this question, in [31] an alternative characterization of the convergence has been developed which is valid for $n \times n$ complex matrices. More precisely, a connection has been established between the convergence of the Magnus series and the existence of multiple eigenvalues of the fundamental matrix $Y(t;\varepsilon)$ for a fixed t, which we denote by $Y_t(\varepsilon)$. By using the theory of analytic matrix functions, and in particular, of the logarithm of an analytic matrix function (such as is done, e.g. in [76]), the following result has been proved in [31]: if the analytic matrix function $Y_t(\varepsilon)$ has an eigenvalue $\rho_0(\varepsilon_0)$ of multiplicity $l > 1$ for a certain ε_0 such that: (a) there is a curve in

the ε-plane joining $\varepsilon = 0$ with $\varepsilon = \varepsilon_0$, and (b) the number of equal terms in $\log \rho_1(\varepsilon_0), \log \rho_2(\varepsilon_0), \ldots, \log \rho_l(\varepsilon_0)$ such that $\rho_k(\varepsilon_0) = \rho_0$, $k = 1, \ldots, l$ is less than the maximum dimension of the elementary Jordan block corresponding to ρ_0, then the radius of convergence of the series $\Omega_t(\varepsilon) = \sum_{k\geq 1} \varepsilon^k \Omega_{t,k}$ verifying $\exp \Omega_t(\varepsilon) = Y_t(\varepsilon)$ is precisely $r = |\varepsilon_0|$. An analysis along the same line has been carried out in [77].

When this criterion is applied to Example 1, it gives, as the radius of convergence of the Magnus series corresponding to Eq. (70) for a fixed t,

$$\Omega_t(\varepsilon) = \sum_{k=1}^{\infty} \varepsilon^k \, \Omega_{t,k}, \tag{79}$$

the value

$$r = |\varepsilon| = \frac{2\pi}{3t}. \tag{80}$$

To get the actual convergence domain of the usual Magnus expansion we have to take $\varepsilon = 1$, and so, from (80), we get $2\pi/(3t) = 1$, or equivalently $t = 2\pi/3$, i.e., the result achieved from the analysis of the exact solution.

With respect to Example 2, one gets [31]

$$|\varepsilon| = \frac{\pi}{|\alpha|(t-1)}.$$

If we now fix $\varepsilon = 1$, the actual t-domain of convergence of the Magnus series is

$$t = 1 + \frac{\pi}{|\alpha|}.$$

Observe that, when $t = 2$, we get $|\alpha| = \pi$ and thus the previous result is recovered: the Magnus series converges only for $|\alpha| < \pi$.

It should also be mentioned that the case of a diagonal matrix $A(t)$ is compatible with this alternative characterization [31].

2.8. Magnus expansion and the BCH formula

The Magnus expansion can also be used to explicitly obtain the terms of the series Z in

$$Z = \log(\mathrm{e}^{X_1} \mathrm{e}^{X_2}),$$

X_1 and X_2 being two non commuting indeterminate variables. As it is well known [41],

$$Z = X_1 + X_2 + \sum_{n=2}^{\infty} G_n(X_1, X_2), \tag{81}$$

where $G_n(X_1, X_2)$ is a homogeneous Lie polynomial in X_1 and X_2 of grade n; in other words, G_n can be expressed in terms of X_1 and X_2 by addition, multiplication by rational numbers and nested commutators. This result is often known as the Baker–Campbell–Hausdorff (BCH) theorem and proves to be very useful in various fields of mathematics (theory of linear differential equations [4], Lie group theory [78], numerical analysis [79]) and theoretical physics (perturbation theory, transformation theory, Quantum Mechanics and Statistical Mechanics [19,80,81]). In particular, in the theory of Lie groups, with this theorem one can explicitly write the operation of multiplication in a Lie group in canonical coordinates in terms of the Lie bracket operation in its algebra and also prove the existence of a local Lie group with a given Lie algebra [78].

If X_1 and X_2 are matrices and one considers the piecewise constant matrix-valued function

$$A(t) = \begin{cases} X_2 & 0 \leq t \leq 1 \\ X_1 & 1 < t \leq 2 \end{cases} \tag{82}$$

then the exact solution of (30) at $t = 2$ is $Y(2) = \mathrm{e}^{X_1} \mathrm{e}^{X_2}$. By computing $Y(2) = \mathrm{e}^{\Omega(2)}$ with recursion (51) one gets for the first terms

$$\begin{aligned} \Omega_1(2) &= X_1 + X_2 \\ \Omega_2(2) &= \frac{1}{2}[X_1, X_2] \\ \Omega_3(2) &= \frac{1}{12}[X_1, [X_1, X_2]] - \frac{1}{12}[X_2, [X_1, X_2]] \\ \Omega_4(2) &= \frac{1}{24}[X_1, [X_2, [X_2, X_1]]]. \end{aligned} \tag{83}$$

In general, each $G_n(X_1, X_2)$ is a linear combination of the commutators of the form $[V_1, [V_2, \ldots, [V_{n-1}, V_n] \ldots]]$ with $V_i \in \{X_1, X_2\}$ for $1 \leq i \leq n$, the coefficients being universal rational constants. This is perhaps one of the reasons why the Magnus expansion is often referred to in the literature as the continuous analogue of the BCH formula. As a matter of fact, Magnus proposed a different method for obtaining the first terms in the series (40) based on (81) [4].

Now we can apply Theorem 9 and obtain the following sharp bound.

Theorem 10. *The Baker–Campbell–Hausdorff series in the form (81) converges absolutely when* $\|X_1\| + \|X_2\| < \pi$.

This result can be generalized, of course, to any number of non commuting operators $X_1, X_2, \ldots, X_q$. Specifically, the series

$$Z = \log(\mathrm{e}^{X_1}\,\mathrm{e}^{X_2}\,\cdots\,\mathrm{e}^{X_q}),$$

converges absolutely if $\|X_1\| + \|X_2\| + \cdots + \|X_q\| < \pi$.

2.9. Preliminary linear transformations

To improve the accuracy and the bounds on the convergence domain of the Magnus series for a given problem, it is quite common to first consider a linear transformation on the system in such a way that the resulting differential equation can be more easily handled in a certain sense, to be specified for each problem. To illustrate the procedure, let us consider a simple example.

Example. Suppose we have the 2×2 matrix

$$A(t) = \alpha(t)X_1 + \beta(t)X_2, \tag{84}$$

where X_1 and X_2 are given by (76) and α and β are complex functions of time, $\alpha, \beta : \mathbb{R} \longrightarrow \mathbb{C}$. Then the exact solution of $Y' = A(t)Y$, $Y(0) = I$ is

$$Y(t) = \begin{pmatrix} \mathrm{e}^{\int_0^t \alpha(s)\mathrm{d}s} & \displaystyle\int_0^t \mathrm{d}s_1 \mathrm{e}^{\int_{s_1}^t \alpha(s_2)\mathrm{d}s_2}\,\beta(s_1)\,\mathrm{e}^{-\int_0^{s_1}\alpha(s_2)\mathrm{d}s_2} \\ 0 & \mathrm{e}^{-\int_0^t \alpha(s)\mathrm{d}s} \end{pmatrix}. \tag{85}$$

Let us factorize the solution as $Y(t) = \tilde{Y}_0(t)\tilde{Y}_1(t)$, with $\tilde{Y}_0(t)$ the solution of the initial value problem defined by

$$\tilde{Y}_0' = A_0(t)\tilde{Y}_0 \qquad A_0(t) = \alpha(t)X_1 = \begin{pmatrix} \alpha(t) & 0 \\ 0 & -\alpha(t) \end{pmatrix} \tag{86}$$

and $\tilde{Y}_0(0) = I$. Then, the equation satisfied by $\tilde{Y}_1$ is

$$\tilde{Y}_1' = A_1(t)\tilde{Y}_1, \quad \text{with } A_1 = \tilde{Y}_0^{-1}A\,\tilde{Y}_0 - A_0 = \begin{pmatrix} 0 & \beta(t)\mathrm{e}^{2\int_0^t \alpha(s)\mathrm{d}s} \\ 0 & 0 \end{pmatrix}, \tag{87}$$

so that the first term of the Magnus expansion applied to (87) already provides the exact solution (85). □

This, of course, is not the typical behavior, but in any case, if the transformation $\tilde{Y}_0$ in the factorization $Y(t) = \tilde{Y}_0(t)\tilde{Y}_1(t)$ is chosen appropriately, the first few terms in the Magnus series applied to the equation satisfied by $\tilde{Y}_1$ usually give very accurate approximations.

Since this kind of preliminary transformation is frequently used in Quantum Mechanics, we specialize the treatment to this particular setting here, and consider Eq. (4) instead. In other words, we write (30) in the more conventional form of the time dependent Schödinger equation

$$\frac{\mathrm{d}U(t)}{\mathrm{d}t} = \tilde{H}(t)U(t), \tag{88}$$

where $\tilde{H} \equiv H/(\mathrm{i}\hbar)$, $\hbar$ is the reduced Planck constant, H is the Hamiltonian and U corresponds to the evolution operator.

As in the example, suppose that $\tilde{H}$ can be split into two pieces, $\tilde{H} = \tilde{H}_0 + \varepsilon\tilde{H}_1$, with $\tilde{H}_0$ a solvable Hamiltonian and $\varepsilon \ll 1$ a small perturbation parameter. In such a situation, one tries to integrate out the $\tilde{H}_0$ piece so as to circumscribe the approximation to the $\tilde{H}_1$ piece. In the case of Eq. (88), this is carried out by means of a linear time-dependent transformation. In Quantum Mechanics, this preliminary linear transformation corresponds to a new evolution picture, such as the interaction or the adiabatic picture.

Among other possibilities, we may factorize the time-evolution operator as

$$U(t) = G(t)U_G(t)G^\dagger(0), \tag{89}$$

where $G(t)$ is a linear transformation whose purpose is yet to be defined. In the new G-Picture, the corresponding time-evolution operator U_G obeys the equation

$$U_G'(t) = \tilde{H}_G(t)U_G(t), \qquad \tilde{H}_G(t) = G^\dagger(t)\tilde{H}(t)G(t) - G^\dagger(t)G'(t). \tag{90}$$

The choice of G depends on the nature of the problem at hand. There is no generic formal recipe to find out the most appropriate G. In the spirit of canonical transformations of Classical Mechanics, one should build up the very U_G

perturbatively. However, the aim here is different because G is defined from the beginning. Two rather common choices are:

- **Interaction picture**. It is well suited when $\tilde{H}_0(t)$ is diagonal in some basis, or else, it is constant. In that case

$$G(t) = \exp\left(\int_0^t \tilde{H}_0(\tau)\mathrm{d}\tau\right) \tag{91}$$

so that

$$\tilde{H}_G(t) = \varepsilon \exp\left(-\int_0^t \tilde{H}_0(\tau)\mathrm{d}\tau\right)\tilde{H}_1(t)\exp\left(\int_0^t \tilde{H}_0(\tau)\mathrm{d}\tau\right). \tag{92}$$

- **Adiabatic picture**. A time scale of the system, much smaller than that of the interaction, defines an adiabatic regime. For instance, suppose that the Hamiltonian operator $H(t)$ depends smoothly on time through the variable $\tau = t/T$, where T determines the time scale and $T \to \infty$. Then the quantum mechanical evolution of the system is described by $\mathrm{d}U/\mathrm{d}t = \tilde{H}(\varepsilon t)U$, with $\varepsilon \equiv 1/T \ll 1$, or equivalently

$$\frac{\mathrm{d}U(\tau)}{\mathrm{d}\tau} = \frac{1}{\varepsilon}\tilde{H}(\tau)U(\tau), \tag{93}$$

with $\tau \equiv \varepsilon t$. In this case the appropriate transformation is a G that renders $\tilde{H}(t)$ instantaneously diagonal, i.e.,

$$G^\dagger(t)\tilde{H}(t)G(t) = E(t) = \mathrm{diag}[E_1(t), E_2(t), \ldots]. \tag{94}$$

The term $G^\dagger G'$ of the new Hamiltonian in (90) is, under adiabatic conditions, very small. Its main diagonal generates the so-called Berry, or geometric, phase [82].

Both types of G do not exclude mutually, but they may be used in succession. As a matter of fact, corrections to the adiabatic approximation must be followed by the former one. In turn, an adiabatic transformation may be iterated, as proposed by Garrido [83] and Berry [84].

In Section 4 we shall extensively use these preliminary linear transformations on several standard problems of Quantum Mechanics to illustrate the practical features of the Magnus expansion.

2.10. Exponential product representations

In contrast to Magnus expansion, much less attention has been paid to solutions of (30) in the form of a product of exponential operators. Both approaches are by no means equivalent, since, in general, the operators Ω_n do not commute with each other. For instance, for a quantum system as in Eq. (88), the ansatz $U = \prod \exp(\Phi_n)$ (where Φ_n are skew-Hermitian operators to be determined) is an alternative to the Magnus expansion, also preserving the unitarity of the time-evolution operator. One such procedure was devised by Fer in 1958 in a paper devoted to the study of systems of linear differential equations [20]. Although the original result obtained by Fer was cited and explicitly stated by Bellman [85, p. 204], sometimes it has been misquoted as a reference for the Magnus expansion [22]. On the other hand, Wilcox associated Fer's name with an interesting alternative infinite product expansion which is indeed a continuous analogue of the Zassenhaus formula [19] (something also attributed to the Fer factorization [2, p. 372]). This, however, also led to some confusion, since his approach is in the spirit of perturbation theory, whereas Fer's original one was essentially nonperturbative. The situation was clarified in [86], where also some applications to Quantum Mechanics were carried out for the first time.

In this section we briefly discuss the main features of the Fer and Wilcox expansions, and how the latter can be derived from the successive terms of the Magnus series. This will clarify the different character of the two expansions. We also include some details on the factorization of the solution proposed by Wei and Norman [17,87]. Finally we provide another interpretation of the Magnus expansion as the continuous analogue of the BCH formula in linear control theory.

2.10.1. Fer method

An intuitive way to introduce Fer formalism is the following [59]. Given the matrix linear system $Y' = A(t)Y$, $Y(0) = I$, we know that

$$Y(t) = \exp(F_1(t)) \tag{95}$$

is the exact solution if A commutes with its time integral $F_1(t) = \int_0^t A(s)\mathrm{d}s$, and $Y(t)$ evolves in the Lie group $\mathcal{G}$ if A lies in its corresponding Lie algebra $\mathfrak{g}$. If the goal is to respect the Lie-group structure in the general case, we need to 'correct' (95) without losing this important feature.

Two possible remedies arise in a quite natural way. The first is just to seek a correction $\Delta(t)$ evolving in the Lie algebra $\mathfrak{g}$ so that

$$Y(t) = \exp\left(F_1(t) + \Delta(t)\right).$$

This is nothing but the Magnus expansion. Alternatively, one may correct with $Y_1(t)$ in the Lie group $\mathcal{G}$,

$$Y(t) = \exp(F_1(t))Y_1(t). \tag{96}$$

This is precisely the approach pursued by Fer, i.e. representing the solution of (30) in the factorized form (96), where (hopefully) Y_1 will be closer to the identity matrix than Y, at least for small t.

The question now is to find the differential equation satisfied by Y_1. Substituting (96) into Eq. (30) we have

$$\frac{\mathrm{d}}{\mathrm{d}t}Y = \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{F_1}\right)Y_1 + \mathrm{e}^{F_1}\frac{\mathrm{d}}{\mathrm{d}t}Y_1 = A\mathrm{e}^{F_1}Y_1, \tag{97}$$

so that, taking into account the expression for the derivative of the exponential map (Lemma 2), one arrives easily at

$$Y_1' = A_1(t)Y_1 \qquad Y_1(0) = I, \tag{98}$$

where

$$A_1(t) = \mathrm{e}^{-F_1}A\mathrm{e}^{F_1} - \int_0^1 \mathrm{d}x\, \mathrm{e}^{-xF_1}A\mathrm{e}^{xF_1}. \tag{99}$$

The above procedure can be repeated to yield a sequence of iterated matrices A_k. After n steps we have the following recursive scheme, known as the Fer expansion:

$$\begin{aligned} &Y = \mathrm{e}^{F_1}\mathrm{e}^{F_2}\cdots\mathrm{e}^{F_n}Y_n \\ &Y_n' = A_n(t)Y_n \qquad Y_n(0) = I, \qquad n = 1, 2, \ldots \end{aligned} \tag{100}$$

with $F_n(t)$ and $A_n(t)$ given by

$$\begin{aligned} F_{n+1}(t) &= \int_0^t A_n(s)\mathrm{d}s \qquad A_0(t) = A(t), \quad n = 0, 1, 2, \ldots \\ A_{n+1} &= \mathrm{e}^{-F_{n+1}}A_n\mathrm{e}^{F_{n+1}} - \int_0^1 \mathrm{d}x\, \mathrm{e}^{-xF_{n+1}}A_n\mathrm{e}^{xF_{n+1}} \\ &= \int_0^1 \mathrm{d}x \int_0^x \mathrm{d}u\, \mathrm{e}^{-(1-u)F_{n+1}}\,[A_n, F_{n+1}]\,\mathrm{e}^{(1-u)F_{n+1}} \\ &= \sum_{j=1}^{\infty} \frac{(-)^j j}{(j+1)!}\mathrm{ad}_{F_{n+1}}^j(A_n), \quad n = 0, 1, 2\ldots. \end{aligned} \tag{101}$$

When, after n steps we impose $Y_n = I$, we are left with an approximation to the exact solution $Y(t)$.

Inspection of the expression of A_{n+1} in (101) reveals an interesting feature of the Fer expansion. If we assume that a perturbation parameter ε is introduced in A (i.e. if we substitute A by εA in the formalism), since F_{n+1} is of the same order in ε as A_n then an elementary recursion shows that the matrix A_n starts with a term of order ε^{2^n} (correspondingly the operator F_n contains terms of order $\varepsilon^{2^{n-1}}$ and higher). This should greatly enhance the rate of convergence of the product in Eq. (100) to the exact solution.

It is possible to derive a bound on the convergence domain in time of the expansion [30]. The idea is just to look for conditions on $A(t)$ which ensure $F_n \to 0$ as $n \to \infty$. As in the case of the Magnus expansion, we take $A(t)$ to be a bounded matrix with $\|A(t)\| \leq k(t) \equiv k_0(t)$. Fer's algorithm, Eqs. (100) and (101), then provides a recursive relation among corresponding bounds $k_n(t)$ for $\|A_n(t)\|$. If we denote $K_n(t) \equiv \int_0^t k_n(s)\mathrm{d}s$, we can write this relation in the generic form $k_{n+1} = f(k_n, K_n)$, which, after integration, gives

$$K_{n+1} = M(K_n). \tag{102}$$

The question now is: when is $K_n \to 0$ as $n \to \infty$? This is certainly so if 0 is a stable fixed point for the iteration of the mapping M and K_0 is within its basin of attraction. To see when this is the case, we have to solve the equation $\xi = M(\xi)$ to find where the next fixed point lies. Let us do it explicitly. By taking norms in the recursive scheme (101) we have

$$\|A_{n+1}\| \leq \int_0^1 \mathrm{d}x \int_0^x \mathrm{d}u\, \mathrm{e}^{2(1-u)K_n}\,\|[A_n, F_{n+1}]\|\,,$$

which can be written as $\|A_{n+1}\| \leq k_{n+1}$, with

$$k_{n+1} = \frac{1 - \mathrm{e}^{2K_n}(1 - 2K_n)}{2K_n}\frac{\mathrm{d}K_n}{\mathrm{d}t}$$

and consequently K_{n+1} is given by Eq. (102) with

$$M(K_n) = \int_0^{K_n} \frac{1 - \mathrm{e}^{2x}(1-2x)}{2x}\mathrm{d}x. \tag{103}$$

That is the mapping we have to iterate. It is clear that $\xi = 0$ is a stable fixed point of M. The next, unstable, fixed point is $\xi = 0.8604065$. So we can conclude that we have a convergent Fer expansion, at least for values of time t, such that

$$\int_0^t \|A(s)\|\mathrm{d}s \leq K_0(t) < 0.8604065. \tag{104}$$

Notice that the bound for the convergence domain provided by this result is smaller than that corresponding to the Magnus expansion (Theorem 9).

2.10.2. Wilcox method

A more tractable form of infinite product expansion has been devised by Wilcox [19] in analogy with the Magnus approach. The idea, as usual, is to treat ε in

$$Y' = \varepsilon A(t)Y, \qquad Y(0) = I \tag{105}$$

as an expansion parameter and to determine the successive factors in the product

$$Y(t) = \mathrm{e}^{W_1}\,\mathrm{e}^{W_2}\,\mathrm{e}^{W_3}\cdots \tag{106}$$

by assuming that W_n is exactly of order ε^n. Hence, it is clear from the very beginning that the methods of Fer and Wilcox give rise indeed to completely different infinite product representations of the solution $Y(t)$.

The explicit expressions of W_1, W_2 and W_3 are given in [19]. It is noteworthy that the operators W_n can be expressed in terms of Magnus operators Ω_k, for which compact formulae and recursive procedures are available. To this end, we simply use the Baker–Campbell–Hausdorff formula to extract formally from the identity

$$\mathrm{e}^{W_1}\,\mathrm{e}^{W_2}\,\mathrm{e}^{W_3}\cdots = \mathrm{e}^{\Omega_1+\Omega_2+\Omega_3+\cdots}, \tag{107}$$

terms of the same order in ε. After a straightforward calculation one finds for the first terms

$$W_1 = \Omega_1, \qquad W_2 = \Omega_2, \qquad W_3 = \Omega_3 - \frac{1}{2}[\Omega_1, \Omega_2], \tag{108}$$

$$W_4 = \Omega_4 - \frac{1}{2}[\Omega_1, \Omega_3] + \frac{1}{6}[\Omega_1, [\Omega_1, \Omega_2]], \quad \text{etc.} \tag{109}$$

The main interest of the Wilcox formalism stems from the fact that it provides explicit expressions for the successive approximations to a solution represented as an infinite product of exponential operators. This offers a useful alternative to the Fer expansion whenever the computation of F_n from Eq. (101) is too cumbersome. We note, in passing, that to first order the three expansions yield the same result: $F_1 = W_1 = \Omega_1$.

2.10.3. Wei–Norman factorization

Now suppose that A and Y in Eq. (30) are linear operators and that $A(t)$ can be expressed in the form

$$A(t) = \sum_{i=1}^{m} u_i(t)X_i, \quad m \text{ finite}, \tag{110}$$

where the $u_i(t)$ are scalar functions of time, and $X_1, X_2, \ldots, X_m$ are time-independent operators. Furthermore, suppose that the Lie algebra $\mathfrak{g}$ generated by the X_i is of finite dimension l (this is obviously true if A and Y are finite dimensional matrix operators). Under these conditions, if $X_1, X_2, \ldots, X_l$ is a basis for $\mathfrak{g}$, the Magnus expansion allows us to express the solution locally in the form $Y(t) = \exp(\sum_{i=1}^{l} f_i(t)X_i)$. Wei and Norman, on the other hand, show that there exists a neighborhood of $t = 0$ in which the solution can be written as a product [17,87]

$$Y(t) = \exp(g_1(t)X_1)\exp(g_2(t)X_2)\cdots\exp(g_l(t)X_l), \tag{111}$$

where the $g_i(t)$ are scalar functions of time. Moreover, the $g_i(t)$ satisfy a set of nonlinear differential equations which depend only on the Lie algebra $\mathfrak{g}$ and the $u_i(t)$'s. These authors also study the conditions under which the solution converges globally, that is, for all t. In particular, this happens for all solvable Lie algebras in a suitable basis, and for any real 2×2 system of equations [87].

In the terminology of Lie algebras and Lie groups, the representation $Y(t) = \exp(\sum_{i=1}^{l} f_i(t)X_i)$ corresponds to the *canonical coordinates of the first kind*, whereas Eq. (111) defines a system of *canonical coordinates of the second kind* [41, 78,88].

This class of factorization has been used in combination with the Fer expansion to obtain closed-form solutions of the Cauchy problem defined by certain classes of parabolic linear partial differential equations [89]. When the algorithm is applied, the solution is written as a finite product of exponentials, depending on certain ordering functions for which convergent approximations are constructed in explicit form.

Notice that the representation (111) is clearly useful when the spectral properties of the individual operators X_i are readily available. Since the X_i are constant and often have simple physical interpretation, the evaluation of the eigenvalues and eigenvectors can be done once for all times, and this may facilitate the computation of the exponentials. This situation arises, in particular, in control theory [88]. The functions $u_i(t)$ are known as the controls, and the operator $Y(t)$ acts on the states of the system, describing how the states are transformed along time.

2.11. The continuous BCH formula

When applied to the equation $Y' = A(t)Y$ with the matrix $A(t)$ given by (110), the Magnus expansion adopts a particularly simple form. Furthermore, by making use of the structure constants of the Lie algebra, it is relatively easy to get explicit expressions for the canonical coordinates of the first kind $f_i(t)$. Let us illustrate the procedure by considering the particular case

$$A(t) = u_1(t)X_1 + u_2(t)X_2.$$

Denoting by $\alpha_i(t) = \int_0^t u_i(s)\mathrm{d}s$, and, for a given function μ,

$$\left(\int_i \mu\right)(t) \equiv \int_0^t u_i(s)\mu(s)\mathrm{d}s,$$

a straightforward calculation shows that the first terms of Ω in the Magnus expansion can be written as

$$\Omega(t) = \beta_1(t)X_1 + \beta_2(t)X_2 + \beta_{12}(t)[X_1, X_2] + \beta_{112}(t)[X_1, [X_1, X_2]] + \beta_{212}(t)[X_2, [X_1, X_2]] + \cdots \tag{112}$$

where

$$\begin{aligned}
&\beta_i = \alpha_i, \quad i = 1, 2,\\
&\beta_{12} = \frac{1}{2}\left(\int_1 \alpha_2 - \int_2 \alpha_1\right),\\
&\beta_{112} = \frac{1}{12}\left(\int_2 \alpha_1^2 - \int_1 \alpha_1\alpha_2\right) - \frac{1}{4}\left(\int_1\int_2 \alpha_1 - \int_1\int_1 \alpha_2\right),\\
&\beta_{212} = \frac{1}{12}\left(\int_2 \alpha_1\alpha_2 - \int_1 \alpha_2^2\right) + \frac{1}{4}\left(\int_2\int_1 \alpha_2 - \int_2\int_2 \alpha_1\right).
\end{aligned} \tag{113}$$

Taking into account the structure constants of the particular finite dimensional Lie algebra under consideration, from (112) one easily gets the functions $f_i(t)$. In the general case, (112) allows us to express Ω as a linear combination of elements of a basis of the free Lie algebra generated by X_1 and X_2. In this case, the recurrence (49)–(50) defining the Magnus expansion can be carried out only with the nested integrals

$$\alpha_{i_1\cdots i_s}(t) \equiv \left(\int_{i_s}\cdots\int_{i_1} 1\right)(t) = \int_0^t \int_0^{t_s}\cdots\int_0^{t_3}\int_0^{t_2} u_{i_s}(t_s)\cdots u_{i_1}(t_1)\mathrm{d}t_1\cdots\mathrm{d}t_s \tag{114}$$

involving the functions $u_1(t)$ and $u_2(t)$. Thus, for instance, the coefficients in (113) can be written (after successive integration by parts) as

$$\begin{aligned}
&\beta_i = \alpha_i, \quad i = 1, 2,\\
&\beta_{12} = \frac{1}{2}\left(\int_1 \alpha_2 - \int_2 \alpha_1\right) = \frac{1}{2}\left(\alpha_{21} - \alpha_{12}\right),\\
&\beta_{112} = \frac{1}{6}\left(\int_2\int_1 \alpha_1 + \int_1\int_1 \alpha_2 - 2\int_1\int_2 \alpha_1\right) = \frac{1}{6}\left(\alpha_{112} + \alpha_{211} - 2\alpha_{121}\right),\\
&\beta_{212} = \frac{1}{6}\left(2\int_2\int_1 \alpha_2 - \int_2\int_2 \alpha_1 - \int_1\int_2 \alpha_2\right) = \frac{1}{6}\left(2\alpha_{212} - \alpha_{122} - \alpha_{221}\right).
\end{aligned}$$

The series (112) expressed in terms of the integrals (114) is usually referred to as the continuous Baker–Campbell–Hausdorff formula [90,91] for the linear case. We will generalize this formalism to the nonlinear case in the next section.

3. Generalizations of the Magnus expansion

In view of the attractive properties of the Magnus expansion as a tool to construct approximate solutions of non-autonomous systems of linear ordinary differential equations, it is hardly surprising that several attempts have been made

along the years, either to extend the procedure to a more general setting or to manipulate the series to achieve further improvements. In this section, we review some of these generalizations, with special emphasis on the treatment of nonlinear differential equations.

First we reconsider an iterative method originally devised by Voslamber [92] for computing approximations $\Omega^{(n)}(t)$ in $Y(t) = \exp(\Omega(t))$ for the linear equation $Y' = A(t)Y$. The resulting approximation may be interpreted as a resummation of terms in the Magnus series and possesses interesting features not shared by the corresponding truncation of the conventional Magnus expansion. Then we adapt the Magnus expansion to the physically relevant case of a periodic matrix $A(t)$ with period T which incorporates, in a natural way, the structure of the solution ensured by the Floquet theorem. Next we go one step further and generalize the Magnus expansion to the so-called nonlinear Lie equation $Y' = A(t, Y)Y$. Finally, we show how the procedure can be applied to *any* nonlinear explicitly time-dependent differential equation. Although the treatment is largely formal, in Section 5 we will see that it is of paramount importance for designing new and highly efficient numerical integration schemes for this class of differential equations. We particularize the treatment to the important case of Hamiltonian systems, and also establish an interesting connection with the Chen–Fliess series for nonlinear differential equations.

3.1. Voslamber iterative method

Let us consider Eq. (30) when there is a perturbation parameter ε in the (in general, complex) matrix A, i.e., Eq. (105). Theorem 9 guarantees that, for sufficiently small values of t, $Y(t;\varepsilon) = \exp \Omega(t;\varepsilon)$, where

$$\Omega(t;\varepsilon) = \sum_{n=1}^{\infty} \varepsilon^n \Omega_n(t). \tag{115}$$

The advantages of this representation and the approximations obtained when the series is truncated have been sufficiently recognized in the treatment done in previous sections. There is, however, a property of the exact solution not shared by any truncation of the series (115) which could be relevant in certain physical applications: $(1/\varepsilon)\sum_{n=1}^{m} \varepsilon^n \Omega_n(t)$ with $m > 1$ is unbounded for $\varepsilon \to \infty$ even when $\Omega(t,\varepsilon)/\varepsilon$ is bounded uniformly with respect to ε under rather general assumptions on the matrix $A(t)$ [92]. Notice that this is the case, in particular, for the adiabatic problem (93).

When Schur's unitary triangularization theorem [44] is applied to the exact solution $Y(t;\varepsilon)$ one has

$$T_\varepsilon = U^\dagger Y U, \tag{116}$$

where T_ε is an upper triangular matrix and U is unitary. In other words, Y is unitarily equivalent to an upper triangular matrix T_ε. Differentiating (116) and using (105), one arrives at

$$T_\varepsilon' = \varepsilon U^\dagger A U T_\varepsilon + \left[T_\varepsilon, U^\dagger U'\right].$$

Since the second term on the right hand side is not upper triangular, it follows at once that

$$T_\varepsilon(t) = \exp\left(\varepsilon \int_0^t (U^\dagger A U)_\triangle \mathrm{d}s\right),$$

where the subscript $\triangle$ denotes the upper triangular part (including terms on the main diagonal) of the corresponding matrix. Taking into account (116) one gets

$$\Omega(t,\varepsilon) = \varepsilon U \left(\int_0^t (U^\dagger A U)_\triangle \mathrm{d}s\right) U^\dagger. \tag{117}$$

Considering now the Frobenius norm (which is unitarily invariant, Section 1.2) of both sides of this equation, one has

$$\begin{aligned} \|\Omega\|_F &= |\varepsilon| \left\| \int_0^t (U^\dagger A U)_\triangle \mathrm{d}s \right\|_F \le |\varepsilon| \int_0^t \|(U^\dagger A U)_\triangle\|_F \mathrm{d}s \\ &\le |\varepsilon| \int_0^t \|U^\dagger A U\|_F \mathrm{d}s = |\varepsilon| \int_0^t \|A\|_F \mathrm{d}s. \end{aligned} \tag{118}$$

If the spectral norm is considered instead, from inequalities (28), (29) and (118), one concludes that

$$\|\Omega\|_2 \le \sqrt{\mathrm{rank}(A)}\, |\varepsilon| \int_0^t \|A\|_2 \mathrm{d}s.$$

In any case, what is important to stress here is that for the exact solution $\Omega(t;\varepsilon)/\varepsilon$ is bounded uniformly with respect to the ε parameter. Voslamber proceeds by deriving an algorithm for generating successive approximations of $Y(t;\varepsilon) = \exp(\Omega(t;\varepsilon))$ which, contrarily to the direct series expansion (115), preserve this property. His point of departure is to get a series expansion for the so-called *dressed derivative of* Ω [93]

$$\Gamma \equiv \mathrm{e}^{\Omega/2}\, \Omega'\, \mathrm{e}^{-\Omega/2}. \tag{119}$$

This is accomplished by inserting (39) in (119). Specifically, one has

$$\begin{aligned}\Gamma &= \mathrm{e}^{\mathrm{ad}_{\Omega/2}}\,\Omega' = \mathrm{e}^{\mathrm{ad}_{\Omega/2}}\mathrm{d}\exp_{\Omega}^{-1}(\varepsilon A) = \mathrm{e}^{\mathrm{ad}_{\Omega/2}}\,\frac{\mathrm{ad}_{\Omega}}{\mathrm{e}^{\mathrm{ad}_{\Omega}}-1}(\varepsilon A)\\ &= \frac{\mathrm{ad}_{\Omega/2}}{\sinh\Omega/2}(\varepsilon A) = \sum_{n=0}^{\infty}\frac{B_n(1/2)}{n!}\mathrm{ad}_{\Omega}^{n}(\varepsilon A)\end{aligned}$$

and finally [92,93]

$$\Gamma = \sum_{n=0}^{\infty}\frac{2^{1-n}-1}{n!}B_n\,\mathrm{ad}_{\Omega}^{n}(\varepsilon A), \tag{120}$$

where, as usual, B_n denote Bernoulli numbers. In order to express Γ as a power series of ε one has to insert the Magnus series (115) into Eq. (120). Then we get

$$\Gamma(t;\varepsilon) = \sum_{n=1}^{\infty}\varepsilon^n\,\Gamma_n(t), \tag{121}$$

where the terms Γ_n can be expressed as a function of Ω_k with $k \le n-2$ through the recursive procedure [93]

$$\begin{aligned}&\Gamma_1 = A, \qquad \Gamma_2 = 0, \\ &\Gamma_n = \sum_{j=2}^{n-1} c_j \sum_{\substack{k_1+\cdots+k_j=n-1\\ k_1\ge 1,\ldots,k_j\ge 1}} \mathrm{ad}_{\Omega_{k_1}}\,\mathrm{ad}_{\Omega_{k_2}}\cdots\mathrm{ad}_{\Omega_{k_j}}A, \quad n\ge 3.\end{aligned} \tag{122}$$

Here

$$c_j \equiv \frac{2^{1-j}-1}{j!}B_j,$$

with $c_{2j+1}=0$, $c_2=-1/24$, $c_4=7/5760$, etc. In particular,

$$\begin{aligned}\Gamma_3 &= -\frac{1}{24}[\Omega_1,[\Omega_1,A]]\\ \Gamma_4 &= -\frac{1}{24}([\Omega_1,[\Omega_2,A]]+[\Omega_2,[\Omega_1,A]]).\end{aligned}$$

Now, from the definition of Γ, Eq. (119), we write

$$\Omega' = \mathrm{e}^{-\Omega/2}\,\Gamma\,\mathrm{e}^{\Omega/2},$$

which, after integration over t, can be used for constructing successive approximations to Ω once the terms Γ_n are known in terms of Ω_k, $k\le n-2$. Thus, the nth approximant $\Omega^{(n)}$ is defined by

$$\Omega^{(n)}(t) = \int_0^t \mathrm{e}^{-\frac{1}{2}\Omega^{(n-1)}(s)}\Gamma^{(n)}(s)\mathrm{e}^{\frac{1}{2}\Omega^{(n-1)}(s)}\mathrm{d}s, \quad n=1,2,\ldots \tag{123}$$

where the ε dependence has been omitted by simplicity and $\Gamma^{(n)} = \sum_{k=1}^{n}\varepsilon^k\Gamma_k$, $\Omega^{(0)} = 0$. The first two approximants explicitly read

$$\begin{aligned}\Omega^{(1)}(t,\varepsilon) &= \varepsilon\,\Omega_1(t) = \varepsilon\int_0^t A(s)\mathrm{d}s\\ \Omega^{(2)}(t,\varepsilon) &= \varepsilon\int_0^t \mathrm{e}^{-\frac{1}{2}\Omega^{(1)}(s,\varepsilon)}A(s)\mathrm{e}^{\frac{1}{2}\Omega^{(1)}(s,\varepsilon)}\mathrm{d}s.\end{aligned} \tag{124}$$

In this approach, the solution is approximated by $Y(t) \simeq \exp(\Omega^{(n)})$. Observe that $\Omega^{(n)}$ contains contributions from an infinity of orders in ε, whereas the nth term in the Magnus series (115) is proportional to ε^n. Furthermore, $\Omega^{(n)}$ contains $\sum_{k=1}^{n}\varepsilon^k\Omega_k$ *and* also higher powers ε^m $(m>n)$. In particular, one easily gets

$$\Omega^{(2)}(t;\varepsilon) = \varepsilon\,\Omega_1(t) + \varepsilon^2\,\Omega_2(t) + \sum_{k=3}^{\infty}\frac{(-1)^{k-1}}{2^{k-1}(k-1)!}\varepsilon^k\int_0^t \mathrm{ad}_{\Omega_1(s)}^{k-1}A(s)\mathrm{d}s.$$

From the structure of the expression (123) it is also possible to find the asymptotic behavior of $\Omega^{(n)}/\varepsilon$ $(n\ge 3)$ for $\varepsilon\to\infty$ and prove that it remains bounded [92], just as the exact solution does. This property of the Voslamber iterative algorithm may lead to better approximations of $Y(t)$ when the parameter ε is not very small, since in that case $\Omega^{(n)}/\varepsilon$ is expected to remain close to Ω/ε, as shown in [93].

3.2. Floquet–Magnus expansion

We now turn our attention to a specific case of Eq. (30) with important physical and mathematical applications, namely when the (complex) $n \times n$ matrix-valued function $A(t)$ is periodic with period T. Then further information is available on the structure of the exact solution, as is given by the celebrated Floquet theorem, which ensures the factorization of the solution in a periodic part and a purely exponential factor. More specifically,

$$Y(t) = P(t)\exp(tF), \tag{125}$$

where F and P are $n \times n$ matrices, $P(t) = P(t+T)$ for all t and F is constant. Thus, albeit a solution of (30) is not, in general, periodic, the departure from periodicity is determined by (125). This result, when applied in quantum mechanics, is referred to as Bloch wave theory [94,95]. It is widely used in problems of solid state physics where space-periodic potentials are quite common. In Nuclear Magnetic Resonance, this structure is exploited as far as either time-dependent periodic magnetic fields or sample spinning are involved [24]. Asymptotic stability of the solution $Y(t)$ is dictated by the nature of the eigenvalues of F, the so-called characteristic exponents of the original periodic system [96].

An alternative manner of interpreting Eq. (125) is to consider the piece $P(t)$, provided it is invertible, to perform a transformation of the solution in such a way that the coefficient matrix corresponding to the new representation has all its matrix entries given by constants. Thus the piece $\exp(tF)$ in (125) may be considered as an exact solution of the system (30) previously moved to a representation where the coefficient matrix is the constant matrix F [76]. The t-dependent change of representation is carried out by $P(t)$. Connecting with Section 2.9, $P(t)$ is the appropriate preliminary linear transformation for periodic systems. Of course, Floquet theorem by itself gives no practical information about this procedure. It just states that such a representation does exist. In fact, a serious difficulty in the study of differential equations with periodic coefficients is that no general method to compute either the matrix $P(t)$ or the eigenvalues of F is known.

Mainly, two ways of exploiting the above structure of $Y(t)$ are found in the literature [97]. The first one consists in performing a Fourier expansion of the formal solution, leading to an infinite system of linear differential equations with constant coefficients. Thus, the t-dependent finite system is replaced with a constant one at the price of handling infinite dimension. Resolution of the truncated system furnishes an approximate solution. The second approach is of perturbative nature. It deals directly with the form (125) by expanding

$$P(t) = \sum_{n=1}^{\infty} P_n(t), \qquad F = \sum_{n=1}^{\infty} F_n. \tag{126}$$

Every term F_n in (126) is fixed so as to ensure $P_n(t) = P_n(t+T)$, which in turn guarantees the Floquet structure (125) at any order of approximation.

Although the Magnus expansion, such as it has been formulated in this work, does not explicitly provide the structure of the solution ensured by Floquet theorem, it can be adapted, without special difficulty, to also cope with this situation. The starting point is to introduce the Floquet form (125) into the differential equation $Y' = A(t)Y$. In that way, the evolution equation for P is obtained:

$$P'(t) = A(t)P(t) - P(t)F, \qquad P(0) = I. \tag{127}$$

The constant matrix F is also unknown and we will determine it so as to ensure $P(t+T) = P(t)$. Now we replace the usual perturbative scheme in Eq. (126) with the exponential ansatz

$$P(t) = \exp(\Lambda(t)), \qquad \Lambda(0) = O. \tag{128}$$

Obviously, $\Lambda(t+T) = \Lambda(t)$ so as to preserve periodicity. Now Eq. (127) conveys

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(\Lambda) = A\exp(\Lambda) - \exp(\Lambda)F, \tag{129}$$

from which, as with the conventional Magnus expansion, it follows readily that

$$\Lambda' = \sum_{k=0}^{\infty} \frac{B_k}{k!}\mathrm{ad}_{\Lambda}^{k}\,(A + (-1)^{k+1}F). \tag{130}$$

This equation is now, in the Floquet context, the analogue of Magnus equation (39). Notice that if we put $F = O$ then (39) is recovered. The next move is to consider the series expansions for Λ and F

$$\Lambda(t) = \sum_{k=1}^{\infty} \Lambda_k(t), \qquad F = \sum_{k=1}^{\infty} F_k, \tag{131}$$

with $\Lambda_k(0) = O$, for all k. Equating terms of the same order in (130) one gets the successive contributions to the series (131). Therefore, the explicit ansatz we are propounding reads

$$Y(t) = \exp\left(\sum_{k=1}^{\infty} \Lambda_k(t)\right)\exp\left(t\sum_{k=1}^{\infty} F_k\right). \tag{132}$$

This can be properly referred as the *Floquet–Magnus expansion*.

Substituting the expansions of Eq. (131) into (130) and equating terms of the same order, one can write

$$\Lambda_n' = \sum_{j=0}^{n-1} \frac{B_j}{j!} \left(W_n^{(j)}(t) + (-1)^{j+1} T_n^{(j)}(t) \right) \quad (n \geq 1). \tag{133}$$

The terms $W_n^{(j)}(t)$ may be obtained by a similar recurrence to that given in Eq. (49)

$$\begin{aligned} W_n^{(j)} &= \sum_{m=1}^{n-j} \left[\Lambda_m, W_{n-m}^{(j-1)} \right] \quad (1 \leq j \leq n-1), \\ W_1^{(0)} &= A, \qquad W_n^{(0)} = O \quad (n > 1), \end{aligned} \tag{134}$$

whereas the terms $T_n^{(j)}(t)$ obey the recurrence relation

$$\begin{aligned} T_n^{(j)} &= \sum_{m=1}^{n-j} \left[\Lambda_m, T_{n-m}^{(j-1)} \right] \quad (1 \leq j \leq n-1), \\ T_n^{(0)} &= F_n \quad (n > 0). \end{aligned} \tag{135}$$

Every F_n is fixed by the condition $\Lambda_n(t+T) = \Lambda_n(t)$. An outstanding feature is that F_n can be determined independently of $\Lambda_n(t)$ as the solution $Y(t) = P(t)\exp(tF)$ shrinks to $Y(T) = \exp(TF)$. Consequently, the conventional Magnus expansion $Y(t) = \exp(\Omega(t))$ computed at $t = T$ must furnish

$$F_n = \frac{\Omega_n(T)}{T}, \quad \text{for all } n. \tag{136}$$

The first contributions to the Floquet–Magnus expansion read, explicitly,

$$\begin{aligned} \Lambda_1(t) &= \int_0^t A(x)\mathrm{d}x - tF_1, \\ F_1 &= \frac{1}{T}\int_0^T A(x)\mathrm{d}x, \\ \Lambda_2(t) &= \frac{1}{2}\int_0^t [A(x) + F_1, \Lambda_1(x)]\,\mathrm{d}x - tF_2, \\ F_2 &= \frac{1}{2T}\int_0^T [A(x) + F_1, \Lambda_1(x)]\,\mathrm{d}x. \end{aligned} \tag{137}$$

Moreover, from the recurrence relations (134) and (135) it is possible to obtain a sufficient condition such that convergence of the series $\sum \Lambda_n$ is guaranteed in the whole interval $t \in [0, T]$ [34]. In fact, one can show that absolute convergence of the Floquet–Magnus series is ensured at least if

$$\int_0^T \|A(t)\|\mathrm{d}t < \xi_F \equiv 0.20925. \tag{138}$$

Notice that convergence of the series $\sum F_n$ is already guaranteed by (136) and the discussion concerning the conventional Magnus expansion in Sections 2.7.2 and 2.7.3. The bound ξ_F in the periodic Floquet case turns out to be smaller than the corresponding bound $r_c = \pi$ in the conventional Magnus expansion. At first sight this could be understood as an impoverishment of the result. However, it has to be recalled that, due precisely to Floquet theorem, once the condition is fulfilled in one period, convergence is assured for any value of time. On the contrary, in the general Magnus case, the bound always gives a running condition.

3.3. Magnus expansions for nonlinear matrix equations

It is possible to extend the procedure leading to the Magnus expansion for the linear equation (30), and obtain approximate solutions for the nonlinear matrix equation

$$Y' = A(t, Y)Y, \qquad Y(0) = Y_0 \in \mathcal{G}, \tag{139}$$

where $\mathcal{G}$ is a matrix Lie group, $A : \mathbb{R}_+ \times \mathcal{G} \longrightarrow \mathfrak{g}$ and $\mathfrak{g}$ denotes the corresponding Lie algebra. Eq. (139) appears in relevant physical fields such as rigid body mechanics, in the calculation of Lyapunov exponents ($\mathcal{G} \equiv \mathrm{SO}(n)$) and other problems arising in Hamiltonian dynamics ($\mathcal{G} \equiv \mathrm{Sp}(n)$). In fact, it can be shown that every differential equation evolving on a matrix Lie group $\mathcal{G}$ can be written in the form (139) [59]. Moreover, the analysis of generic differential equations defined in homogeneous spaces can be reduced to the Lie-group Eq. (139) [98].

In [99] a general procedure for devising Magnus expansions for the nonlinear equation (139) is introduced. It is based on applying Picard's iteration on the associated differential equation in the Lie algebra and retaining in each iteration the terms necessary to increase the order, while maintaining the explicit character of the expansion. The resulting methods are thus explicit by design and are expressed in terms of integrals.

As usual, the starting point in the formalism is to represent the solution of (139) in the form

$$Y(t) = \exp(\Omega(t, Y_0))Y_0. \tag{140}$$

Then one obtains the differential equation satisfied by Ω:

$$\Omega' = \mathrm{d}\exp_{\Omega}^{-1}\left(A(t, \mathrm{e}^{\Omega}Y_0)\right), \qquad \Omega(0) = O, \tag{141}$$

where $\mathrm{d}\exp_{\Omega}^{-1}$ is given by (38). Now, as in the linear case, one can apply Picard's iteration to Eq. (141), giving instead

$$\begin{aligned}\Omega^{[m+1]}(t) &= \int_0^t \mathrm{d}\exp^{-1}_{\Omega^{[m]}(s)} A(s, \mathrm{e}^{\Omega^{[m]}(s)}Y_0)\mathrm{d}s \\ &= \int_0^t \sum_{k=0}^{\infty} \frac{B_k}{k!}\mathrm{ad}^k_{\Omega^{[m]}(s)}A(s, \mathrm{e}^{\Omega^{[m]}(s)}Y_0)\mathrm{d}s, \quad m \geq 0.\end{aligned}$$

The next step in getting explicit approximations is to truncate appropriately the $\mathrm{d}\exp^{-1}$ operator in the above expansion. Roughly speaking, when the whole series for $\mathrm{d}\exp^{-1}$ is considered, the power series expansion of the iterate function $\Omega^{[k]}(t)$, $k \geq 1$, only reproduces the expansion of the solution $\Omega(t)$ up to a certain order, say m. In consequence, the (infinite) power series of $\Omega^{[k]}(t)$ and $\Omega^{[k+1]}(t)$ differ in terms $\mathcal{O}(t^{m+1})$. The idea is then to discard in $\Omega^{[k]}(t)$ all terms of order greater than m. This of course requires careful analysis of each term in the expansion. For instance, $\Omega^{[0]} = O$ implies that $(\Omega^{[1]})' = A(t, Y_0)$ and therefore

$$\Omega^{[1]}(t) = \int_0^t A(s, Y_0)\mathrm{d}s = \Omega(t, Y_0) + \mathcal{O}(t^2).$$

Since

$$A(s, \mathrm{e}^{\Omega^{[1]}(s)}Y_0) = A(0, Y_0) + \mathcal{O}(s)$$

it follows at once that

$$-\frac{1}{2}\int_0^t [\Omega^{[1]}(s), A(s, \mathrm{e}^{\Omega^{[1]}(s)}Y_0)]\mathrm{d}s = \mathcal{O}(t^3).$$

When this second term in $\Omega^{[2]}(t)$ is included and $\Omega^{[3]}$ is computed, it turns out that $\Omega^{[3]}$ reproduces correctly the expression of $\Omega^{[2]}$ up to $\mathcal{O}(t^2)$. Therefore we truncate $\mathrm{d}\exp^{-1}$ at the $k = 0$ term and take

$$\Omega^{[2]}(t) = \int_0^t A(s, \mathrm{e}^{\Omega^{[1]}(s)}Y_0)\mathrm{d}s.$$

With greater generality, we let

$$\begin{aligned}\Omega^{[1]}(t) &= \int_0^t A(s, Y_0)\mathrm{d}s \\ \Omega^{[m]}(t) &= \sum_{k=0}^{m-2}\frac{B_k}{k!}\int_0^t \mathrm{ad}^k_{\Omega^{[m-1]}(s)}A(s, \mathrm{e}^{\Omega^{[m-1]}(s)}Y_0)\mathrm{d}s, \quad m \geq 2\end{aligned} \tag{142}$$

and take the approximation $\Omega(t) \approx \Omega^{[m]}(t)$. This results in an explicit approximate solution that involves a linear combination of multiple integrals of nested commutators, so that $\Omega^{[m]}(t) \in \mathfrak{g}$ for all $m \geq 1$. It can also be proved that $\Omega^{[m]}(t)$, once inserted in (140), provides an explicit approximation $Y^{[m]}(t)$ for the solution of (139) that is correct up to terms $\mathcal{O}(t^{m+1})$ [99]. In addition, $\Omega^{[m]}(t)$ reproduces exactly the sum of the first m terms in the Ω series of the usual Magnus expansion for the linear equation $Y' = A(t)Y$. It makes sense, then, to regard the scheme (142) as an explicit Magnus expansion for the nonlinear equation (139).

This procedure can be easily adapted to construct an exponential representation of the solution for the differential system

$$Y' = [A(t, Y), Y], \qquad Y(0) = Y_0 \in \mathrm{Sym}(n). \tag{143}$$

Here $\mathrm{Sym}(n)$ stands for the set of $n \times n$ symmetric real matrices and the (sufficiently smooth) function A maps $\mathbb{R}_+ \times \mathrm{Sym}(n)$ into $\mathfrak{so}(n)$, the Lie algebra of $n \times n$ real skew-symmetric matrices. It is well known that the solution itself remains in $\mathrm{Sym}(n)$ for all $t \geq 0$. Furthermore, the eigenvalues of $Y(t)$ are independent of time, i.e., $Y(t)$ has the same eigenvalues as Y_0. This remarkable qualitative feature of the system (143) is the reason why it is called an *isospectral flow*. Such flows have

several interesting applications in physics and applied mathematics, from molecular dynamics to micromagnetics to linear algebra [100].

Since $Y(t)$ and $Y(0)$ share the same spectrum, there exists a matrix function $Q(t) \in \mathrm{SO}(n)$ such that $Y(t)Q(t) = Q(t)Y(0)$ or, equivalently,

$$Y(t) = Q(t)Y_0Q^{\mathrm{T}}(t). \tag{144}$$

Then, by inserting (144) into (143), it is clear that the time evolution of $Q(t)$ is described by

$$Q' = A(t, QY_0Q^{\mathrm{T}})Q, \qquad Q(0) = I, \tag{145}$$

i.e., an equation of type (139). Yet there is another possibility: if we seek the orthogonal matrix solution of (145) as $Q(t) = \exp(\Omega(t))$ with Ω skew-symmetric,

$$Y(t) = \mathrm{e}^{\Omega(t)}Y_0\,\mathrm{e}^{-\Omega(t)}, \quad t \geq 0, \Omega(t) \in \mathfrak{so}(n), \tag{146}$$

then the corresponding equation for Ω reads

$$\Omega' = \mathrm{d}\exp_{\Omega}^{-1}\left(A(t, \mathrm{e}^{\Omega}Y_0\mathrm{e}^{-\Omega})\right), \quad \Omega(0) = O. \tag{147}$$

In a similar way as for Eq. (141), we apply Picard's iteration to (147) and truncate the $\mathrm{d}\exp^{-1}$ series at $k = m - 2$. Now we can also consistently truncate the operator

$$\mathrm{Ad}_{\Omega}Y_0 \equiv \mathrm{e}^{\Omega}Y_0\mathrm{e}^{-\Omega} = \mathrm{e}^{\mathrm{ad}_{\Omega}}Y_0$$

and the outcome still lies in $\mathfrak{so}(n)$. By doing so, we replace the computation of one matrix exponential by several commutators.

In the end, the scheme reads

$$\begin{aligned}
\Omega^{[1]}(t) &= \int_0^t A(s, Y_0)\mathrm{d}s \\
\Theta_{m-1}(t) &= \sum_{l=0}^{m-1} \frac{1}{l!}\mathrm{ad}^l_{\Omega^{[m-1]}(t)}Y_0 \\
\Omega^{[m]}(t) &= \sum_{k=0}^{m-2} \frac{B_k}{k!}\int_0^t \mathrm{ad}^k_{\Omega^{[m-1]}(s)}A(s, \Theta_{m-1}(s))\mathrm{d}s, \quad m \geq 2
\end{aligned} \tag{148}$$

and, as before, one has $\Omega(t) = \Omega^{[m]}(t) + \mathcal{O}(t^{m+1})$. Thus

$$\begin{aligned}
\Theta_1(t) &= Y_0 + [\Omega^{[1]}(t), Y_0] \\
\Omega^{[2]}(t) &= \int_0^t A(s, \Theta_1(s))\mathrm{d}s \\
\Theta_2(t) &= Y_0 + [\Omega^{[2]}(t), Y_0] + \frac{1}{2}[\Omega^{[2]}(t), [\Omega^{[2]}(t), Y_0]] \\
\Omega^{[3]}(t) &= \int_0^t A(s, \Theta_2(s))\mathrm{d}s - \frac{1}{2}\int_0^t [\Omega^{[2]}(s), A(s, \Theta_2(s))]\mathrm{d}s
\end{aligned}$$

and so on. Observe that this procedure preserves the isospectrality of the flow, since the approximation $\Omega^{[m]}(t)$ lies in $\mathfrak{so}(n)$ for all $m \geq 1$ and $t \geq 0$. It is also equally possible to develop a formalism based on rooted trees in this case, in a similar way as for the standard Magnus expansion.

Example. The double bracket equation

$$Y' = [[Y, N], Y], \qquad Y(0) = Y_0 \in \mathrm{Sym}(n) \tag{149}$$

was introduced by Brockett [101] and Chu & Driessel [102] to solve certain standard problems in applied mathematics, although similar equations also appear in the formulation of physical theories such as micromagnetics [103]. Here N is a constant $n \times n$ symmetric matrix. It clearly constitutes an example of an isospectral flow with $A(t, Y) \equiv [Y, N]$. When the procedure (148) is applied to (149), one reproduces exactly the expansion obtained in [104] with the convergence domain established in [105]. □

3.4. Treatment of general nonlinear equations

As a matter of fact, the Magnus expansion can be formally generalized to any nonlinear explicitly time-dependent differential equation. Given the importance of the expansion, it has, indeed, been (re)derived a number of times along the years in different settings. We have to mention, in this respect, the work of Agrachev and Gamkrelidze [106–108], and

Strichartz [53]. In the context of Hamiltonian dynamical systems, the expansion was first proposed in [109] and subsequently applied in a more general context in [110] with the aim of designing new numerical integration algorithms.

By introducing nonstationary vector fields and flows, it turns out that one gets a linear differential equation in terms of operators which can be analyzed in exactly the same way as in Section 2. Thus it is in principle possible to build approximate solutions of the differential equation which preserve some geometric properties of the exact solution. The corresponding Magnus series expansion allows us to write the formal solution, and then different approximations can be obtained by truncating the series. Obviously, this formal expansion presents two difficulties in order to render a useful algorithm in practice: (i) it is not evident what the domain of convergence is, and (ii) some device has to be designed to compute the exponential map once the series is truncated.

Next we briefly summarize the main ideas involved in the procedure. To begin with, let us consider the autonomous equation

$$\mathbf{x}' = \mathbf{f}(\mathbf{x}), \qquad \mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{R}^n. \tag{150}$$

If φ^t denotes the exact flow of (150), i.e. $\mathbf{x}(t) = \varphi^t(\mathbf{x}_0)$, then for each infinitely differentiable map $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, $g(\varphi^t(\mathbf{y}))$ admits the representation

$$g(\varphi^t(\mathbf{y})) = \Phi^t[g](\mathbf{y}) \tag{151}$$

where the operator Φ^t acts on differentiable functions [111]. To be more specific, let us introduce the Lie derivative (or Lie operator) associated with $\mathbf{f}$,

$$L_{\mathbf{f}} = \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i}. \tag{152}$$

It acts on differentiable functions $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ (see [112, Chap. 8] for more details) as

$$L_{\mathbf{f}}F(\mathbf{y}) = F'(\mathbf{y})\mathbf{f}(\mathbf{y}),$$

where $F'(\mathbf{y})$ denotes the Jacobian matrix of F. It follows from the chain rule that, for the solution $\varphi^t(\mathbf{x}_0)$ of (150),

$$\frac{\mathrm{d}}{\mathrm{d}t}F(\varphi^t(\mathbf{x}_0)) = (L_{\mathbf{f}}F)(\varphi^t(\mathbf{x}_0)), \tag{153}$$

and applying the operator iteratively one gets

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}F(\varphi^t(\mathbf{x}_0)) = (L_{\mathbf{f}}^k F)(\varphi^t(\mathbf{x}_0)), \quad k \geq 1.$$

Therefore, the Taylor series of $F(\varphi^t(\mathbf{x}_0))$ at $t = 0$ is given by

$$F(\varphi^t(\mathbf{x}_0)) = \sum_{k\geq 0} \frac{t^k}{k!}(L_{\mathbf{f}}^k F)(\mathbf{x}_0) = \exp(tL_{\mathbf{f}})[F](\mathbf{x}_0). \tag{154}$$

Now, putting $F(\mathbf{y}) = \mathrm{Id}(\mathbf{y}) = \mathbf{y}$, the identity map, this is just the Taylor series of the solution itself

$$\varphi^t(\mathbf{x}_0) = \sum_{k\geq 0} \frac{t^k}{k!}(L_{\mathbf{f}}^k \mathrm{Id})(\mathbf{x}_0) = \exp(tL_{\mathbf{f}})[\mathrm{Id}](\mathbf{x}_0).$$

If we substitute F by g in (154) and compare with (151), then it is clear that $\Phi^t[g](\mathbf{y}) = \exp(tL_{\mathbf{f}})[g](\mathbf{y})$. The object $\exp(tL_{\mathbf{f}})$ is called the Lie transform associated with $\mathbf{f}$.

At this point, let us suppose that $\mathbf{f}(\mathbf{x})$ can be split as $\mathbf{f}(\mathbf{x}) = \mathbf{f}_1(\mathbf{x}) + \mathbf{f}_2(\mathbf{x})$, in such a way that the systems

$$\mathbf{x}' = \mathbf{f}_1(\mathbf{x}), \qquad \mathbf{x}' = \mathbf{f}_2(\mathbf{x})$$

have flows φ_1^t and φ_2^t, respectively, so that

$$g(\varphi_i^t(\mathbf{y})) = \exp(tL_{\mathbf{f}_i})[g](\mathbf{y}) \quad i = 1, 2.$$

Then, for their composition one has

$$g(\varphi_2^t \circ \varphi_1^s(\mathbf{y})) = \exp(sL_{\mathbf{f}_1}) \exp(tL_{\mathbf{f}_2})[g](\mathbf{y}). \tag{155}$$

This is precisely formula (154) with $\mathbf{f} = \mathbf{f}_1$, t replaced with s and with $F(\mathbf{y}) = \exp(tL_{\mathbf{f}_2})[g](\mathbf{y})$. Notice that the indices 1 and 2, as well as s and t to the left and right of Eq. (155), are permuted. In other words, *the Lie transforms appear in the reverse order to their corresponding maps* [79].

The Lie derivative $L_{\mathbf{f}}$ satisfies some remarkable properties. Given two functions ψ_1, ψ_2, it can be easily verified that

$$L_{\mathbf{f}}(\alpha_1\psi_1 + \alpha_2\psi_2) = \alpha_1 L_{\mathbf{f}}\psi_1 + \alpha_2 L_{\mathbf{f}}\psi_2, \quad \alpha_1, \alpha_2 \in \mathbb{R}$$
$$L_{\mathbf{f}}(\psi_1\,\psi_2) = (L_{\mathbf{f}}\psi_1)\psi_2 + \psi_1\, L_{\mathbf{f}}\psi_2$$

and by induction we can prove the Leibniz rule

$$L_{\mathbf{f}}^k(\psi_1\,\psi_2) = \sum_{i=0}^{k} \binom{k}{i} \left(L_{\mathbf{f}}^i\psi_1\right)\left(L_{\mathbf{f}}^{k-i}\psi_2\right),$$

with $L_{\mathbf{f}}^i\psi = L_{\mathbf{f}}\left(L_{\mathbf{f}}^{i-1}\psi\right)$ and $L_{\mathbf{f}}^0\psi = \psi$, justifying the name of Lie derivative. In addition, given two vector fields $\mathbf{f}$ and $\mathbf{g}$, then

$$\alpha_1 L_{\mathbf{f}} + \alpha_2 L_{\mathbf{g}} = L_{\alpha_1\mathbf{f}+\alpha_2\mathbf{g}},$$
$$[L_{\mathbf{f}}, L_{\mathbf{g}}] = L_{\mathbf{f}}L_{\mathbf{g}} - L_{\mathbf{g}}L_{\mathbf{f}} = L_{\mathbf{h}},$$

where $\mathbf{h}$ is another vector field corresponding to the Lie bracket of the vector fields $\mathbf{f}$ and $\mathbf{g}$, denoted by $\mathbf{h} = (\mathbf{f}, \mathbf{g})$, whose components are

$$h_i = (\mathbf{f}, \mathbf{g})_i = L_{\mathbf{f}}g_i - L_{\mathbf{g}}f_i = \sum_{j=1}^{n}\left(f_j\frac{\partial g_i}{\partial x_j} - g_j\frac{\partial f_i}{\partial x_j}\right). \tag{156}$$

Moreover, from (151) and (153) (replacing F with g) we can write

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^t[g](\mathbf{x}_0) = \frac{\mathrm{d}}{\mathrm{d}t}g(\varphi^t(\mathbf{x}_0)) = (L_{\mathbf{f}}g)(\varphi^t(\mathbf{x}_0)) = \Phi^t L_{\mathbf{f}}[g](\mathbf{x}_0).$$

Particularizing to the function $g(\mathbf{x}) = \mathrm{Id}_j(\mathbf{x}) = x_j$, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^t[\mathrm{Id}_j](\mathbf{y}) = \Phi^t L_{\mathbf{f}(\mathbf{y})}[\mathrm{Id}_j](\mathbf{y}), \quad j = 1, \ldots, n, \mathbf{y} = \mathbf{x}_0$$

or, in short,

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^t = \Phi^t L_{\mathbf{f}(\mathbf{y})}, \quad \mathbf{y} = \mathbf{x}_0, \tag{157}$$

i.e., a linear differential equation for the operator Φ^t. Notice that, as expected, Eq. (157) admits as formal solution

$$\Phi^t = \exp(tL_{\mathbf{f}(\mathbf{y})}), \quad \mathbf{y} = \mathbf{x}_0. \tag{158}$$

We can follow the same steps for the non-autonomous equation

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}), \tag{159}$$

where, now, the operational equation to be solved is

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^t = \Phi^t L_{\mathbf{f}(t,\mathbf{y})}, \quad \mathbf{y} = \mathbf{x}_0. \tag{160}$$

To simplify notation, from now on we consider $\mathbf{x}_0$ as a set of coordinates such that $\mathbf{f}(t, \mathbf{x}_0)$ is a differentiable function of $\mathbf{x}_0$. Since $L_{\mathbf{f}}$ is a linear operator, we can then directly use the Magnus series expansion to obtain the formal solution of (160) as $\Phi^t = \exp(L_{\mathbf{w}(t,\mathbf{x}_0)})$, with $\mathbf{w} = \sum_i \mathbf{w}_i$. The first two terms are now

$$\mathbf{w}_1(t, \mathbf{x}_0) = \int_0^t \mathbf{f}(s, \mathbf{x}_0)\mathrm{d}s$$
$$\mathbf{w}_2(t, \mathbf{x}_0) = -\frac{1}{2}\int_0^t \mathrm{d}s_1 \int_0^{s_1} \mathrm{d}s_2(\mathbf{f}(s_1, \mathbf{x}_0), \mathbf{f}(s_2, \mathbf{x}_0)). \tag{161}$$

Observe that the sign of $\mathbf{w}_2$ is changed when compared with Ω_2 in (43), and the integrals only affect the explicit time-dependent part of the vector field. In general, due to the structure of Eqs. (30) and (160), the expression for $\mathbf{w}_n(t, \mathbf{x}_0)$ can be obtained from the corresponding $\Omega_n(t)$ in the linear case by applying the following rules:

(1) replace $A(t_i)$ by $\mathbf{f}(t_i, \mathbf{x}_0)$;
(2) replace the commutator $[\cdot, \cdot]$ by the Lie bracket (156);
(3) change the sign in $\mathbf{w}_n(t, \mathbf{x}_0)$ for even n.

Once $\mathbf{w}^{[n]} = \sum_{i=1}^{n} \mathbf{w}_i(t, \mathbf{x}_0)$ is computed, it still remains to evaluate the action of the Lie transform $\exp(L_{\mathbf{w}(t,\mathbf{x}_0)})$ on the initial conditions $\mathbf{x}_0$. At time $t = T$, this can be seen as the 1-flow solution of the autonomous differential equation

$$\mathbf{y}' = \mathbf{w}^{[n]}(T, \mathbf{y}), \quad \mathbf{y}(0) = \mathbf{x}_0, \tag{162}$$

since $\mathbf{y}(1) = \exp(L_{\mathbf{w}(T,\mathbf{x}_0)})\mathbf{x}_0 = \mathbf{x}(T)$.

Although this is arguably the most direct way to construct a Magnus expansion for *arbitrary* time dependent nonlinear differential equations, it is by no means the only one. In particular, Agrachev and Gamkrelidze [107,108] obtain a similar expansion by transforming (160) into the integral equation

$$\Phi^t = \mathrm{Id} + \int_0^t \Phi^s \vec{X}_s \mathrm{d}s \tag{163}$$

which is subsequently solved by successive approximations. Here, for clarity, we have denoted $\vec{X}_s \equiv L_{\mathbf{f}(s,\mathbf{x}_0)}$. Then one gets the formal series

$$\begin{aligned}\Phi^t &= \mathrm{Id} + \int_0^t \mathrm{d}t_1 \vec{X}_{t_1} + \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \vec{X}_{t_2}\vec{X}_{t_1} + \cdots \\ &= \mathrm{Id} + \sum_{m=1}^{\infty} \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \cdots \int_0^{t_{m-1}} \mathrm{d}t_m \, \vec{X}_{t_m} \cdots \vec{X}_{t_1}.\end{aligned} \tag{164}$$

An object with this shape is called a *formal chronological series* [107], and the set of all formal chronological series can be endowed with a real associative algebra structure. It is then possible to show that there exists an absolutely continuous formal chronological series

$$V_t(\vec{X}_t) = \sum_{m=1}^{\infty} \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \cdots \int_0^{t_{m-1}} \mathrm{d}t_m \, G_m(\vec{X}_{t_1}, \ldots, \vec{X}_{t_m}) \tag{165}$$

such that

$$\Phi^t = \exp(V_t(\vec{X}_t)).$$

Here $G_m(\vec{X}_{t_1}, \ldots, \vec{X}_{t_m})$ are Lie polynomials, homogeneous of the first grade in each variable, which can be algorithmically constructed. In particular,

$$\begin{aligned}&G_1(\vec{X}_{t_1}) = \vec{X}_{t_1} \\ &G_2(\vec{X}_{t_1}, \vec{X}_{t_2}) = \frac{1}{2}[\vec{X}_{t_2}, \vec{X}_{t_1}] \\ &G_3(\vec{X}_{t_1}, \vec{X}_{t_2}, \vec{X}_{t_3}) = \frac{1}{6}([\vec{X}_{t_3}, [\vec{X}_{t_2}, \vec{X}_{t_1}]] + [[\vec{X}_{t_3}, \vec{X}_{t_2}], \vec{X}_{t_1}]).\end{aligned}$$

The series (165) in general diverges, even if the Lie operator $\vec{X}_t$ is analytic [106]. Nevertheless, in certain cases convergence holds. For instance, if $\vec{X}_t$ belongs to a Banach Lie algebra $\mathcal{B}$ for all $t \in \mathbb{R}$, where one has a norm satisfying $\|[X, Y]\| \leq \|X\|\|Y\|$ for all $X, Y \in \mathcal{B}$ and $\int_0^t \|\vec{X}_s\| \mathrm{d}s \equiv \int_0^t \|L_{\mathbf{f}(s,\mathbf{x}_0)}\| \mathrm{d}s \leq 0.44$, then $V_t(\vec{X}_t)$ converges absolutely in $\mathcal{B}$ [107]. As a matter of fact, an argument analogous to that used in [30,70] may allow us to improve this bound and get convergence for

$$\int_0^t \|\vec{X}_s\| \mathrm{d}s \leq \frac{1}{2} \int_0^{2\pi} \frac{1}{2 + \frac{x}{2}(1 - \cot\frac{x}{2})} \mathrm{d}x = 1.08686870\ldots.$$

3.4.1. Treatment of Hamiltonian systems

We have seen how the algebraic setting we have developed for linear systems of differential equations may be extended formally to nonlinear systems. We will next review how it can be adapted to the important class of Hamiltonian systems. In this context, the role of a Lie bracket of vector fields (156) is played by the classical Poisson bracket [113].

The Lie algebraic presentation of Hamiltonian systems in Classical Mechanics has been approached in different ways, and the Magnus expansion invoked in this context by diverse authors [114–116]. More explicit use of the Magnus expansion is done in [109] where the evolution operator for a classical system is constructed and its differential equation analyzed.

To particularize to this situation the preceding general treatment, let us consider a system with l degrees of freedom and phase space variables $\mathbf{x} = (\mathbf{q}, \mathbf{p}) = (q_1, \ldots, q_l, p_1, \ldots, p_l)$, where (q_i, p_i), $i = 1, \ldots, l$ are the usual pairs of canonical conjugate coordinate and momentum, respectively. By defining the Poisson bracket of two scalar functions $F(\mathbf{q}, \mathbf{p})$ and $G(\mathbf{q}, \mathbf{p})$ of phase space variables in the conventional way [113]

$$\{F, G\} \equiv \sum_{i=1}^{l} \left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i} \right),$$

we have

$$\{F, G\} = \sum_{i,j=1}^{2l} \frac{\partial F}{\partial x_i} J_{ij} \frac{\partial G}{\partial x_j},$$

and in particular

$$\{x_i, x_j\} = J_{ij}.$$

Here J is the basic symplectic matrix appearing in Eq. (20) (with $n = l$). With these definitions, the set of (sufficiently smooth) functions on phase space acquires the structure of a Lie algebra and we can associate with any such function $F(\mathbf{x})$ a Lie operator

$$L_F = \sum_{i,j=1}^{2l} \frac{\partial F}{\partial x_i} J_{ij} \frac{\partial}{\partial x_j} \tag{166}$$

which acts on the same set of functions as $L_F G = \{F, G\}$. It is then a simple exercise to show that the set of all Lie operators is also a Lie algebra under the usual commutator $[L_F, L_G] = L_F L_G - L_G L_F$ and furthermore

$$[L_F, L_G] = L_{\{F,G\}}.$$

Given the Hamiltonian function $H(\mathbf{q}, \mathbf{p}, t) : \mathbb{R}^{2l} \times \mathbb{R} \to \mathbb{R}$, where $\boldsymbol{q}, \boldsymbol{p} \in \mathbb{R}^l$, the equations of motion are

$$\mathbf{q}' = \nabla_{\mathbf{p}} H, \qquad \mathbf{p}' = -\nabla_{\mathbf{q}} H, \tag{167}$$

or, equivalently, in terms of $\mathbf{x}$,

$$\mathbf{x}' = J \nabla_{\mathbf{x}} H.$$

It is then elementary to show that the Lie operator L_{-H} is nothing but the Lie derivative $L_{\mathbf{f}}$ (152) associated with the function

$$\mathbf{f} = J \nabla_{\mathbf{x}} H.$$

Therefore, the operational equation (160) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} \Phi_H^t = \Phi_H^t L_{-H(\mathbf{y},t)}, \quad \mathbf{y} = \mathbf{x}_0 \tag{168}$$

and the previous treatment also holds in this setting. As a result, the Magnus expansion reads

$$\Phi_H^t = \exp(L_W), \tag{169}$$

where $W = \sum_{i=1}^{\infty} W_i$ and the first two terms are

$$W_1(\mathbf{x}_0) = -\int_0^t H(\mathbf{x}_0, s)\mathrm{d}s \tag{170}$$

$$W_2(\mathbf{x}_0) = -\frac{1}{2} \int_0^t \mathrm{d}s_1 \int_0^{s_1} \mathrm{d}s_2 \{H(\mathbf{x}_0, s_1), H(\mathbf{x}_0, s_2)\}.$$

3.5. Magnus expansion and the Chen–Fliess series

Suppose that $\mathbf{f}$ in Eq. (159) has the form $\mathbf{f}(t, \mathbf{x}) = \sum_{i=1}^{m} u_i(t) \mathbf{f}_i(\mathbf{x})$, i.e., we are dealing with the nonlinear differential equation

$$\mathbf{x}'(t) = \sum_{i=1}^{m} u_i(t)\, \mathbf{f}_i(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{p}, \tag{171}$$

where $u_i(t)$ are integrable functions of time. Systems of the form (171) appear, for instance, in nonlinear control theory. In that context, the functions u_i are the controls and $\mathbf{f}_i$ are related to the non-varying geometry of the system. Observe that this problem constitutes the natural (nonlinear) generalization of the case studied in Section 2.11.

One of the most basic procedures for obtaining $\mathbf{x}(T)$ for a given T is by applying simple Picard iteration. For an analytic *output function* $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, from (153) it is clear that

$$\frac{\mathrm{d}}{\mathrm{d}t} g(\mathbf{x}(t)) = (L_{(\sum u_i \mathbf{f}_i)} g)(\mathbf{x}(t)) = \sum_{i=1}^{m} u_i(t) (E_i g)(\mathbf{x}(t)), \qquad g(\mathbf{x}(0)) = g(\mathbf{p}), \tag{172}$$

where, for simplicity, we have denoted by E_i the Lie derivative $L_{\mathbf{f}_i}$. This can be particularized to the case $g = x_i$, the ith component function.

By rewriting (172) as an equivalent integral equation and iterating we get

$$\begin{aligned} g(\mathbf{x}(t)) &= g(\mathbf{p}) + \int_0^t \sum_{i_1=1}^m u_{i_1}(t_1)(E_{i_1}g)(\mathbf{x}(t_1))\mathrm{d}t_1 \\ &= g(\mathbf{p}) + \int_0^t \sum_{i_1=1}^m u_{i_1}(t_1)\left((E_{i_1}g)(\mathbf{p}) + \int_0^{t_1}\sum_{i_2=1}^m u_{i_2}(t_2)(E_{i_2}E_{i_1}g)(\mathbf{x}(t_2))\mathrm{d}t_2\right)\mathrm{d}t_1 \\ &= g(\mathbf{p}) + \int_0^t \sum_{i_1=1}^m u_{i_1}(t_1)\left((E_{i_1}g)(\mathbf{p}) + \int_0^{t_1}\sum_{i_2=1}^m u_{i_2}(t_2)\left((E_{i_2}E_{i_1}g)(\mathbf{p})\right.\right. \\ &\quad \left.\left. + \int_0^{t_2}\sum_{i_3=1}^m u_{i_3}(t_3)(E_{i_3}E_{i_2}E_{i_1}g)(\mathbf{x}(t_3))\mathrm{d}t_3\right)\mathrm{d}t_2\right)\mathrm{d}t_1 \end{aligned} \tag{173}$$

and so on. Notice that, in this expression, the time dependence of the solution is separated from the non-varying geometry of the system, which is contained in the vector fields E_i and need to be computed only once, at the beginning of the calculation. Next, we reverse the names of the integration variables and indices used (e.g., rename i_1 to become i_3 and vice versa), so that

$$\begin{aligned} g(\mathbf{x}(t)) &= g(\mathbf{p}) + \sum_{i_1=1}^m \left(\int_0^t u_{i_1}(t_1)\mathrm{d}t_1\right)(E_{i_1}g)(\mathbf{p}) + \sum_{i_2=1}^m\sum_{i_1=1}^m \left(\int_0^t\int_0^{t_2} u_{i_2}(t_2)u_{i_1}(t_1)\mathrm{d}t_1\mathrm{d}t_2\right)(E_{i_1}E_{i_2}g)(\mathbf{p}) \\ &\quad + \sum_{i_3=1}^m\sum_{i_2=1}^m\sum_{i_1=1}^m \left(\int_0^t\int_0^{t_3}\int_0^{t_2} u_{i_3}(t_3)u_{i_2}(t_2)u_{i_1}(t_1)\mathrm{d}t_1\mathrm{d}t_2\mathrm{d}t_3\right)(E_{i_1}E_{i_2}E_{i_3}g)(\mathbf{p}) + \cdots. \end{aligned} \tag{174}$$

Observe that the indices in the Lie derivatives and in the integrals are in the opposite order. This procedure can be further iterated, thus yielding the formal infinite series

$$g(\mathbf{x}(t)) = g(\mathbf{x}(0)) + \sum_{s\geq 1}\sum_{i_1\cdots i_s}\int_0^t\int_0^{t_s}\cdots\int_0^{t_3}\int_0^{t_2} u_{i_s}(t_s)\cdots u_{i_1}(t_1)\mathrm{d}t_1\cdots\mathrm{d}t_s E_{i_1}\cdots E_{i_s}g(\mathbf{x}(0)), \tag{175}$$

where each $i_j \in L = \{1,\ldots,m\}$. An expression of the form (175) is referred to as the Chen–Fliess series, and it can be proved that, under certain circumstances, it actually converges uniformly to the solution of (172) [90]. This series originates in K.T. Chen's work [117] on geometric invariants and iterated integrals of paths in $\mathbb{R}^n$. Later, Fliess [118] applied the theory to the analysis of control systems.

One of the great advantages of the Chen–Fliess series is that it can be manipulated with purely algebraic and combinatorial tools, instead of working directly with nested integrals. To emphasize this aspect, observe that each term in the series can be identified by a sequence of indices or *word* $w = i_1 i_2\cdots i_s$ in the *alphabet* L through the following two maps:

$$\begin{aligned} &\mathcal{M}_1 : w = i_1 i_2\cdots i_s \longmapsto \left(g \mapsto (E_w g)(\mathbf{p}) = (E_{i_1}E_{i_2}\cdots E_{i_s}g)(\mathbf{p})\right), \\ &\mathcal{M}_2 : w = i_1 i_2\cdots i_s \longmapsto \left(u \mapsto \int_0^t u_{i_s}(t_s)\int_0^{t_s}\cdots\int_0^{t_2} u_{i_1}(t_1)\mathrm{d}t_1\cdots\mathrm{d}t_{s-1}\mathrm{d}t_s\right). \end{aligned}$$

In fact, the nested integral appearing in the map $\mathcal{M}_2$ can be expressed in a simple way, as we did for the linear case in (114)

$$\alpha_{i_1\cdots i_s} = \int_0^t u_{i_s}(t_s)\int_0^{t_s}\cdots\int_0^{t_2} u_{i_1}(t_1)\mathrm{d}t_1\cdots\mathrm{d}t_{s-1}\mathrm{d}t_s. \tag{176}$$

With this notation, the series of linear differential operators appearing at the right-hand side of (175) can be written in the compact form [91]

$$\sum_{w\in L^*}\alpha_w E_w, \tag{177}$$

where L^* denotes the set of words on the alphabet $L = \{1, 2, \ldots, m\}$, the function α_w is given by (176) for each word $w \in L^*$ and

$$E_w = E_{i_1}\cdots E_{i_s}, \quad \text{if } w = i_1\cdots i_s \in L^*.$$

It was proved by Chen that the series (177) is an exponential Lie series [117], i.e., it can be rewritten as the exponential of a series of vector fields obtained as nested commutators of $E_1,\ldots,E_m$. Such an expression is referred to in nonlinear control

as the formal analogue of a continuous Baker–Campbell–Hausdorff formula and also as the logarithm of the Chen–Fliess series [119].

Notice the similarities of this procedure with the more general treatment carried out in Section 3.4 for the nonlinear differential equation (150). Thus, expression (164) constitutes the generalization of (175) to an arbitrary function $\mathbf{f}$ in (150). Conversely, the logarithm of the Chen–Fliess series can be viewed as the corresponding nonlinear Magnus series for the particular nonlinear system (171).

From these considerations, it is clear that, in principle, one can obtain an explicit formula for the terms of the logarithm of the Chen–Fliess series in a basis of the Lie algebra generated by $E_1, \ldots, E_m$, but this problem has been only recently solved for any number of operators m and arbitrary order, using labeled rooted trees [91]. Thus, for instance, when $m = 2$, it holds that

$$\sum_{w\in L^*} \alpha_w E_w = \exp(\beta_1 E_1 + \beta_2 E_2 + \beta_{12}[E_1, E_2] + \beta_{112}[E_1, [E_1, E_2]] + \beta_{212}[E_2, [E_1, E_2]] + \cdots), \tag{178}$$

where, not surprisingly, the expressions of the β coefficients are given by (113) with the corresponding change of sign in β_{12} due to the nonlinear character of Eq. (171).

Another relevant consequence of the connection between Magnus series and the Chen–Fliess series is the following: the Lie series defining the logarithm of the Chen–Fliess series can be obtained explicitly from the recurrence (49)–(50), valid in principle for the linear case. Of course, the successive terms of the Chen–Fliess series itself can be generated by expanding the exponential.

4. Illustrative examples

After having reviewed, in the preceding two sections, the main theoretical aspects of the Magnus expansion and other exponential methods, in this section we gather some examples of their application. All of them are standard problems of Quantum Mechanics where the exact solution for the evolution operator $U(t)$ is well known. Due to their simplicity, higher order computations are possible with a reasonable amount of effort. The comparison between approximate and exact analytic results may help the reader to grasp the advantages as well as the technical difficulties of the methods we have analyzed.

The examples considered here are treated in [15,34,86,93], although some results are unpublished material, in particular those involving highest order computations. In Section 4.1 we present results concerning the most straightforward way of dealing with ME, namely computations in the Interaction Picture. In Section 4.2 an application of ME in the adiabatic basis is developed. Section 4.3 is devoted to illustrating the exponential infinite-product expansions of Fer and Wilcox. An example on the application of the iterative version of ME by Voslamber is given in Section 4.4. Eventually, Section 4.5 contains an application of the Floquet–Magnus formalism.

4.1. ME in the Interaction Picture

We illustrate the application of ME in the Interaction Picture (see Section 2.9) by means of two simple time-dependent physical systems frequently encountered in the literature, for which exact solutions are available: the time-dependent forced harmonic oscillator, and a particle of spin $\frac{1}{2}$ in a constant magnetic field. In the first case we fix $\hbar = 1$ for convenience.

As we will see, ME in the Interaction Picture is appropriate whenever the characteristic time scale of the perturbation is shorter than the proper time scale of the system.

To illustrate and evaluate the quality of the various approximations for the time-evolution operator, we compute the transition probabilities among non-perturbed eigenstates, induced by the small perturbation.

4.1.1. Linearly forced harmonic oscillator

The Hamiltonian function describing a linearly driven harmonic oscillator reads ($\hbar = 1$)

$$H = H_0 + V(t), \quad \text{with } H_0 = \frac{1}{2}\omega_0(p^2 + q^2),\ V(t) = \sqrt{2} f(t) q \tag{179}$$

and $f(t)$ is real. Here q and p stand for the position and momentum operators satisfying $[q, p] = i$ and ω_0 gives the energy level spacing in absence of the perturbation $V(t)$. We introduce the usual operators $a_\pm \equiv \frac{1}{\sqrt{2}}(q \mp \mathrm{i}p)$, so that $[a_-, a_+] = 1$. With this notation we have

$$H_0 = \omega_0 \left(a_+ a_- + \frac{1}{2} \right), \qquad V = f(t)(a_+ + a_-). \tag{180}$$

The eigenstates of H_0 are denoted by $|n\rangle$, so that $H_0|n\rangle = \omega_0(n + \frac{1}{2})|n\rangle$, where n stands for the quantum number. With this notation $n = 0$ corresponds to the ground state.

For simplicity in the computations we choose $\omega_0 = 1$. In accordance with the prescriptions in Section 2.9, the Hamiltonian in the Interaction Picture is given by (92) and reads

$$H_I(t) = \mathrm{e}^{\mathrm{i}H_0 t}\, V(t)\mathrm{e}^{-\mathrm{i}H_0 t} = f(t)(\mathrm{e}^{\mathrm{i}t}a_+ + \mathrm{e}^{-\mathrm{i}t}a_-). \tag{181}$$

Accordingly, the evolution operator is factorized as

$$U(t,0) = \exp(-\mathrm{i}H_0 t)U_I(t,0), \tag{182}$$

where the new evolution operator U_I is obtained from $U_I' = \tilde{H}_I(t)U_I \equiv -\mathrm{i}H_I(t)U_I$.

The infinite Magnus series terminates in the present example. It happens because the second order Magnus approximant, which involves the computation of

$$\begin{aligned}[\tilde{H}_I(t_1), \tilde{H}_I(t_2)] &= f(t_1)f(t_2)\left(\mathrm{e}^{\mathrm{i}(t_1-t_2)}[a_+, a_-] + \mathrm{e}^{-\mathrm{i}(t_1-t_2)}[a_-, a_+]\right)\\ &= 2\mathrm{i}f(t_1)f(t_2)\sin(t_2 - t_1)\end{aligned} \tag{183}$$

reduces to a scalar function. Thus the Magnus series in the Interaction Picture furnishes the exact evolution operator irrespective of $f(t)$:

$$\begin{aligned}U_I(t,0) &= \exp\left(\int_0^t \mathrm{d}t_1 \tilde{H}_I(t_1) - \frac{1}{2}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 [\tilde{H}_I(t_1), \tilde{H}_I(t_2)]\right)\\ &= \exp\left(-\mathrm{i}(\alpha a_+ + \alpha^* a_-) - \mathrm{i}\beta\right)\\ &= \exp(-\mathrm{i}\alpha a_+)\exp(-\mathrm{i}\alpha^* a_-)\exp(-\mathrm{i}\beta - |\alpha|^2/2),\end{aligned} \tag{184}$$

where we have defined

$$\alpha \equiv \int_0^t \mathrm{d}t_1 f(t_1)\mathrm{e}^{\mathrm{i}t_1}, \tag{185}$$

$$\beta \equiv \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 f(t_1)f(t_2)\sin(t_2 - t_1). \tag{186}$$

Eqs. (184) and (182) yield the exact time-evolution operator for the linearly forced harmonic oscillator Hamiltonian (179) [120].

To compute transition probabilities between free harmonic oscillator states of quantum numbers n and m,

$$P_{n\to m} = |\langle m|U_I|n\rangle|^2, \tag{187}$$

the last form in (184) is most convenient. Specifically, assuming that the oscillator was initially in its ground state $|0\rangle$, we get in particular the familiar Poisson distribution for the transition probabilities

$$P_{0\to n} = \frac{1}{n!}|\alpha|^{2n}\exp(-|\alpha|^2). \tag{188}$$

4.1.2. Two-level quantum systems

The generic Hamiltonian for a two-level quantum system can be written down in the form

$$H(t) = \begin{pmatrix} E_1(t) & C(t) \\ C^*(t) & E_2(t)\end{pmatrix} \tag{189}$$

where $E_1(t), E_2(t)$ are real functions and $C(t)$ is, in general, a complex function of t. We define the solvable piece of the Hamiltonian as the diagonal matrix

$$H_0(t) = \begin{pmatrix} E_1(t) & 0 \\ 0 & E_2(t)\end{pmatrix} \tag{190}$$

and all the time-dependent interaction described by the function $C(t)$ is considered as a perturbation. In the Interaction Picture, the new Hamiltonian reads (see (92))

$$H_I(t) = \begin{pmatrix} 0 & C(t)\exp\left(\mathrm{i}\int_0^t \mathrm{d}t'\omega(t')\right) \\ C^*(t)\exp\left(-\mathrm{i}\int_0^t \mathrm{d}t'\omega(t')\right) & 0\end{pmatrix} \tag{191}$$

with $\omega = (E_1 - E_2)/\hbar$. Suppose now that H_0 is time-independent. Then $U(t) = \exp(\tilde{H}_0 t)U_I(t)$. Without loss of generality, the H_0 may be rendered traceless, so that $E_1 = -E_2 \equiv E$. Thus $\pm E$ denote the eigenenergies associated to the eigenvectors

$|+\rangle \equiv (1,0)^{\mathrm{T}}$, $|-\rangle \equiv (0,1)^{\mathrm{T}}$ of H_0, the unperturbed system. In terms of Pauli matrices the Hamiltonian in this case may be expressed as

$$H(t) = \frac{1}{2}\hbar\omega\sigma_3 + f(t)\sigma_1 + g(t)\sigma_2, \tag{192}$$

where $f = \mathrm{Re}(C)$ and $g = -\mathrm{Im}(C)$.

Since H_0 is diagonal, the transition probability between eigenstates $|+\rangle$, $|-\rangle$ of H_0 is simply

$$P(t) = |\langle +|U_I(t)|-\rangle|^2. \tag{193}$$

As the evaluation of (193) requires the computation and manipulation of exponential matrices involving Pauli matrices, formulas (18) and (19) in Section 1.2 come in handy here.

Next, we study two particular cases of interaction for which the exact solution of the time evolution operator admits an analytic expression.

1- Rectangular step. Suppose that in (192) $g = 0$, namely,

$$H(t) = \frac{1}{2}\hbar\omega\sigma_3 + f(t)\sigma_1 \tag{194}$$

with $f = 0$ for $t < 0$ and $f = V_0$ for $t \geq 0$. Alternatively, if we restrict ourselves to compute an observable such as the transition probability, this example is equivalent to a rectangular mound (or rectangular barrier) of width $T = t$ above. The exact solution for this problem reads

$$U(t,0) = \exp\left(-\mathrm{i}\left(\frac{\omega}{2}\sigma_3 + \frac{V_0}{\hbar}\sigma_1\right)t\right), \tag{195}$$

which yields the exact transition probability

$$P_{ex} = \frac{4\gamma^2}{4\gamma^2+\xi^2}\sin^2\sqrt{\gamma^2+\xi^2/4} \tag{196}$$

between eigenstates $|+\rangle$, $|-\rangle$ of H_0. Here we have denoted $\gamma \equiv V_0 t/\hbar$ and $\xi \equiv \omega t$.

The Interaction Picture is defined here by the explicit integration of the diagonal piece in the Hamiltonian, so that $U = \exp(-\mathrm{i}\xi\sigma_3/2)U_I$, where U_I stands for the time evolution operator in the Interaction Picture and obeys

$$U_I' = \tilde{H}_I(t)U_I, \qquad U_I(0) = I \tag{197}$$

with

$$H_I(t) = f(t)(\sigma_1\cos\xi - \sigma_2\sin\xi). \tag{198}$$

A computation with the usual time-dependent perturbation theory gives for the first orders (formula (52))

$$P_{pt}^{(1)} = P_{pt}^{(2)} = \frac{4\gamma^2}{\xi^2}\sin^2(\xi/2) \tag{199}$$

$$P_{pt}^{(3)} = P_{pt}^{(4)} = \frac{\gamma^2}{\xi^2}\left[2\sin\frac{\xi}{2} - \frac{\gamma^2}{3\xi^2}\left(9\sin\frac{\xi}{2} + \sin\frac{3\xi}{2} - 6\xi\cos\frac{\xi}{2} + 4\sin^3\frac{\xi}{2}\right)\right]^2.$$

Notice that $P_{pt}^{(i)} > 1$ may happen in the equations above because the unitary character of the operator $U(t)$ is not preserved by the usual time-dependent perturbation formalism.

In this example it is not difficult to compute the first four terms in the Magnus series corresponding to $U_I(t) = \exp\Omega(t)$ in (197). To facilitate the notation, we define $s = \sin\xi$ and $c = \cos\xi$. The Magnus approximants in the Interaction Picture may be written down in terms of Pauli matrices and read explicitly

$$\begin{aligned}
\Omega_1 &= -\mathrm{i}\frac{\gamma}{\xi}[\sigma_1 s + \sigma_2(1-c)] \\
\Omega_2 &= -\mathrm{i}\left(\frac{\gamma}{\xi}\right)^2\sigma_3(s-\xi) \\
\Omega_3 &= -\mathrm{i}\left(\frac{\gamma}{\xi}\right)^3\frac{1}{3}\{\sigma_1[3\xi(1+c) - (5+c)s] + \sigma_2[(3\xi - s)s - 4(1-c)]\} \\
\Omega_4 &= -\mathrm{i}\left(\frac{\gamma}{\xi}\right)^4\frac{1}{3}\sigma_3[(4c+5)\xi - (c+8)s].
\end{aligned} \tag{200}$$

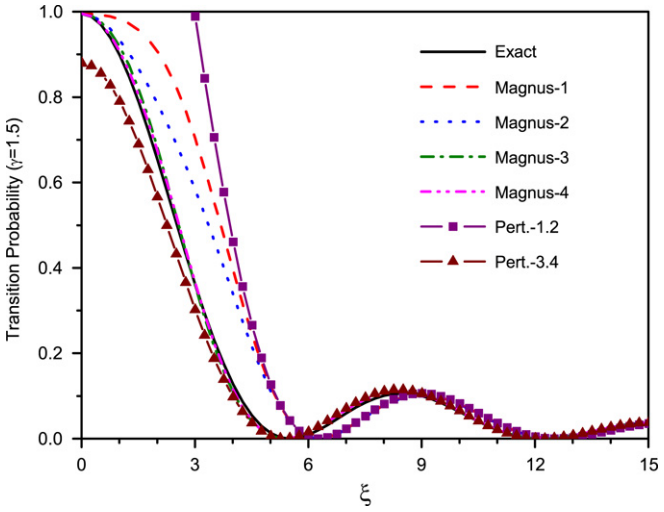


Fig. 3. Rectangular step: Transition probabilities as a function of ξ, with $\gamma = 1.5$. The solid line corresponds to the exact result (196). Broken lines stand for approximations obtained via ME, and lines with symbols correspond to perturbation theory, according to the legend. Computations up to fourth order, in the Interaction Picture.

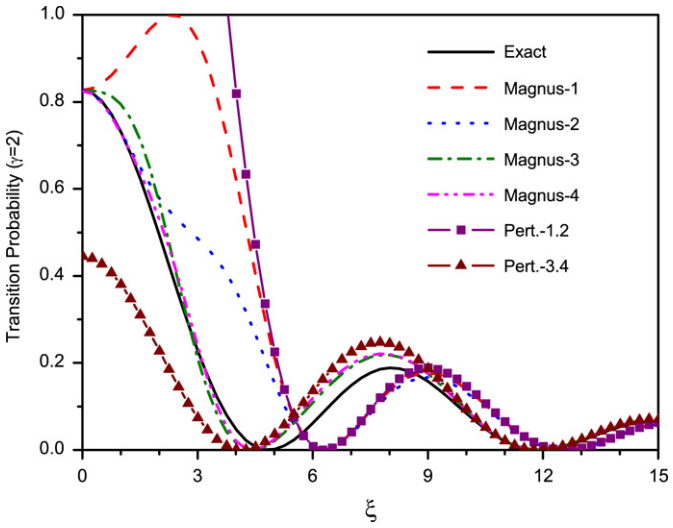


Fig. 4. Rectangular step: Transition probabilities as a function of ξ, with $\gamma = 2$. Lines are coded as in Fig. 3. Computations up to fourth order, in the Interaction Picture.

The first two formulae for the approximate transition probabilities are, respectively

$$P_M^{(1)} = \sin^2\left(\frac{2\gamma}{\xi}\sin(\xi/2)\right)$$

$$P_M^{(2)} = \frac{4\gamma^2}{\xi^2}\frac{\sin^2\lambda}{\lambda^2}\sin^2(\xi/2), \quad \lambda = [4\sin^2(\xi/2) + \frac{\gamma^2}{\xi^2}(\sin\xi - \xi)^2]^{1/2}. \tag{201}$$

We omit explicit expressions for $P_M^{(3)}$ and $P_M^{(4)}$ since they are quite involved. However, we include their outputs in Figs. 3–5, where we plot the first to fourth order approximate transition probabilities with ME in the Interaction Picture and compare them to the exact case, and also with perturbation theory outputs. In Figs. 3 and 4 we set $\gamma = 1.5$ and $\gamma = 2$ respectively, whereas in Fig. 5 we fix $\xi = 1$.

We observe that for the Magnus expansion in the Interaction Picture, the smaller the value of the parameter ξ, the better the approximate solution works. As a matter of fact, in the sudden limit, $\xi \ll 1$, ME furnishes the exact result; unlike perturbation theory. As far as the intensity of the perturbation γ increases, the quality of the approximations spoils. This effect is much more dramatic for the standard perturbation theory.

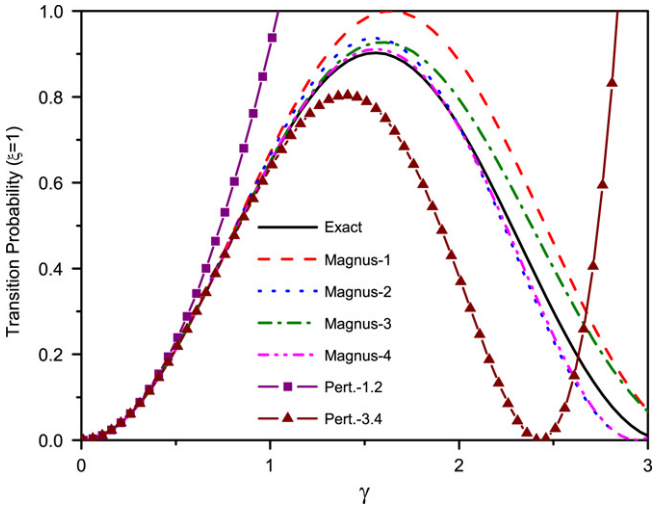


Fig. 5. Rectangular step: Transition probabilities as a function of γ, with $\xi = 1$. Lines are coded as in Fig. 3. Computations up to fourth order, in the Interaction Picture.

On the other hand, it is clear from (198) that

$$\int_{-\infty}^{t} \|H_I(t_1)\|_2 \, \mathrm{d}t_1 = \int_0^t |f(t_1)| \, \mathrm{d}t_1 = V_0 \, t,$$

whence $\int_{-\infty}^{t} \|\tilde{H}_I(t_1)\|_2 \, \mathrm{d}t_1 = \gamma$, and thus Theorem 9 guarantees that the Magnus expansion in the Interaction Picture is convergent if $\gamma < \pi$. Notice that this is always the case for the parameters considered in Figs. 3–5. The estimate $\gamma < \pi$ for the convergence domain in the Interaction Picture should be compared with the corresponding one in the Schrödinger picture:

$$\int_{-\infty}^{t} \|\tilde{H}(t_1)\|_2 \, \mathrm{d}t_1 = \sqrt{\gamma^2 + \frac{\xi^2}{4}} < \pi.$$

Notice then that, as pointed out in Section 2.9, a change of picture allows us to improve the convergence of the Magnus expansion.

2- Hyperbolic secant step: Rosen–Zener model. In the Rosen–Zener Hamiltonian [121] the interaction $C(t)$ in (189) is given by the real function $V(t) = V_0 \operatorname{sech}(t/T)$, where T determines the time-scale. We will use the notation $\gamma = \pi V_0 T/\hbar$ and $\xi = \omega T = 2ET/\hbar$.

The corresponding Hamiltonian in terms of Pauli matrices is

$$H(t) = E\sigma_3 + V(t)\sigma_1 \equiv \mathbf{a}(t) \cdot \boldsymbol{\sigma}, \quad V(t) = V_0/\cosh(t/T), \tag{202}$$

with $\mathbf{a} \equiv (V(t), 0, E)$. In the Interaction Picture one has

$$H_I(s) = V(s)(\sigma_1 \cos(\xi s) - \sigma_2 \sin(\xi s)) \tag{203}$$

in terms of the dimensionless time-variable $s = t/T$. Notice that ξ measures the ratio between the *interaction* time T and the *internal* time of the system $\hbar/2E$. From (203), and after straightforward calculation, the first and second ME operators are readily obtained.

The exact result for the transition probability (provided the time interval extends from $-\infty$ to $+\infty$), as well as perturbation theory and Magnus expansion up to second order read [15]

$$\begin{aligned}
P_{ex} &= \frac{\sin^2 \gamma}{\cosh^2(\pi\xi/2)} \\
P_{pt}^{(1)} &= P_{pt}^{(2)} = \frac{\gamma^2}{\cosh^2(\pi\xi/2)} \\
P_M^{(1)} &= \sin^2[\gamma/\cosh(\pi\xi/2)] \\
P_M^{(2)} &= \frac{\sin^2\lambda}{\lambda^2} \frac{\gamma^2}{\cosh^2(\pi\xi/2)} \\
\lambda &= \gamma \left[\frac{1}{\cosh^2(\pi\xi/2)} + \frac{\gamma^2 g^2(\xi)}{\pi^4} \right]^{1/2}, \quad g(\xi) = 8\xi \sum_{k=0}^{\infty} \frac{2k+1}{[(2k+1)^2 + \xi^2]^2}.
\end{aligned} \tag{204}$$

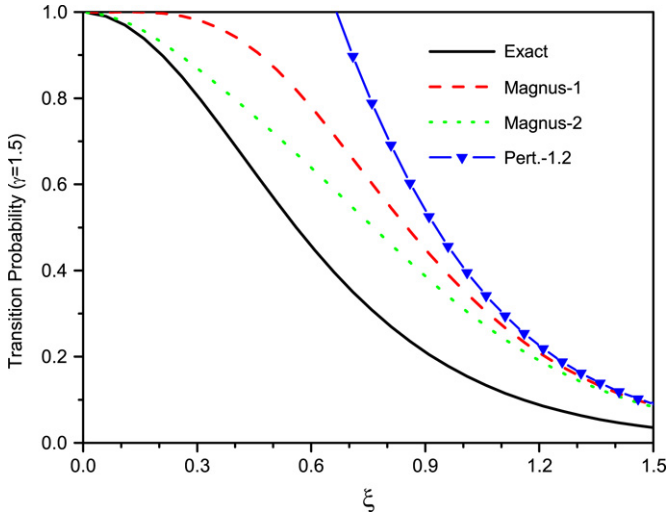


Fig. 6. Rosen–Zener model: Transition probabilities (204) as a function of ξ, with $\gamma = 1.5$. The solid line stands for the exact result. Broken lines stand for approximations obtained via ME, and triangles correspond to perturbation theory, according to the legend. Computations up to second order, in the Interaction Picture.

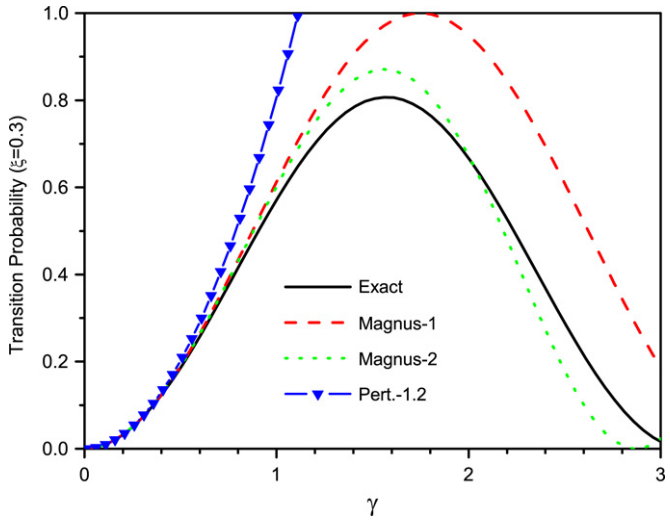


Fig. 7. Rosen–Zener model: Transition probabilities as a function of γ, with $\xi = 0.3$. Lines are coded as in Fig. 6. Computations up to second order, in the Interaction Picture.

In Figs. 6 and 7 we plot some results from the formulae in (204). In Fig. 6 we take $\gamma = 1.5$ and in Fig. 7 we set $\xi = 0.3$.

Similarly to the case of the rectangular step, we observe in Fig. 6 that *the Magnus expansion works better in the sudden regime* defined by $\xi \ll 1$, namely, when the internal time of the system $\hbar/2E$ is much larger than the time scale T of the perturbation. Also, Fig. 7 illustrates how the approximations spoil as far as the intensity γ increases. Notice, in both figures, the unitarity violation of the approximation built with the usual perturbation theory.

In this case, a simple calculation shows that

$$\int_{-\infty}^{\infty} \|\tilde{H}_I(t)\|_2 \, \mathrm{d}t = (1/\hbar) \int_{-\infty}^{\infty} |V(t)| \, \mathrm{d}t = V_0 \pi T/\hbar = \gamma,$$

and thus the Magnus series converges at least for $\gamma < \pi$.

4.2. ME in the Adiabatic Picture

Here we will illustrate the effect of using the Adiabatic Picture introduced in Section 2.9. The use of this type of preliminary transformation is convenient whenever the time scale of the interaction is much larger than the proper time of the unperturbed system.

Since an adiabatic regime conveys a smooth profile for the perturbation, namely, existence of derivatives, the case of the rectangular step cannot be properly used for the sake of illustration.

4.2.1. Linearly forced harmonic oscillator

For the linearly driven harmonic oscillator, the procedure yields the exact solution, as in the preceding Section 4.1.2, albeit the method is a bit more technically involved [120].

4.2.2. Rosen–Zener model

Following [122] we will deal with the Rosen–Zener model (see Section 4.1.2 and [15]) since it allows a clear illustration of the adiabatic regime.

The preliminary linear transformation $G(s)$ defined in (94) for the Hamiltonian (202) is given by $G(s) = \hat{\mathbf{b}} \cdot \boldsymbol{\sigma}$, where the unit vector $\hat{\mathbf{b}}$ points in the direction $\mathbf{b} = \hat{\mathbf{a}} + \hat{\mathbf{k}}$ ($\hat{\mathbf{k}}$ = unit vector along the z-axis). Remember that $s = t/T$ is the dimensionless time-variable. The evolution operator then gets factorized as

$$U_G(s) = G^\dagger(s) U(s) G(s_0), \tag{205}$$

which, according to (90), satisfies the equation

$$\frac{\mathrm{d}U_G}{ds} = \tilde{H}_G(s) U_G. \tag{206}$$

Here $\tilde{H}_G \equiv -\mathrm{i}H_G/\hbar$ is given by

$$\tilde{H}_G(s) = \frac{T}{\mathrm{i}\hbar} a\sigma_3 - \mathrm{i}\frac{\theta'}{2}\sigma_2, \tag{207}$$

with $a^2 = E_0^2 + V_0^2/\cosh^2 s$ and $\cot\theta = (E_0/V_0)\cosh s$.

Next, in analogy to (203), we introduce the Adiabatic Interaction Picture which allows us to integrate the diagonal piece of $\tilde{H}_G(s)$. The time-evolution operator gets eventually factorized as

$$U_G(s) = \exp\left((-\mathrm{i}T/\hbar)\int_0^\infty \mathrm{d}s'\, a(s')\sigma_3\right) U_G^{(I)}(s) \exp\left((-\mathrm{i}T/\hbar)\int_{-\infty}^0 \mathrm{d}s'\, a(s')\sigma_3\right), \tag{208}$$

where $U_G^{(I)}(s)$ obeys the equation

$$\frac{\mathrm{d}U_G^{(I)}}{ds} = \tilde{H}_G^{(I)}(s) U_G^{(I)}, \tag{209}$$

with

$$\tilde{H}_G^{(I)}(s) = -\mathrm{i}(\theta'/2)[\sigma_1 \sin A(s) + \sigma_2 \cos A(s)], \tag{210}$$

and

$$A(s) = \frac{2T}{\hbar}\int_0^s \mathrm{d}s'\, a(s') = \frac{\xi}{2}\ln\frac{1+\rho}{1-\rho} + \frac{2\gamma}{\pi}\arctan\frac{2\gamma}{\pi\xi}\rho. \tag{211}$$

We have introduced the definition

$$\rho = \{1 - [1 + (\pi\xi/2\gamma)]\sin^2\theta\}^{1/2}, \tag{212}$$

in terms of the dimensionless strength parameter $\gamma = \pi V_0 T/\hbar$ and θ. Using the ME to first order in the adiabatic basis (which coincides with the fixed one at $s = \pm\infty$), one finds the spin-flip approximate (first order) transition probability

$${}^{ad}P_M^{(1)} = \sin^2\left[\int_0^{\theta_0} \mathrm{d}\theta \sin A(s(\theta))\right]. \tag{213}$$

In Fig. 8 we compare the numerical results given by the new approximation (213) with the exact formula P_{ex} in (204). For the sake of illustration, we also plot the results in the usual Interaction Picture to first order in ME (see $P_M^{(1)}$ in (204)). The gain achieved when using the adiabatic ME in the intermediate regime (i.e., moderate values of ξ) is of note, although only the first order is considered. Here, the adiabatic regime corresponds to large values of $\xi = 2ET/\hbar$ ($\varepsilon = 1/T \ll 1$).

It should also be noted that, for the Hamiltonian (210) one has

$$\int_{s_0}^s \|\tilde{H}_G^{(I)}(s_1)\|_2 \mathrm{d}s_1 = \frac{1}{2}|\theta(s) - \theta(s_0)| < \frac{1}{2}2\pi = \pi$$

and thus the convergence condition given by Theorem 9 is always satisfied. In other words, for this example *the Magnus expansion is always convergent in the Adiabatic Interaction Picture*.

More involved illustrative examples of ME in the Adiabatic Picture may be found in [120,123].

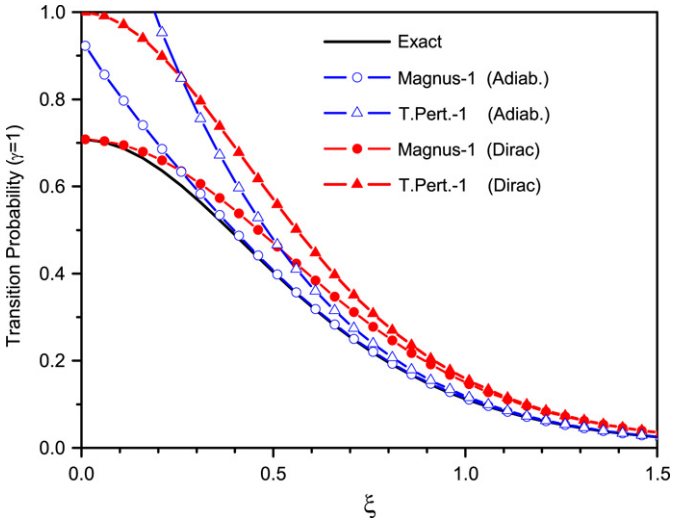


Fig. 8. Rosen–Zener model: Transition probability as a function of ξ, with $\gamma = 1$. The solid line stands for the exact result in (204). The remaining lines stand for first order computations with ME (circles) and perturbation theory (triangles). Open symbols are for the Adiabatic Picture and solid symbols for the Interaction Picture.

4.3. Fer and Wilcox infinite-product expansions

Next we illustrate the use of Fer and Wilcox infinite-product expansions by using the same time-dependent systems as before.

4.3.1. Linearly forced harmonic oscillator

Since the commutator $[a_-, a_+]$ is a c-number, Fer iterated Hamiltonians $H^{(n)}$ with $n > 1$ eventually vanish so that $F_n = 0$ for $n > 2$. The Wilcox operators W_n with $n > 2$ in eq. (106) vanish for the same reason. Thus, in this particular case, the second-order approximation in either method leads to the exact solution of the Schrödinger equation, just as ME did. To sum up, the final result reads

$$U_I = \mathrm{e}^{\Omega_1+\Omega_2} = \mathrm{e}^{W_1}\mathrm{e}^{W_2} = \mathrm{e}^{F_1}\mathrm{e}^{F_2} = \mathrm{e}^{-\mathrm{i}\beta}\mathrm{e}^{-\mathrm{i}(\alpha a_+ + \alpha^* a_-)}, \tag{214}$$

where $\alpha(t)$ and $\beta(t)$ are given in (185) and (186), respectively.

4.3.2. Two-level quantum system: Rectangular step

For the Hamiltonian (194), the first-order Fer and Wilcox operators in the Interaction Picture verify

$$F_1 = W_1 = \Omega_1 = \int_0^t \mathrm{d}t_1 \tilde{H}_I(t_1), \tag{215}$$

where $\tilde{H}_I(t)$ is given by (198). The explicit expression is collected in (200) (first equation). Analogously, the second equation there also corresponds to the second-order Wilcox operator $W_2 = \Omega_2$.

To proceed further with Fer's method, we must calculate the modified Hamiltonian $\tilde{H}^{(1)}$ in (96). After straightforward algebra, one eventually obtains

$$\tilde{H}^{(1)} = \frac{1}{2\theta}\left(\frac{\sin^2\theta}{\theta} - \sin 2\theta + \frac{1}{\theta}\left(\frac{\sin 2\theta}{2\theta} - \cos 2\theta\right) F_1\right)[F_1, \tilde{H}_I], \tag{216}$$

where $\theta = (2\gamma/\xi)\sin(\xi/2)$ (notice that $\tilde{H}^{(1)}$ and therefore F_2 depend on σ_1 and σ_2, while W_2 is proportional to σ_3). Since it does not seem possible to derive an analytical expression for F_2, the corresponding matrix elements have been computed by replacing the integral by a conveniently chosen quadrature.

The transition probability $P(t)$ from an initial state with spin up to a state with spin down (or vice versa) is given by (193). This expression has been computed assuming $U_I \simeq \mathrm{e}^{F_1} = \mathrm{e}^{W_1}$, $U_I \simeq \mathrm{e}^{F_1}\mathrm{e}^{F_2}$, $U_I \simeq \mathrm{e}^{W_1}\mathrm{e}^{W_2}$, $U_I \simeq \mathrm{e}^{F_1}\mathrm{e}^{F_2}\mathrm{e}^{F_3}$ and $U_I \simeq \mathrm{e}^{W_1}\mathrm{e}^{W_2}\mathrm{e}^{W_3}$, and the results have been compared with the exact analytical solution (196).

In Figs. 9 and 10 we show the transition probability P as a function of ξ for two different values of γ ($\gamma = 1.5$ and $\gamma = 2$, respectively), while in Fig. 11 we have plotted P versus γ for ξ fixed. Notice that the second order in the Wilcox expansion does not contribute to the transition probability (this is similar to what happens in perturbation theory). On the other hand, Fer's second-order approximation is already in remarkable agreement with the exact result, whereas the third order cannot even be distinguished from the exact result in Fig. 10 at the cost of a much larger computational effort. Wilcox approximants preserve unitarity but do not give acceptable approximations.

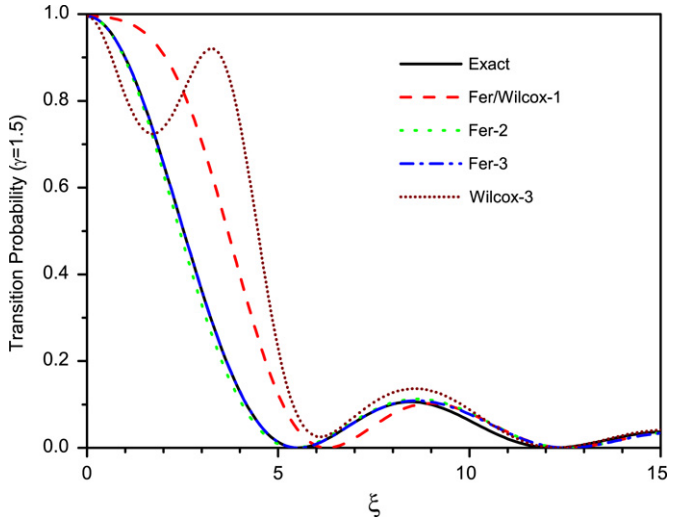


Fig. 9. Rectangular step: Transition probability as a function of ξ, with $\gamma = 1.5$. The solid line corresponds to the exact result (196). Broken lines stand for Fer and Wilcox approximations up to third order. Computations up to third order, in the Interaction Picture.

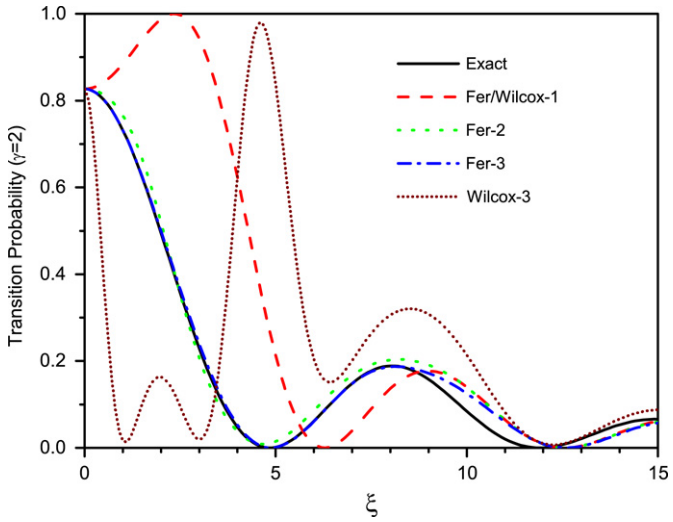


Fig. 10. Rectangular step: Transition probability as a function of ξ, with $\gamma = 2$. Lines are coded as in Fig. 9. Computations up to third order, in the Interaction Picture.

4.4. Voslamber iterative method

Just to keep the same structure as in preceding subsections, we mention that the Voslamber iterative method of Section 3.1 also yields the exact solution for the linearly driven harmonic oscillator after computing the second iteration.

Next, for the two-level system with a rectangular step, described by the Hamiltonian (194), we compute the second iterate $\Omega^{(2)}$ and compare with second order ME approximation for U_I. As a test, we shall again obtain the transition probability $P(t)$ given by (196). The expression (193) will be calculated here, assuming: $U_I \simeq \exp \Omega^{(1)} = \exp \Omega_1$, $U_I \simeq \exp(\Omega_1 + \Omega_2)$ and $U_I \simeq \exp \Omega^{(2)}$.

The second order Magnus approximation to the transition probability is given by (201), whereas the second iterate is obtained from (124),

$$\Omega^{(2)}(t) = -\mathrm{i}\frac{\gamma}{\omega}\int_0^{\omega t} \{[\sin^2(\Delta) + \cos^2(\Delta)\cos\xi]\,\sigma_1 + \cos^2(\Delta)\sin\xi\,\sigma_2 - \sin(2\Delta)\sin(\xi/2)\,\sigma_3\}\mathrm{d}\xi\,, \tag{217}$$

where $\Delta \equiv \frac{\gamma}{\omega}|\sin(\xi/2)|$, $\gamma = V_0 t/\hbar$, $\xi = \omega t$. Since it does not seem possible to derive an analytical expression for $\Omega^{(2)}$, the corresponding matrix elements have been computed by approximating the integral in (217) with a sufficiently accurate quadrature.

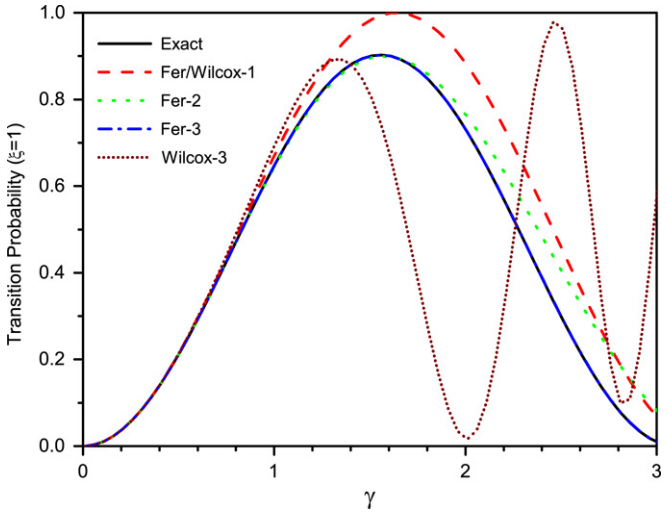


Fig. 11. Rectangular step: Transition probability as a function of γ, with $\xi = 1$. Lines are coded as in Fig. 9. Computations up to third order, in the Interaction Picture.

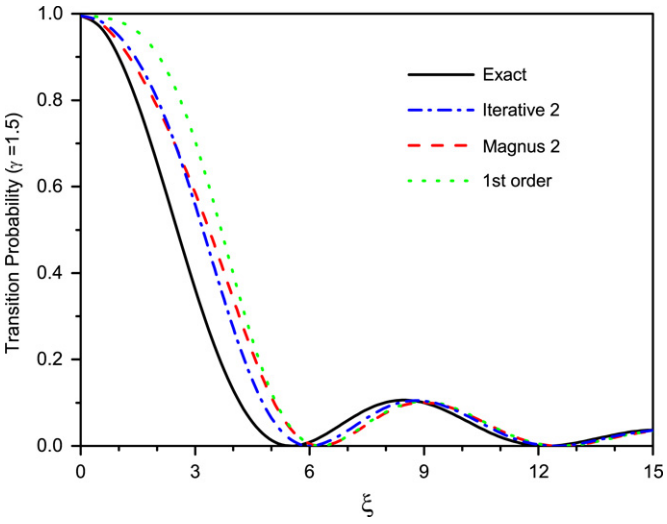


Fig. 12. Rectangular step: Transition probability in the two level system as a function of ξ, for $\gamma = 1.3$: Exact result (204) (solid line), second iterate of Voslamber method (dashed line), second order ME (dotted line) and first Voslamber iterate (or order in ME) (dash-dotted line). Computations are done in the Interaction Picture.

In Fig. 12, the various approximated transition probabilities, as well as the exact result (196), have been plotted as a function of ξ for a fixed value $\gamma = 1.5$. We observe that the approximation from the second iterate keeps the trend of the exact solution in a better way than the second order Magnus approximation does.

In Fig. 13 we have plotted the corresponding transition probabilities versus γ for fixed value $\xi = 1$. Although locally, the second order Magnus approximation may be more accurate, it seems that the trend of the exact solution is mimicked for a longer interval of γ.

As it has already been pointed out above, the Magnus expansion works better the more sudden the perturbation. Thus, the re-summation involved in the iterative method slightly improves that issue. Further results on the present example may be found in [93].

4.5. Linear periodic systems: Floquet–Magnus formalism

Next we deal with a periodically driven harmonic oscillator, where the infinite expansions obtained from Floquet–Magnus recurrences (see Section 3.2), are utterly summed. Comparison with the exact solution illustrates the feasibility of the method.

The particular system we consider is described by the Hamiltonian (179) with $f(t) = \frac{\beta}{\sqrt{2}} \cos \omega t$ and $\omega_0 < \omega$. Once the recurrences of the Floquet–Magnus expansion (see Section 3.2) are explicitly computed for several orders, their general term

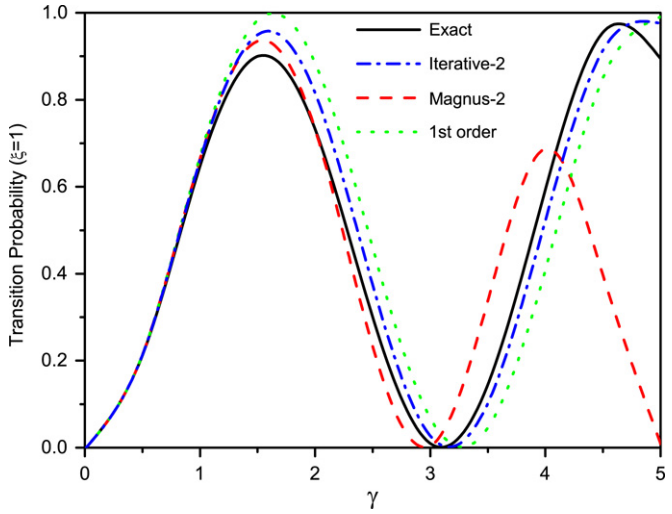


Fig. 13. Rectangular step: Transition probability in the two level system as a function of γ, for $\xi = 1$. Lines are coded as in Fig. 12. Computations are done in the Interaction Picture.

may be guessed by inspection. For the Floquet operator we get

$$F = -\mathrm{i}\left[\frac{\omega_0}{2}\left(p^2+q^2\right) - \beta\frac{\omega_0}{\omega}\sum_{k=0}^{\infty}\left(\frac{\omega_0}{\omega}\right)^k q + \beta^2\frac{\omega_0}{4\omega^2}\sum_{k=0}^{\infty}(2k+1)\left(\frac{\omega_0}{\omega}\right)^k\right], \tag{218}$$

and the associated transformation results from

$$\begin{aligned}\Lambda(t) = {} & \mathrm{i}\beta^2\frac{\omega_0}{\omega^3}\left[\sin(\omega t)\sum_{k=0}^{\infty}(k+1)\left(\frac{\omega_0}{\omega}\right)^k - \frac{\omega t}{2}\sum_{k=0}^{\infty}\left(\frac{\omega_0}{\omega}\right)^k\right] \\ & - \mathrm{i}\left[\sin(\omega t)\, q + \omega_0(\cos(\omega t)-1)\, p\right]\frac{\beta}{\omega}\sum_{k=0}^{\infty}\left(\frac{\omega_0}{\omega}\right)^k.\end{aligned} \tag{219}$$

The resulting series may be summed in closed form, thus yielding the Floquet operator

$$F = -\mathrm{i}\frac{\omega_0}{2}\left[\left(q - \frac{\beta}{\omega_0(\rho^2-1)}\right)^2 + p^2\right] - \mathrm{i}\frac{\beta^2}{4\omega_0\left(\rho^2-1\right)}, \tag{220}$$

with $\rho \equiv \omega/\omega_0$. Its eigenvalues are the so-called *Floquet eigenenergies* [124]

$$E_n = \omega_0\left(n+\frac{1}{2}\right) + \frac{\beta^2}{4\omega_0\left(\rho^2-1\right)}. \tag{221}$$

The corresponding Λ transformation after summation of the series in (219) is

$$\Lambda(t) = \mathrm{i}\frac{\beta/\omega_0}{1-\rho^2}\left[(\rho\sin\omega t)\, q + (\cos\omega t - 1)p + \left(\frac{2\rho^2}{1-\rho^2} + \cos\omega t\right)\frac{\beta\sin\omega t}{4\omega}\right]. \tag{222}$$

Notice that, as they should, both operators are skew-Hermitian and reproduce the exact solution of the problem.

5. Numerical integration methods based on the Magnus expansion

5.1. Introduction

The Magnus expansion, as formulated in Section 2, has found extensive use in mathematical physics, quantum chemistry, control theory etc, essentially as a perturbative tool in the treatment of the linear equation

$$Y' = A(t)Y, \qquad Y(t_0) = Y_0. \tag{223}$$

When the recurrence (49)–(50) is applied, one is able to explicitly get the successive terms Ω_k in the series, defining Ω as linear combinations of multivariate integrals containing commutators acting iteratively on the coefficient matrix A, as in

(42)–(46). As a result, with this scheme, analytical approximations to the exact solution are constructed explicitly. These approximate solutions are fairly accurate inside the convergence domain, especially when high order terms in the Magnus series are taken into account, as illustrated by the examples considered in Section 4.

There are several drawbacks, however, involved in the procedure developed so far, especially when one tries to find accurate approximations to the solution for very long times. The first one is implicitly contained in the analysis done in Section 2: the size of the convergence domain of the Magnus series may be relatively small. The logarithm of the exact solution $Y(t)$ may have complex singularities and this implies that no series expansion can converge beyond the first singularity. This disadvantage may be, up to some point, avoided by using different pictures (e.g. the transformations shown in Section 2.9) in order to increase the convergence domain of the Magnus expansion, a fact also illustrated by several examples in Section 4. Unfortunately, these preliminary transformations sometimes either do not guarantee convergence, or the rate of convergence of the series is very slow. In that case, accurate results can only be obtained provided a large number of terms in the series are taken into account.

The second drawback is the increasingly complex structure of the terms Ω_k in the Magnus series: each Ω_k is a k-multivariate integral involving a linear combination of $(k-1)$-nested commutators of A evaluated at different times t_i, $i = 1, \ldots, k$. Although in some cases these expressions can be computed explicitly (for instance, when the elements of A and its commutators are polynomial or trigonometric functions), in general a special procedure has to be designed to approximate multivariate integrals and reduce the number of commutators involved.

When the entries of the coefficient matrix $A(t)$ are complicated functions of time or they are only known for certain values of t, numerical approximation schemes are unavoidable. In many cases, it is thus desirable to obtain numerical approximations to the exact solution at many different times. This section is devoted precisely to the Magnus series expansion as a tool for building numerical integrators for Eq. (223).

Before embarking on exposing the technical details contained in this construction, let us first introduce several concepts which are commonplace in the context of the numerical integration of differential equations.

Given the general (nonlinear) ordinary differential equation (ODE)

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}), \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \in \mathbb{C}^d, \tag{224}$$

standard numerical integrators, such as Runge–Kutta and multistep methods, proceed as follows. First the whole time interval $[t_0, t_f]$, is split into N subintervals, $[t_{n-1}, t_n]$, $n = 1, \ldots, N$, with $t_N = t_f$, and subsequently the value of $\mathbf{x}(t_n)$ is approximated with a time-stepping advance procedure of the form

$$\mathbf{x}_{n+1} = \Phi(h_n, \mathbf{x}_n, \ldots, h_0, \mathbf{x}_0) \tag{225}$$

starting from $\mathbf{x}_0$. Here the map Φ depends on the specific numerical method and $h_n = t_{n+1} - t_n$ are the time steps. For simplicity in the presentation, we consider a constant time step h, so that $t_n = t_0 + n\,h$. In this way, one gets $\mathbf{x}_{n+1}$ as an approximation to $\mathbf{x}(t_{n+1})$. In other words, the exact evolution of the system (224) is replaced by the discrete or numerical flow (225). The simplest of all numerical methods for (224) is the *explicit Euler scheme*

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h\,\mathbf{f}(t_n, \mathbf{x}_n). \tag{226}$$

It computes approximations $\mathbf{x}_n$ to the values $\mathbf{x}(t_n)$ of the solution using one explicit evaluation of $\mathbf{f}$ at the already computed value $\mathbf{x}_{n-1}$. In general, the numerical method (225) is said to be of order p if, assuming $\mathbf{x}_n = \mathbf{x}(t_n)$, then $\mathbf{x}_{n+1} = \mathbf{x}(t_{n+1}) + \mathcal{O}(h^{p+1})$. Thus, in particular, Euler method is of order one.

Of course, more elaborate and efficient general purpose algorithms, using several $\mathbf{f}$ evaluations per step, have been proposed along the years for the numerical treatment of Eq. (224). In fact, any standard software package and program library contains dozens of routines aimed at providing numerical approximations with several degrees of accuracy, including (explicit and implicit) Runge–Kutta methods, linear multistep methods, extrapolation schemes etc, with fixed or adaptive step size. They are designed in such a way that the user has to provide only the initial condition and the function $\mathbf{f}$ to obtain approximations at any given time.

This being the case, one could ask the following question: if general purpose integrators are widely available for the integration of the linear equation (223) (which is a particular case of (224)), what is the point of designing new and somehow sophisticated algorithms for this specific problem?

It turns out that, in the same way as for classical time-dependent perturbation theory, the qualitative properties of the exact solution are not preserved by the numerical approximations obtained by standard integrators. This motivates the study of the Magnus expansion, with the ultimate goal of constructing numerical integration methods. We will show that highly accurate schemes can, indeed, be designed, which in addition preserve qualitative properties of the system. The procedure can also incorporate the tools developed for the analytical treatment, such as preliminary linear transformations, to end up with improved numerical algorithms.

We first summarize the main features of the well know class of Runge–Kutta methods as representative of integrators of the form (225). They are introduced for the general nonlinear ODE (224) and are subsequently adapted to the linear case (223).

5.2. Runge–Kutta methods

The Runge–Kutta (RK) class of methods are possibly the most frequently used algorithms for numerically solving ODEs. Among them, perhaps the most successful during more than half a century has been the 4th-order method, which, applied to Eq. (224), provides the following numerical approximation for the integration step $t_n \mapsto t_{n+1} = t_n + h$:

$$\begin{aligned}
\mathbf{Y}_1 &= \mathbf{x}_n \\
\mathbf{Y}_2 &= \mathbf{x}_n + \frac{h}{2}\mathbf{f}(t_n, \mathbf{Y}_1) \\
\mathbf{Y}_3 &= \mathbf{x}_n + \frac{h}{2}\mathbf{f}\left(t_n + \frac{h}{2}, \mathbf{Y}_2\right) \\
\mathbf{Y}_4 &= \mathbf{x}_n + h\mathbf{f}\left(t_n + \frac{h}{2}, \mathbf{Y}_3\right) \\
\mathbf{x}_{n+1} &= \mathbf{x}_n + \frac{h}{6}\left(\mathbf{f}(t_n, \mathbf{Y}_1) + 2\mathbf{f}\left(t_n + \frac{h}{2}, \mathbf{Y}_2\right) + 2\mathbf{f}\left(t_n + \frac{h}{2}, \mathbf{Y}_3\right) + \mathbf{f}(t_n + h, \mathbf{Y}_4)\right).
\end{aligned} \tag{227}$$

Notice that the function $\mathbf{f}$ can always be computed explicitly because each $\mathbf{Y}_i$ depends only on the $\mathbf{Y}_j$, $j < i$, previously evaluated. To measure the computational cost of the method, it is usual to consider that the evaluation of the function $\mathbf{f}(t, \mathbf{x})$ is the most consuming part. In this sense, scheme (227) requires four evaluations, which is precisely the number of *stages* (or inner steps) in the algorithm.

The general class of s-stage Runge–Kutta methods are characterized by the real numbers a_{ij}, b_i $(i, j = 1, \ldots, s)$ and $c_i = \sum_{j=1}^{s} a_{ij}$, as

$$\begin{aligned}
\mathbf{Y}_i &= \mathbf{x}_n + h\sum_{j=1}^{s} a_{ij}\,\mathbf{f}(t_n + c_j h, \mathbf{Y}_j), \quad i = 1, \ldots, s \\
\mathbf{x}_{n+1} &= \mathbf{x}_n + h\sum_{i=1}^{s} b_i\,\mathbf{f}(t_n + c_i h, \mathbf{Y}_i),
\end{aligned} \tag{228}$$

where $\mathbf{Y}_i$, $i = 1, \ldots, s$ are the intermediate stages. For simplicity, the associated coefficients are usually displayed with the so-called *Butcher tableau* [125,126] as follows:

$$\begin{array}{c|ccc}
c_1 & a_{11} & \ldots & a_{1s} \\
\vdots & \vdots & & \vdots \\
c_s & a_{s1} & \ldots & a_{ss} \\
\hline
 & b_1 & \ldots & b_s
\end{array} \tag{229}$$

If $a_{ij} = 0, j \geq i$, then the intermediate stages $\mathbf{Y}_i$ can be evaluated recursively and the method is explicit. In that case, the zero a_{ij} coefficients (in the upper triangular part of the tableau) are omitted for clarity. With this notation, 'the' 4th-order Runge–Kutta method (227) can be expressed as

$$\begin{array}{c|cccc}
0 & & & & \\
\frac{1}{2} & \frac{1}{2} & & & \\
\frac{1}{2} & 0 & \frac{1}{2} & & \\
1 & 0 & 0 & 1 & \\
\hline
 & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6}
\end{array} \tag{230}$$

Otherwise, the scheme is implicit and requires one to numerically solve a system of $s\,d$ nonlinear equations of the form

$$\mathbf{y} = \mathbf{X}_n + h\,\mathbf{G}(h, \mathbf{y}), \tag{231}$$

where $\mathbf{y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_s)^{\mathrm{T}}$, $\mathbf{X}_n = (\mathbf{x}_n, \ldots, \mathbf{x}_n)^{\mathrm{T}} \in \mathbb{R}^{sd}$, and $\mathbf{G}$ is a function which depends on the method. A standard procedure to get $\mathbf{x}_{n+1}$ from (231) is applying simple iteration:

$$\mathbf{y}^{[j]} = \mathbf{X}_n + h\,\mathbf{G}(h, \mathbf{y}^{[j-1]}), \quad j = 1, 2, \ldots. \tag{232}$$

When h is sufficiently small, the iteration starts with $\mathbf{y}^{[0]} = \mathbf{X}_n$ and stops once $\|\mathbf{y}^{[j]} - \mathbf{y}^{[j-1]}\|$ is smaller than a prefixed tolerance. Of course, more sophisticated techniques can be used [126].

After these general considerations, let us now turn our attention to the linear equation (223). When dealing with numerical methods applied to this equation, it is important to keep in mind that the relevant small parameter here is no longer the norm of the matrix $A(t)$ as in the analytical treatment, but the time step h. For this reason, the concept of order of

accuracy of a method is different in this context. Now $\mathbf{f}(t,\mathbf{x}) = A(t)\mathbf{x}$ and the general class of s-stage Runge–Kutta methods adopt the more compact form

$$\mathbf{Y}_i = \mathbf{x}_n + h\sum_{j=1}^{s} a_{ij}A_j\mathbf{Y}_j, \quad i = 1,\ldots,s$$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h\sum_{i=1}^{s} b_iA_i\mathbf{Y}_i, \tag{233}$$

with $A_i = A(t_n + c_ih)$. In terms of matrices, this is equivalent to

$$\begin{Bmatrix}\mathbf{Y}_1\\ \vdots\\ \mathbf{Y}_s\end{Bmatrix} = \begin{Bmatrix}\mathbf{x}_n\\ \vdots\\ \mathbf{x}_n\end{Bmatrix} + h\widetilde{A}\begin{Bmatrix}\mathbf{Y}_1\\ \vdots\\ \mathbf{Y}_s\end{Bmatrix}, \quad \text{with } \widetilde{A} = \begin{pmatrix} a_{11}A_1 & \cdots & a_{1s}A_s\\ \vdots & & \vdots\\ a_{s1}A_1 & \cdots & a_{ss}A_s\end{pmatrix}$$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h\begin{pmatrix} b_1A_1 & \cdots & b_sA_s\end{pmatrix}\begin{Bmatrix}\mathbf{Y}_1\\ \vdots\\ \mathbf{Y}_s\end{Bmatrix}, \tag{234}$$

so that the application of the method for the integration step $t_n \mapsto t_{n+1} = t_n + h$ can also be written as

$$\mathbf{x}_{n+1} = \left(I_d + h\begin{pmatrix} b_1A_1 & \cdots & b_sA_s\end{pmatrix}\left(I_{sd} - h\widetilde{A}\right)^{-1}\mathbb{I}_{sd\times d}\right)\mathbf{x}_n, \tag{235}$$

where $\mathbb{I}_{sd\times d} = (I_d, I_d, \ldots, I_d)^{\mathrm{T}}$ and I_d is the $d \times d$ identity matrix. For instance, taking $s = 2$ and using Matlab this can be easily implemented as follows

$A = [a11 * A1 \;\; a12 * A2 \,;\; a21 * A1 \;\; a22 * A2];$
$x = (Id + h * [b1 * A1 \;\; b2 * A2] * ((I2d - h * A) \setminus [Id \;\; Id]')) * x;$

where $b1, b2, a11, \ldots, a22$ are the coefficients of the method, $A1, A2$ correspond to A_1, A_2 and Id, $I2d$ are the identity matrices of dimension d and $2d$, respectively.

There is extensive literature regarding RK methods built for many different purposes [125,126]. It is therefore reasonable to look for the most appropriate scheme to be used for each problem. In practice, explicit RK methods are preferred, because its implementation is usually simpler. They typically require more stages than implicit methods and, in general, a higher number of evaluations of the matrix $A(t)$, although this is not always the case. For instance, the 4-stage fourth order method (227) only requires two new evaluations of $A(t)$ per step ($A(t_n + h/2)$ and $A(t_n + h)$), which is precisely the minimum number of evaluations needed by any fourth order method. In our numerical examples, we will also use the 7-stage sixth order method with coefficients [125, p. 203-205]

$$\begin{array}{c|ccccccc}
0 & & & & & & & \\
\frac{5-\sqrt{5}}{10} & \frac{5-\sqrt{5}}{10} & & & & & & \\
\frac{5+\sqrt{5}}{10} & \frac{-\sqrt{5}}{10} & \frac{5+2\sqrt{5}}{10} & & & & & \\
\frac{5-\sqrt{5}}{10} & \frac{-15+7\sqrt{5}}{20} & \frac{-1+\sqrt{5}}{4} & \frac{15-7\sqrt{5}}{10} & & & & \\
\frac{5+\sqrt{5}}{10} & \frac{5-\sqrt{5}}{60} & 0 & \frac{1}{6} & \frac{15+7\sqrt{5}}{60} & & & \\
\frac{5-\sqrt{5}}{10} & \frac{5+\sqrt{5}}{60} & 0 & \frac{9-5\sqrt{5}}{12} & \frac{1}{6} & \frac{-5+3\sqrt{5}}{10} & & \\
1 & \frac{1}{6} & 0 & \frac{-55+25\sqrt{5}}{12} & \frac{-25-7\sqrt{5}}{12} & 5-2\sqrt{5} & \frac{5+\sqrt{5}}{2} & \\
\hline
 & \frac{1}{12} & 0 & 0 & 0 & \frac{5}{12} & \frac{5}{12} & \frac{1}{12}
\end{array} \tag{236}$$

Observe that this method only requires three new evaluations of the matrix $A(t)$ per step. This is, in fact, the minimum number for a sixth-order method. Other implicit RK schemes widely used in the literature involving the minimum number of stages at each order are based on Gauss–Legendre collocation points [126]. For instance, the corresponding methods of order four and six (with two and three stages, respectively) have the following coefficients:

$$\begin{array}{c|cc}
\frac{3-\sqrt{3}}{6} & \frac{1}{4} & \frac{3-2\sqrt{3}}{12}\\
\frac{3+\sqrt{3}}{6} & \frac{3+2\sqrt{3}}{12} & \frac{1}{4}\\
\hline
 & \frac{1}{2} & \frac{1}{2}
\end{array}
\qquad
\begin{array}{c|ccc}
\frac{5-\sqrt{15}}{10} & \frac{5}{36} & \frac{2}{9}-\frac{\sqrt{15}}{15} & \frac{5}{36}-\frac{\sqrt{15}}{30}\\
\frac{1}{2} & \frac{5}{36}+\frac{\sqrt{15}}{24} & \frac{2}{9} & \frac{5}{36}+\frac{\sqrt{15}}{24}\\
\frac{5+\sqrt{15}}{10} & \frac{5}{36}+\frac{\sqrt{15}}{30} & \frac{2}{9}+\frac{\sqrt{15}}{15} & \frac{5}{36}\\
\hline
 & \frac{5}{18} & \frac{4}{9} & \frac{5}{18}
\end{array} \tag{237}$$

5.3. Preservation of qualitative properties

Notice that the numerical solution provided by the class of Runge–Kutta schemes, and in general by an integrator of the form (225), is constructed as a sum of vectors in $\mathbb{R}^d$. Let us point out some (undesirable) consequences of this fact. Suppose, for instance, that $\mathbf{x}$ is a vector known to evolve on a sphere. One does not expect that $\mathbf{x}_n$ built as a sum of two vectors as in (225), preserves this feature of the exact solution, whereas approximations of the form $\mathbf{x}_{n+1} = Q_n\mathbf{x}_n$, with Q_n an orthogonal matrix, clearly lie on the sphere.

In Section 3.3 we have introduced isospectral flows (143), which include, as a particular case, the system

$$Y' = [A(t), Y], \qquad Y(t_0) = Y_0, \tag{238}$$

with A a skew-symmetric matrix. As we have shown there, if Y_0 is a symmetric matrix, the exact solution can be factorized as $Y(t) = Q(t)Y_0Q^{\mathrm{T}}(t)$, with $Q(t)$ an orthogonal matrix satisfying the equation

$$Q' = A(t)Q, \qquad Q(0) = I \tag{239}$$

and, in addition, $Y(t)$ and $Y(0)$ have the same eigenvalues. This is also true when Y is Hermitian, A is skew-Hermitian and Q is unitary, in which case (238) and (239) can be interpreted as particular examples of the Heisenberg and Schrödinger equations in Quantum Mechanics, respectively.

When a numerical scheme of the form (225) is applied to (239), in general, the approximations Q_n will no longer be unitary matrices and therefore Y_n and Y_0 will not be unitarily similar. As a result, the isospectral character of the system (238) is lost in the numerical description. Observe that explicit Runge–Kutta methods employ the ansatz, that locally the solution of the differential equation behaves like a polynomial in t, so that one cannot expect the approximate solution to be a unitary matrix. In this sense, explicit Runge–Kutta methods present the same drawbacks in the numerical analysis of differential equations as the standard time-dependent perturbation theory, when looking for analytical approximations.

Implicit Runge–Kutta methods, on the other hand, can be considered as rational approximations, and in some cases the outcome they provide is a unitary matrix. For this class of methods, however, matrices of relatively large sizes have to be inverted, making the algorithms computationally expensive. Furthermore, the evolution of many systems (including highly oscillatory problems) cannot be efficiently approximated, either by polynomial or by rational approximations.

With all these considerations in mind, a pair of questions arise in a quite natural way:

(Q1) Is it possible to design numerical integration methods for Eq. (224) such that the corresponding numerical approximations still preserve the main qualitative features of the exact solution?
(Q2) Since the Magnus expansion constitutes an extremely useful procedure for obtaining analytical (as opposed to numerical) approximate solutions to the linear evolution equation (223), is it feasible to construct efficient numerical integration schemes from the general formulation exposed in Section 2?

It turns out that both questions can be answered in the affirmative. As a matter of fact, it has been in trying to address (Q1), that the field of *geometric numerical integration* has developed during the last few years. Here, the aim is to reproduce the qualitative features of the solution of the differential equation which is being discretized, in particular its geometric properties. The motivation for developing such structure-preserving algorithms arises independently in areas of research as diverse as celestial mechanics, molecular dynamics, control theory, particle accelerators physics, and numerical analysis [59,79,127–129].

Although diverse, the systems appearing in these areas have one important common feature. They all preserve some underlying geometric structure which influences the qualitative nature of the phenomena they produce. In the field of geometric numerical integration, these properties are built into the numerical method, which gives the method an improved qualitative behavior, but also allows for a significantly more accurate long-time integration than with general-purpose methods.

In addition to the construction of new numerical algorithms, an important aspect of geometric integration is the explanation of the relationship between preservation of the geometric properties of a numerical method and the observed favorable error propagation in long-time integration.

Geometric numerical integration has been an active and interdisciplinary research area during the last decade, and nowadays is the subject of intensive development. Perhaps the most familiar examples of *geometric integrators* are symplectic integration algorithms in classical Hamiltonian dynamics, symmetric integrators for reversible systems and methods preserving first integrals and numerical methods on manifolds [79].

A particular class of geometric numerical integrators are the so-called *Lie-group methods*. If the matrix $A(t)$ in (223) belongs to a Lie algebra $\mathfrak{g}$, the aim of Lie-group methods is to construct numerical solutions staying in the corresponding Lie group $\mathcal{G}$ [59].

With respect to question (Q2) above, it will be the subject of the next subsection, where the main issues involved in the construction of numerical integrators based on the Magnus expansion are analyzed. The methods thus obtained, preserve the qualitative properties of the system and, in addition, are highly competitive with other more conventional numerical schemes, with respect to accuracy and computational effort. They constitute, therefore, clear examples of Lie-group methods.

5.4. Magnus integrators for linear systems

Since the Magnus series only converges locally, as we have pointed out before, when the length of the time-integration interval exceeds the bound provided by Theorem 9, and in the spirit of any numerical integration method, the usual procedure consists of dividing the time interval $[t_0, t_f]$ into N steps such that the Magnus series converges in each subinterval $[t_{n-1}, t_n]$, $n = 1, \ldots, N$, with $t_N = t_f$. In this way, the solution of (223) at the final time t_f is represented by

$$Y(t_N) = \prod_{n=1}^{N} \exp(\Omega(t_n, t_{n-1}))Y_0, \tag{240}$$

and the series $\Omega(t_n, t_{n-1})$ has to be appropriately truncated.

Early steps in this approach were taken in [21,130], where second and fourth order numerical schemes were used for calculations of collisional inelasticity and potential scattering, respectively. Those authors were well aware that the resulting integration method "would become practical only when the advantage of being able to use bigger step sizes outweighs the disadvantage in having to evaluate the integrals involved in the Magnus series, and then doing the exponentiation" [130]. Later on, by following a similar approach, Devries [131] designed a numerical procedure for determining a fourth order approximation to the propagator employed in the integration of the single channel Schrödinger equation, but it was in the pioneering work [37] where Iserles and Nørsett carried out the first systematic study of the Magnus expansion with the aim of constructing numerical integration algorithms for linear problems. To design the new integrators, the explicit time dependency of each term Ω_k had to be analyzed, in particular its order of approximation in time to the exact solution.

Generally speaking, the process of rendering the Magnus expansion, a practical numerical integration algorithm, involves three steps. First, the Ω series is truncated at an appropriate order. Second, the multivariate integrals in the truncated series $\Omega^{[p]} = \sum_{i=1}^{p} \Omega_i$ are replaced by conveniently chosen approximations. Third, the exponential of the matrix $\Omega^{[p]}$ has to be computed. We now briefly consider the first two issues, whereas the general problem of evaluating the matrix exponential will be treated in Section 5.6.

We have shown in Section 2.6 that the Magnus expansion is time symmetric. As a consequence of Eq. (66), if $A(t)$ is analytic and one evaluates its Taylor series centered around the midpoint of a particular subinterval $[t_n, t_n + h]$, then each term in Ω_k is an odd function of h, and thus $\Omega_{2s+1} = \mathcal{O}(h^{2s+3})$ for $s \geq 1$. Equivalently, $\Omega^{[2s-2]} = \Omega + \mathcal{O}(h^{2s+1})$ and $\Omega^{[2s-1]} = \Omega + \mathcal{O}(h^{2s+1})$. In other words, for achieving an integration method of order $2s$ ($s > 1$) only terms up to Ω_{2s-2} in the Ω series are required [37,39]. For this reason, in general, only even order methods are considered.

Once the series expansion is truncated up to an appropriate order, the multidimensional integrals involved have to be computed or at least conveniently approximated. Although, at first glance, this seems to be a quite difficult enterprise, it turns out that their very structure allows one to approximate all the multivariate integrals appearing in Ω just by evaluating $A(t)$ at the nodes of a univariate quadrature [37].

To illustrate how this task can be accomplished, let us expand the matrix $A(t)$ around the midpoint, $t_{1/2} \equiv t_n + h/2$, of the subinterval $[t_n, t_{n+1}]$,

$$A(t) = \sum_{j=0}^{\infty} a_j \left(t - t_{1/2}\right)^j, \quad \text{where } a_j = \frac{1}{j!} \left.\frac{\mathrm{d}^j A(t)}{dt^j}\right|_{t=t_{1/2}}, \tag{241}$$

and insert the series (241) into the recurrence defining the Magnus expansion (49)–(50). In this way one gets, explicitly, the expression of Ω_k up to order h^6 as

$$\begin{aligned}
\Omega_1 &= h a_0 + h^3 \frac{1}{12} a_2 + h^5 \frac{1}{80} a_4 + \mathcal{O}(h^7) \\
\Omega_2 &= h^3 \frac{-1}{12} [a_0, a_1] + h^5 \left(\frac{-1}{80} [a_0, a_3] + \frac{1}{240} [a_1, a_2] \right) + \mathcal{O}(h^7) \\
\Omega_3 &= h^5 \left(\frac{1}{360} [a_0, a_0, a_2] - \frac{1}{240} [a_1, a_0, a_1] \right) + \mathcal{O}(h^7) \\
\Omega_4 &= h^5 \frac{1}{720} [a_0, a_0, a_0, a_1] + \mathcal{O}(h^7),
\end{aligned} \tag{242}$$

whereas $\Omega_5 = \mathcal{O}(h^7)$, $\Omega_6 = \mathcal{O}(h^7)$ and $\Omega_7 = \mathcal{O}(h^9)$. Here we write for clarity $[a_{i_1}, a_{i_2}, \ldots, a_{i_{l-1}}, a_{i_l}] \equiv [a_{i_1}, [a_{i_2}, [\ldots, [a_{i_{l-1}}, a_{i_l}] \ldots]]]$. Notice that, as anticipated, only odd powers of h appear in Ω_k and, in particular, $\Omega_{2i+1} = \mathcal{O}(h^{2i+3})$ for $i > 1$.

Let us denote $\alpha_i \equiv h^i a_{i-1}$. Then $[\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_{l-1}}, \alpha_{i_l}]$ is an element of order $h^{i_1 + \cdots + i_l}$. In fact, the matrices $\{\alpha_i\}_{i \geq 1}$ can be considered as the generators (with grade i) of a graded free Lie algebra $\mathcal{L}(\alpha_1, \alpha_2, \ldots)$ [62]. It turns out that it is possible to build methods of order $p \equiv 2s$ by considering only terms involving $\alpha_1, \ldots, \alpha_s$ in Ω. Then, these terms can be approximated

by appropriate linear combinations of the matrix $A(t)$ evaluated at different points. In particular, up to order four, we have to approximate

$$\Omega = \alpha_1 - \frac{1}{12}[\alpha_1, \alpha_2] + \mathcal{O}(h^5), \tag{243}$$

whereas, up to order six, the relevant expression is

$$\Omega = \alpha_1 + \frac{1}{12}\alpha_3 - \frac{1}{12}[\alpha_1, \alpha_2] + \frac{1}{240}[\alpha_2, \alpha_3] + \frac{1}{360}[\alpha_1, \alpha_1, \alpha_3] - \frac{1}{240}[\alpha_2, \alpha_1, \alpha_2] + \frac{1}{720}[\alpha_1, \alpha_1, \alpha_1, \alpha_2] + \mathcal{O}(h^7). \tag{244}$$

In order to present methods which can be easily adapted for different quadrature rules, we introduce the averaged (or generalized momentum) matrices

$$A^{(i)}(h) \equiv \frac{1}{h^i}\int_{t_n}^{t_n+h} \left(t - t_{1/2}\right)^i A(t)\mathrm{d}t = \frac{1}{h^i}\int_{-h/2}^{h/2} t^i A(t + t_{1/2})\mathrm{d}t \tag{245}$$

for $i = 0, \ldots, s-1$. If their exact evaluation is not possible, or is computationally expensive, a numerical quadrature may be used instead. Suppose that b_i, c_i, $(i = 1, \ldots, k)$, are the weights and nodes of a particular quadrature rule, say X (we will use $X = G$ for Gauss–Legendre quadratures and $X = NC$ for Newton–Cotes quadratures) of order p, respectively [45],

$$A^{(0)} = \int_{t_n}^{t_n+h} A(t)\mathrm{d}t = h\sum_{i=1}^{k} b_i A_i + \mathcal{O}(h^{p+1}),$$

with $A_i \equiv A(t_n + c_i h)$. Then it is possible to approximate all the integrals $A^{(i)}$ (up to the required order) by using only the evaluations A_i at the nodes c_i of the quadrature rule required to compute $A^{(0)}$. Specifically,

$$A^{(i)} = h\sum_{j=1}^{k} b_j \left(c_j - \frac{1}{2}\right)^i A_j, \quad i = 0, \ldots, s-1, \tag{246}$$

or equivalently, $A^{(i)} = h\sum_{j=1}^{k}\left(Q_X^{(s,k)}\right)_{ij} A_j$ with $\left(Q_X^{(s,k)}\right)_{ij} = b_j\left(c_j - \frac{1}{2}\right)^i$.

In particular, if fourth and sixth order Gauss–Legendre quadrature rules are considered, then for $s = k = 2$ we have [45]

$$b_1 = b_2 = \frac{1}{2}, \qquad c_1 = \frac{1}{2} - \frac{\sqrt{3}}{6}, \qquad c_2 = \frac{1}{2} + \frac{\sqrt{3}}{6},$$

to order four, whereas for $s = k = 3$,

$$b_1 = b_3 = \frac{5}{18}, \qquad b_2 = \frac{4}{9}, \qquad c_1 = \frac{1}{2} - \frac{\sqrt{15}}{10}, \qquad c_2 = \frac{1}{2}, \qquad c_3 = \frac{1}{2} + \frac{\sqrt{15}}{10},$$

to order six, so that

$$Q_G^{(2,2)} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{\sqrt{3}}{12} & \frac{\sqrt{3}}{12} \end{pmatrix}, \qquad Q_G^{(3,3)} = \begin{pmatrix} \frac{5}{18} & \frac{4}{9} & \frac{5}{18} \\ -\frac{\sqrt{15}}{36} & 0 & \frac{\sqrt{15}}{36} \\ \frac{1}{24} & 0 & \frac{1}{24} \end{pmatrix}. \tag{247}$$

Furthermore, substituting (241) into (245) we find (neglecting higher order terms)

$$A^{(i)} = \sum_{j=1}^{s}\left(T^{(s)}\right)_{ij}\alpha_j \equiv \sum_{j=1}^{s}\frac{1-(-1)^{i+j}}{(i+j)2^{i+j}}\alpha_j, \quad 0 \le i \le s-1. \tag{248}$$

If this relation is inverted (to order four, $s = 2$, and six, $s = 3$) one has

$$R^{(2)} \equiv (T^{(2)})^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 12 \end{pmatrix}, \qquad R^{(3)} = \begin{pmatrix} \frac{9}{4} & 0 & -15 \\ 0 & 12 & 0 \\ -15 & 0 & 180 \end{pmatrix} \tag{249}$$

respectively, so that the corresponding expression of α_i in terms of $A^{(i)}$ or A_j is given by

$$\alpha_i = \sum_{j=1}^{s} \left(R^{(s)}\right)_{ij} A^{(j-1)} = h \sum_{j=1}^{k} \left(R^{(s)} Q_X^{(s,k)}\right)_{ij} A_j. \tag{250}$$

Thus, by virtue of (250) we can write $\Omega(h)$ in terms of the univariate integrals (245) or in terms of any desired quadrature rule. In this way, one gets the final numerical approximations to Ω. Fourth and sixth-order methods can be obtained by substituting in (243) and (244), respectively. The algorithm then provides an approximation for $Y(t_{n+1})$ starting from $Y_n \approx Y(t_n)$, with $t_{n+1} = t_n + h$.

The observant reader surely has noticed that, up to order h^6, there are more terms involved in (242) than those considered in (243) or (244): specifically, $\frac{1}{12}\alpha_3$ in (243) and $\frac{1}{80}\alpha_5$ and $-\frac{1}{80}[\alpha_1, \alpha_4]$ in (244). The reason is that $\Omega^{[6]}$ can be approximated by $A^{(0)}, A^{(1)}, A^{(2)}$ up to order h^6 and then, these omitted terms are automatically reproduced when either $A^{(0)}, A^{(1)}, A^{(2)}$ are evaluated analytically or are approximated by any symmetric quadrature rule of order six or higher.

Another important issue involved in any approximation based on the Magnus expansion is the number of commutators appearing in Ω. As is already evident from (242), this number rapidly increases with the order, and so it might constitute a major factor in the overall computational cost of the resulting numerical methods. It is possible, however, to design an optimization procedure aimed at reducing this number to a minimum [40]. For instance, a straightforward counting of the number of commutators in (244) suggests that it seems necessary to compute seven commutators up to order six in h, whereas the general analysis carried out in [40] shows that this can be done with only three commutators. More specifically, the scheme

$$\begin{aligned} C_1 &= [\alpha_1, \alpha_2], \\ C_2 &= -\frac{1}{60}[\alpha_1, 2\alpha_3 + C_1] \\ \Omega^{[6]} &\equiv \alpha_1 + \frac{1}{12}\alpha_3 + \frac{1}{240}[-20\alpha_1 - \alpha_3 + C_1, \alpha_2 + C_2], \end{aligned} \tag{251}$$

verifies that $\Omega^{[6]} = \Omega + \mathcal{O}(h^7)$. Three is, in fact, the minimum number of commutators required to get a sixth-order approximation to Ω.

This technique to reduce the number of commutators is, indeed, valid for any element in a graded free Lie algebra. It has been used, in particular, to obtain approximations to the Baker–Campbell–Hausdorff formula up to a given order, with the minimum number of commutators [132].

As an illustration, next we provide the relevant expressions for integration schemes of order 4 and 6, which readily follow from the previous analysis.

Order 4. Choosing the Gauss–Legendre quadrature rule, one has to evaluate

$$A_1 = A\left(t_n + \left(\frac{1}{2} - \frac{\sqrt{3}}{6}\right)h\right), \qquad A_2 = A\left(t_n + \left(\frac{1}{2} + \frac{\sqrt{3}}{6}\right)h\right) \tag{252}$$

and thus, taking $Q_G^{(2,2)}$ in (247), $R^{(2)}$ in (249) and substituting in (250) we find

$$\alpha_1 = \frac{h}{2}(A_1 + A_2), \qquad \alpha_2 = \frac{h\sqrt{3}}{12}(A_2 - A_1). \tag{253}$$

Then, by replacing in (243) we obtain

$$\begin{aligned} \Omega^{[4]}(h) &= \frac{h}{2}(A_1 + A_2) - h^2\frac{\sqrt{3}}{12}[A_1, A_2] \\ Y_{n+1} &= \exp(\Omega^{[4]}(h))Y_n. \end{aligned} \tag{254}$$

Alternatively, evaluating A at equispaced points, with $k = 3$ and $c_1 = 0, c_2 = 1/2, c_3 = 1$; $b_1 = b_3 = 1/6, b_2 = 2/3$ (i.e., using the Simpson rule to approximate $\int_{t_n}^{t_n+h} A(s)\mathrm{d}s$),

$$A_1 = A(t_n), \qquad A_2 = A\left(t_n + \frac{h}{2}\right), \qquad A_3 = A(t_n + h)$$

we have instead $\alpha_1 = \frac{h}{6}(A_1 + 4A_2 + A_3)$, $\alpha_2 = h(A_3 - A_1)$, and then

$$\Omega^{[4]}(h) = \frac{h}{6}(A_1 + 4A_2 + A_3) - \frac{h^2}{72}[A_1 + 4A_2 + A_3, A_3 - A_1]. \tag{255}$$

It should be noticed that other possibilities not directly obtainable from the previous analysis are equally valid. For instance, one could consider

$$\Omega^{[4]}(h) = \frac{h}{6}(A_1 + 4A_2 + A_3) - \frac{h^2}{12}[A_1, A_3]. \tag{256}$$

Although apparently more A evaluations are necessary in (255) and (256), this is not actually the case, since A_3 can be reused at the next integration step.

Order 6. In terms of Gauss–Legendre collocation points, one has

$$A_1 = A\left(t_n + \left(\frac{1}{2} - \frac{\sqrt{15}}{10}\right)h\right), \qquad A_2 = A\left(t_n + \frac{1}{2}h\right), \qquad A_3 = A\left(t_n + \left(\frac{1}{2} + \frac{\sqrt{15}}{10}\right)h\right)$$

and similarly we obtain

$$\alpha_1 = hA_2, \qquad \alpha_2 = \frac{\sqrt{15}h}{3}(A_3 - A_1), \qquad \alpha_3 = \frac{10h}{3}(A_3 - 2A_2 + A_1), \tag{257}$$

which are then inserted in (251) to get the approximation $Y_{n+1} = \exp(\Omega^{[6]})Y_n$.

If the matrix $A(t)$ is only known at equispaced points, we can use the Newton–Cotes (NC) quadrature values with $s = 3$ and $k = 5$, $b_1 = b_5 = 7/90$, $b_2 = b_4 = 32/90$, $b_3 = 12/90$ and $c_j = (j-1)/4$, $j = 1, \ldots, 5$. Then, using the corresponding matrix $Q_{NC}^{(3,5)}$ from (246) we get

$$\begin{aligned}
\alpha_1 &= \frac{1}{60}\left(-7(A_1 + A_5) + 28(A_2 + A_4) + 18A_3\right) \\
\alpha_2 &= \frac{1}{15}\left(7(A_5 - A_1) + 16(A_4 - A_2)\right) \\
\alpha_3 &= \frac{1}{3}\left(7(A_1 + A_5) - 4(A_2 + A_4) - 6A_3\right).
\end{aligned} \tag{258}$$

Both schemes involve the minimum number of commutators (three) and require three or four evaluations of the matrix $A(t)$ per integration step (observe that A_5 can be reused in the next step in the Newton–Cotes implementation because $c_1 = 0$ and $c_5 = 1$).

Higher orders can be treated in a similar way. For instance, an 8th-order Magnus method can be obtained with only six commutators [40]. Also variable step size techniques can be easily implemented [39,133].

5.4.1. From Magnus to Fer and Cayley methods

For arbitrary matrix Lie groups, it is feasible to design numerical methods based also on the Fer and Wilcox expansions, whereas for the J-orthogonal group (Eq. (21)) the Cayley transform also maps the Lie algebra onto the Lie group [41] and thus it allows us to build a new class of Lie-group methods. Here we briefly show how these integration methods can be easily constructed from the previous schemes, based on Magnus. In other words, if the solution of (223) in a neighborhood of t_0 is written as

$$Y(t_0 + h) = \mathrm{e}^{\Omega(h)}\, Y_0 \quad \text{(Magnus)} \tag{259}$$

$$= \mathrm{e}^{F_1(h)}\mathrm{e}^{F_2(h)} \cdots Y_0 \quad \text{(Fer)} \tag{260}$$

$$= \mathrm{e}^{S_1(h)}\mathrm{e}^{S_2(h)} \cdots \mathrm{e}^{S_2(h)}\mathrm{e}^{S_1(h)} Y_0 \quad \text{(Symmetric Fer)} \tag{261}$$

$$= \left(I - \frac{1}{2}C(h)\right)^{-1}\left(I + \frac{1}{2}C(h)\right) Y_0 \quad \text{(Cayley)} \tag{262}$$

one may express the functions $F_i(h)$, $S_i(h)$ and $C(h)$ in terms of the successive approximations to Ω and, by using the same techniques as in the previous section, obtain the new methods. As the schemes based on the Wilcox factorization are quite similar as the Fer methods, they will not be considered here.

5.4.2. Fer based methods

To obtain integration methods based on the Fer factorization (260) one applies the Baker–Campbell–Hausdorff (BCH) formula after equating to the Magnus expansion (259). More specifically, in the domain of convergence of expansions (259) and (260) we can write

$$\mathrm{e}^{\Omega(h)} = \mathrm{e}^{F_1(h)}\,\mathrm{e}^{F_2(h)}\,\cdots,$$

where $F_1 = \Omega_1$ is the first term in the Magnus series, $F_2 = O(h^3)$ and $F_3 = O(h^7)$. Then, a p-th order algorithm with $3 \leq p \leq 6$, based on the Fer expansion requires one to compute $F_2^{[p]}$ such that

$$Y(t_n + h) = \mathrm{e}^{F_1(h)}\, \mathrm{e}^{F_2^{[p]}(h)}\, Y(t_n) + \mathcal{O}(h^{p+1}). \tag{263}$$

Taking into account that

$$\mathrm{e}^{\Omega^{[p]}(h)} = \mathrm{e}^{F_1(h)}\, \mathrm{e}^{F_2^{[p]}(h)} + \mathcal{O}(h^{p+1}),$$

we have ($F_1 = \Omega_1$)

$$\mathrm{e}^{F_2^{[p]}(h)} = \mathrm{e}^{-\Omega_1(h)}\, \mathrm{e}^{\Omega^{[p]}(h)} + \mathcal{O}(h^{p+1}).$$

Then, by using the BCH formula and simple algebra to remove higher order terms we obtain to order four

$$F_2^{[4]} = -\frac{1}{12}\left([\alpha_1, \alpha_2] - \frac{1}{2}[\alpha_1, \alpha_1, \alpha_2]\right), \tag{264}$$

so that two commutators are needed in this case. A sixth-order method can be similarly obtained with four commutators [40]. These methods are slightly more expensive than their Magnus counterpart, and they do not preserve the time-symmetry of the exact solution. This can be fixed by the self-adjoint version of the Fer factorization in the form (261) proposed in [134] and presented in a more efficient way in [40]. The schemes based on (261) up to order six are given by

$$Y(t_n + h) = \mathrm{e}^{S_1(h)} \mathrm{e}^{S_2^{[p]}(h)} \mathrm{e}^{S_1(h)} Y_n \tag{265}$$

with $S_1 = \Omega_1/2$. A fourth-order method is given by

$$S_2^{[4]}(h) = -\frac{1}{12}[\alpha_1, \alpha_2] \tag{266}$$

and a sixth-order one by

$$\begin{aligned} s_1 &= [\alpha_1, \alpha_2] \\ r_1 &= \frac{1}{120}\,[\alpha_1, -4\alpha_3 + 3s_1] \\ S_2^{[6]}(h) &= \frac{1}{240}\,[-20\alpha_1 - \alpha_3 + s_1,\ \alpha_2 + r_1]\,. \end{aligned} \tag{267}$$

To complete the formulation of the scheme, the α_i have to be expressed in terms of the matrices A_i evaluated at the quadrature points (e.g., Eqs. (257) or (258)).

5.4.3. Cayley-transform methods

We have seen in Section 1.2 that, for the J-orthogonal group $\mathrm{O}_J(n)$, the Cayley transform (23) provides a useful alternative to the exponential mapping, relating the Lie algebra to the Lie group. This fact is particularly important for numerical methods where the evaluation of the exponential matrix is the most computation-intensive part of the algorithm.

If the solution of Eq. (223) is written as

$$Y(t) = \left(I - \frac{1}{2}C(t)\right)^{-1} \left(I + \frac{1}{2}C(t)\right) Y_0 \tag{268}$$

then $C(t) \in \mathrm{o}_J(n)$ satisfies the equation [135]

$$C' = A - \frac{1}{2}[C, A] - \frac{1}{4}CAC, \quad t \geq t_0, \qquad C(t_0) = 0. \tag{269}$$

Time-symmetric methods of order 4 and 6 have been obtained, based on the Cayley transform (268), by expanding the solution of (269) in a recursive manner and constructing quadrature formulae for the multivariate integrals that appear in the procedure [135–137]. It turns out that efficient Cayley based methods can be built directly from Magnus based integrators [40]. In particular, we get:

Order 4:

$$C^{[4]} = \Omega^{[4]}\left(I - \frac{1}{12}(\Omega^{[4]})^2\right) = \alpha_1 - \frac{1}{12}[\alpha_1, \alpha_2] - \frac{1}{12}\alpha_1^3 + O(h^5), \tag{270}$$

where $C^{[4]} = C(h) + O(h^5)$.

Order 6:

$$C^{[6]} = \Omega^{[6]}\left(I - \frac{1}{12}(\Omega^{[6]})^2\left(I - \frac{1}{10}(\Omega^{[6]})^2\right)\right) = C(h) + O(h^7). \tag{271}$$

Three matrix–matrix products are required in addition to the three commutators involved in the computation of $\Omega^{[6]}$, for a total of nine matrix–matrix products per step.

5.5. Numerical illustration: The Rosen–Zener model revisited

Next we apply the previous numerical schemes to the integration of the differential equation governing the evolution of a particular quantum two-level system. Our purpose here is to illustrate the main issues involved, and compare the different approximations obtained with both the analytical treatment done in Section 4 and the exact result. Specifically, we consider the Rosen–Zener model in the Interaction Picture already analyzed in Section 4.1.2. In this case, the equation to be solved is $U_I' = \tilde{H}_I(t)U_I$, or equivalently, Eq. (223) with $Y(t) = U_I(t)$ and coefficient matrix ($\hbar = 1$)

$$A(t) = \tilde{H}_I(t) = -\mathrm{i}V(s)\,(\sigma_1\cos(\xi s) - \sigma_2\sin(\xi s)) \equiv -\mathrm{i}\,\mathbf{b}(s)\cdot\boldsymbol{\sigma}. \tag{272}$$

Here $V(s) = V_0/\cosh(s)$, $\xi = \omega T$ and $s = t/T$. Given the initial condition $|+\rangle \equiv (1, 0)^{\mathrm{T}}$ at $t = -\infty$, our purpose is to get an approximate value for the transition probability to the state $|-\rangle \equiv (0, 1)^{\mathrm{T}}$ at $t = +\infty$. Its exact expression is collected in the first line of Eq. (204), which we reproduce here for the reader's convenience:

$$P_{ex} = |(U_I)_{12}(+\infty, -\infty)|^2 = \frac{\sin^2\gamma}{\cosh^2(\pi\xi/2)}, \tag{273}$$

with $\gamma = \pi V_0 T$.

To obtain, in practice, a numerical approximation to P_{ex}, we have to integrate the equation in a sufficiently large time interval. We take the initial condition at $s_0 = -25$ and the numerical integration is carried out until $s_f = 25$. Then, we determine $(U_I)_{12}(s_f, s_0)$.

As a first numerical test, we take a fixed (and relatively large) time step h such that the whole numerical integration in the time interval $s \in [s_0, s_f]$ is carried out with 50 evaluations of the vector $\mathbf{b}(s)$ for all methods. In this way, their computational cost is similar.

To illustrate the qualitative behavior of Magnus integrators in comparison with standard Runge–Kutta schemes, the following methods are considered:

- **Explicit first-order Euler** (E1): $Y_{n+1} = Y_n + hA(t_n)Y_n$ with $t_{n+1} = t_n + h$ and $h = 1$ (solid lines with squares in the figures).
- **Explicit fourth-order Runge–Kutta** (RK4), i.e., scheme (227) with $h = 2$, since only two evaluations of $\mathbf{b}(s)$ per step are required in the linear case (solid lines with triangles).
- **Second-order Magnus** (M2): we consider the midpoint rule (one evaluation per step) to approximate Ω_1 taking $h = 1$ (dashed lines), i.e.,

$$Y_{n+1} = \exp\left(-\mathrm{i}h\,\mathbf{b}_n\cdot\boldsymbol{\sigma}\right)Y_n = \left(\cos(hb_n)I - \mathrm{i}\frac{\sin(hb_n)}{hb_n}\,\mathbf{b}_n\cdot\boldsymbol{\sigma}\right)Y_n \tag{274}$$

  with $\mathbf{b}_n \equiv \mathbf{b}(t_n + h/2)$ and $b_n = \|\mathbf{b}_n\|$. The trapezoidal rule is equally valid by considering $\mathbf{b}_n \equiv (\mathbf{b}(t_n) + \mathbf{b}(t_n + h))/2$.
- **Fourth-order Magnus** (M4). Using the fourth-order Gauss–Legendre rule to approximate the integrals and taking $h = 2$, one has the scheme (254) which, for this problem, reads

$$\begin{aligned}
&\mathbf{b}_1 = \mathbf{b}(t_n + c_1 h), \qquad \mathbf{b}_2 = \mathbf{b}(t_n + c_2 h),\\
&\mathbf{d} = \frac{h}{2}(\mathbf{b}_1 + \mathbf{b}_2) - h^2\frac{\sqrt{3}}{6}\mathrm{i}(\mathbf{b}_1\times\mathbf{b}_2)\\
&Y_{n+1} = \exp\left(-\mathrm{i}h\,\mathbf{d}\cdot\boldsymbol{\sigma}\right)Y_n
\end{aligned} \tag{275}$$

  with $c_1 = \frac{1}{2} - \frac{\sqrt{3}}{6}$, $c_2 = \frac{1}{2} + \frac{\sqrt{3}}{6}$ (dotted lines).

We choose $\xi = 0.3$ and $\xi = 1$, and each numerical integration is carried out for different values of γ in the range $\gamma \in [0, 2\pi]$. We plot the corresponding approximations to the transition probability in a similar way as in Fig. 7 for the analytical treatment. Thus we obtain the plots of Fig. 14. As expected, the performance of the methods deteriorates as γ increases. Notice, also, that the qualitative behavior of the different numerical schemes is quite similar as that exhibited by the analytical approximations. Euler and Runge–Kutta methods do not preserve unitarity and may lead to transition probabilities greater than 1 (just like the standard perturbation theory). On the other hand, for sufficiently small values of γ (inside the convergence domain of the Magnus series) the fourth-order Magnus method improves the result achieved by the second-order, whereas for large values of γ, a higher order method does not necessarily lead to a better approximation.

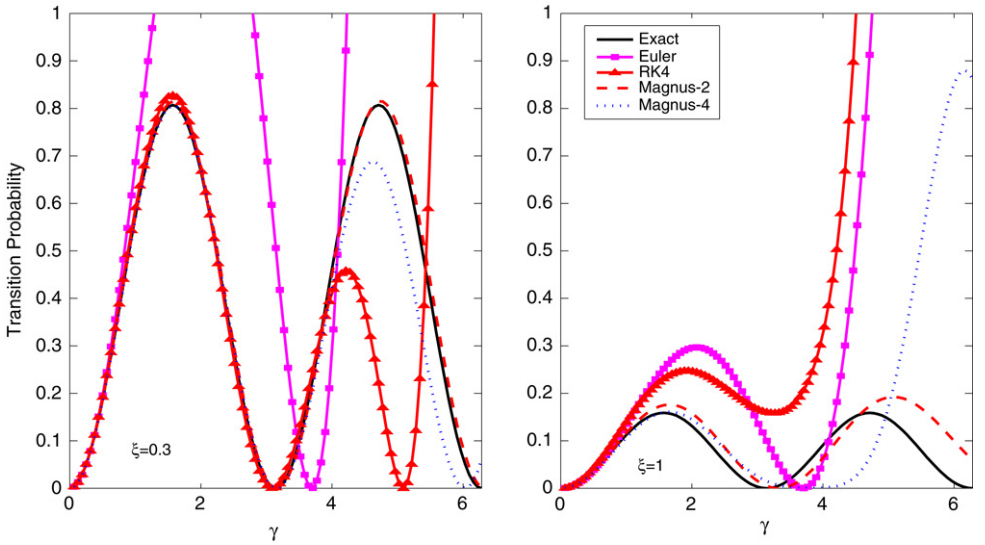


Fig. 14. Rosen–Zener model: Transition probabilities as a function of γ, with $\xi = 0.3$ and $\xi = 1$. The curves are coded as follows. Solid line represents the exact result; E1: solid lines with squares; RK4: solid lines with triangles; M2: dashed lines; M4: dashed-broken lines (indistinguishable from exact result in left panel).

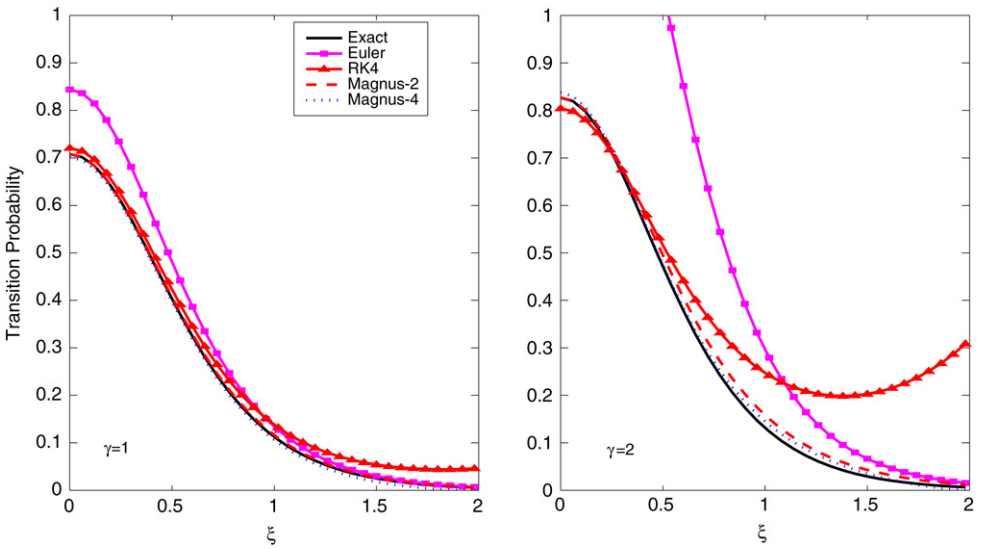


Fig. 15. Rosen–Zener model: Transition probabilities as a function of ξ, with $\gamma = 2$ and $\gamma = 5$. Lines are coded as in Fig. 14.

In Fig. 15 we repeat the experiment, now taking $\gamma = 1$ and $\gamma = 2$, and for different values of ξ. In the first case, only the Euler method differs considerably from the exact solution, and in the second case this happens for both RK methods.

To increase the accuracy, one can always take smaller time steps, but then the number of evaluations of $A(t)$ increases, and this may be computationally very expensive for some problems due to the complexity of the time dependence and/or the size of the matrix. In those cases, it is important to have reliable numerical methods, providing the desired accuracy as fast as possible or, alternatively, leading to the best possible accuracy at a given computational cost.

A good perspective of the overall performance of a given numerical integrator is provided by the so-called efficiency diagram. This efficiency plot is obtained by carrying out the numerical integration with different time steps, corresponding to different numbers of evaluations of $A(t)$. For each run, one compares the corresponding approximation with the exact solution, and plots the error as a function of the total number of matrix evaluations. The results are better illustrated in a double logarithmic scale. In that case, the slope of the curves should correspond, in the limit of very small time steps, to (minus) the order of accuracy of the method.

To illustrate this issue, in Fig. 16 we collect the efficiency plots of the previous schemes when $\xi = 0.3$ with $\gamma = 10$ (left) and $\gamma = 100$ (right). We have also included the results obtained by several higher order integrators, namely the sixth-order RK method (RK6) whose coefficients are collected in (229) and the sixth-order Magnus integrator (M6) given

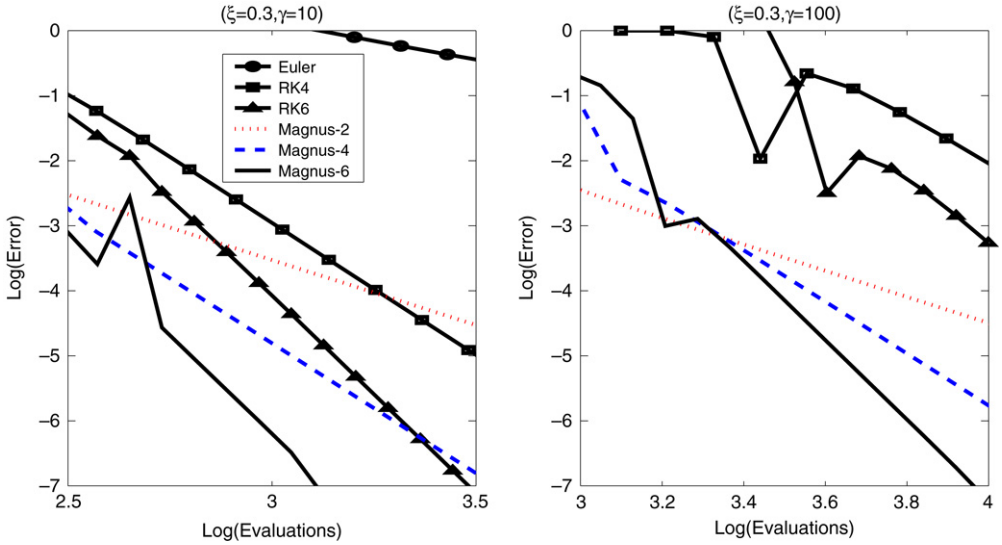


Fig. 16. Rosen–Zener model: Error in the transition probabilities versus the number of evaluations of the Hamiltonian $H_I(s)$ for $\xi = 0.3$ and $\gamma = 10$ (left panel) and $\gamma = 100$ (right panel).

by (251) and (257). We clearly observe that the Euler method is, by far, the worst choice if accurate results are desired. Notice the (double) logarithmic scale of the plots: thus, for instance, when $\gamma = 10$ the range goes approximately from 300 to 3000 evaluations of $A(t)$. Magnus integrators, in addition to providing results in SU(2) by construction (as the exact solution), show a better efficiency than Runge–Kutta schemes for these examples, and this efficiency increases with the value of γ considered. The implicit Runge–Kutta–Gauss–Legendre methods (237) show slightly better results than the explicit RK methods, but still considerably worse than Magnus integrators.

If we compare Magnus integrators of different orders of accuracy, we observe that the most efficient scheme is the second order method M2 when relatively low accuracies are desired. For higher accuracy, however, it is necessary to carry out a thorough analysis of the computational cost of the methods for a given problem before asserting the convenience of M4 or M6 with respect to higher order schemes. For a fixed time step h, the computational cost of a certain family of methods (such as those based on Magnus) usually increases with the order. However, if one fixes the number of $A(t)$ evaluations, this is not necessarily the case (sometimes higher order methods require more commutators but less exponentials).

Let us now compare the performance of the Magnus methods with respect to other Lie-group solvers, namely Fer and Cayley methods. We repeat the same experiments as in Fig. 16 but, for clarity, only the results for the 6th-order methods are shown. We consider the symmetric-Fer method given by (265) and (267) and the Cayley method (268) with (271), using in both cases the Gauss–Legendre quadrature. The results obtained are collected in Fig. 17. We clearly observe that the relative performance of the Cayley method deteriorates by increasing the value of γ similarly to the RK6. In spite of preserving the qualitative properties, this example shows that for some problems, polynomial or rational approximations do not perform efficiently. Here, in particular, the Magnus scheme is slightly more efficient than the symmetric Fer method, although for other problems their performance is quite similar.

5.6. Computing the exponential of a matrix

We have seen that the numerical schemes based on the Magnus expansion provide excellent results when applied to Eq. (223) with coefficient matrix (272). In fact, they are even more efficient than several Runge–Kutta algorithms. Of course, for this particular example, the number of $A(t)$ evaluations is a good indication of the computational cost required by the numerical schemes, since the evaluation of $\exp(\Omega)$ can be done analytically by means of formula (18). In general, however, the matrix exponential also has to be approximated numerically, and thus the performance of the numerical integration algorithms based on the Magnus expansion strongly depends on this fact. It may be that the evaluation of $\exp(C)$, where C is a (real or complex) $N \times N$ matrix, represents the major factor in the overall computational cost required by this class of algorithms and is probably one of the most problematic aspects.

As a matter of fact, the approximation of the matrix exponential is among the oldest and most extensively researched topics in numerical mathematics. Although many efficient algorithms have been developed, the problem is still far from being solved in general. It seems, then, reasonable to briefly summarize here, some of the most widely used procedures in this context.

Let us begin with two obvious but important remarks. (i) First, one has to distinguish whether it is necessary to evaluate the full matrix $\exp(C)$ or only the product $\exp(C)v$ for some given vector v. In the latter case, special algorithms can be designed, requiring a much reduced computational effort. This is especially true when C is large and sparse

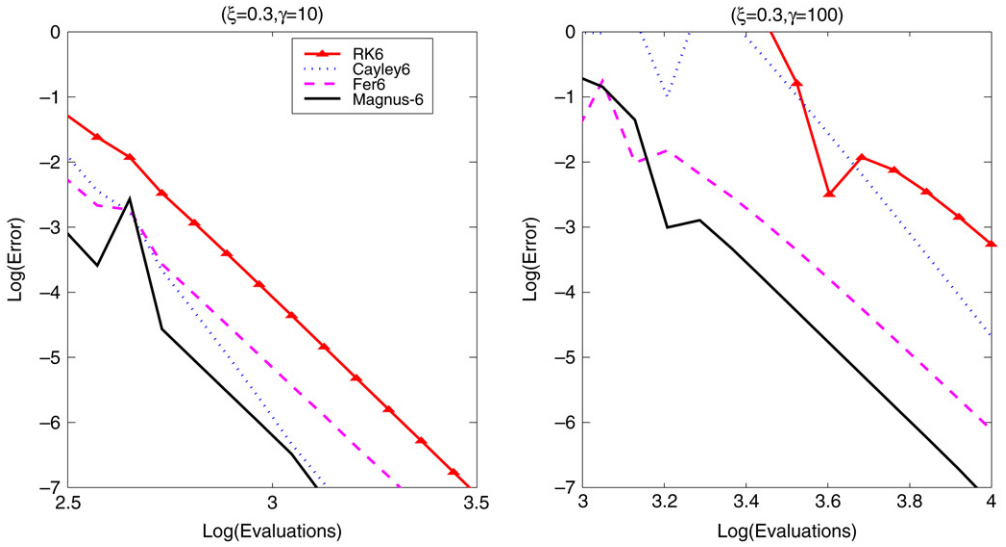


Fig. 17. Rosen–Zener model: Same as Fig. 16 where we compare the performance of the 6th-order Magnus, symmetric-Fer, Cayley and RK6.

(as often happens with matrices arising from the spatial discretization of partial differential equations). (ii) Second, for the numerical integration methods based on ME, one has to compute $\exp(C(h))$, where $C(h) = \mathcal{O}(h^p)$, h is a (not too large) step size and $p \geq 1$. In other words, the matrices to be exponentiated typically have a small norm (usually restricted by the convergence bounds of the expansion).

In any case, prior to the computation of $\exp(C)$, it is important to have as much information about the exponent C as possible. Thus, for instance, if the matrix C resides in a Lie algebra, then $\exp(C)$ belongs to the corresponding Lie group and one has to decide whether this qualitative property has to be exactly preserved, or if constructing a sufficiently accurate approximation (e.g., at a higher order than the order of the integrator itself) is enough. Also, when C can be split into different parts, one may consider a factorization of the form $\exp(C) \approx \exp(C_1)\exp(C_2)\cdots\exp(C_m)$ if each individual exponential is easy to evaluate exactly.

An important reference in this context is [138] and its updated version [139], where up to nineteen (or twenty in [139]) different numerical algorithms for computing the exponential of a matrix are carefully reviewed. An extensive software package for computing the matrix exponential is Expokit, developed by R. Sidje, with Fortran and Matlab versions available [140,141]. In addition to computing the matrix-valued function $\exp(C)$ for small, dense matrices C, Expokit has functions for computing the vector-valued function $\exp(C)v$ for both small, dense matrices and large, sparse matrices.

5.6.1. *Scaling and squaring with Padé approximation*

Among one of the least dubious ways of computing $\exp(C)$ is by scaling and squaring in combination with a diagonal Padé approximation [139]. The procedure is based on a fundamental property of the exponential function, namely

$$\mathrm{e}^C = (\mathrm{e}^{C/j})^j$$

for any integer j. The idea, then, is to choose j to be a power of two for which $\exp(C/j)$ can be reliably and efficiently computed, and then to form the matrix $(\exp(C/j))^j$ by repeated squaring. If the integer j is chosen as the smallest power of two for which $\|C\|/j < 1$, then $\exp(C/j)$ can be satisfactorily computed by diagonal Padé approximants of order, say, m. This is roughly the method used by the built-in function `expm` in Matlab.

For the integrators based on the Magnus expansion, as $C = \mathcal{O}(h^p)$ with $p \geq 1$, one usually gets good approximations with relatively small values of j and m.

As is well known, diagonal Padé approximants map the Lie algebra $\mathrm{o}_J(n)$ to the J-orthogonal Lie group $\mathrm{O}_J(n)$ and thus also constitute a valid alternative to the evaluation of the exponential matrix in Magnus-based methods for this particular Lie group. More specifically, if $B \in \mathrm{o}_J(n)$, then $\psi_{2m}(tB) \in \mathrm{O}_J(n)$ for sufficiently small $t \in \mathbb{R}$, with

$$\psi_{2m}(\lambda) = \frac{P_m(\lambda)}{P_m(-\lambda)}, \tag{276}$$

provided the polynomials P_m are generated according to the recurrence

$$P_0(\lambda) = 1, \qquad P_1(\lambda) = 2 + \lambda$$
$$P_m(\lambda) = 2(2m-1)P_{m-1}(\lambda) + \lambda^2 P_{m-2}(\lambda).$$

Moreover, $\psi_{2m}(\lambda) = \exp(\lambda) + \mathcal{O}(\lambda^{2m+1})$ and ψ_2 corresponds to the Cayley transform (262), whereas for $m = 2, 3$ we have

$$\psi_4(\lambda) = \left(1 + \frac{1}{2}\lambda + \frac{1}{12}\lambda^2\right) \Big/ \left(1 - \frac{1}{2}\lambda + \frac{1}{12}\lambda^2\right)$$

$$\psi_6(\lambda) = \left(1 + \frac{1}{2}\lambda + \frac{1}{10}\lambda^2 + \frac{1}{120}\lambda^3\right) \Big/ \left(1 - \frac{1}{2}\lambda + \frac{1}{10}\lambda^2 - \frac{1}{120}\lambda^3\right).$$

Thus, we can combine the optimized approximations to Ω obtained in Section 5.4 for Magnus based methods with diagonal Padé approximants, up to the corresponding order, to obtain time-symmetric integration schemes preserving the algebraic structure of the problem without computing the matrix exponential. For instance, the "Magnus–Padé" methods thus obtained involve, in addition to one matrix inversion, 3 and 8 matrix–matrix products for order 4 and 6, respectively.

Observe that since $\Omega^{[2n]} = \mathcal{O}(h)$ then

$$\psi_{2m}(\Omega^{[2n]}) = \exp(\Omega^{[2n]}) + \mathcal{O}(h^{2k+1}),$$

where $k = \min\{m, n\}$. With $m = n$ we have a method of order $2n$. However, for some problems, this rational approximation to the exponential may be not very accurate depending on the eigenvalues of $\Omega^{[2n]}$. In this case, one may take $m > n$, thus giving a better approximation to the exponential and a more accurate result by slightly increasing the computational cost of the method. Of course, when the norm of the matrix $\Omega^{[2n]}$ is not so small, this technique can be combined with scaling and squaring [142].

Instead of using Padé approximants for the exponential of the scaled matrix $B \equiv C/2^k$, Najfeld and Havel [143] propose a rational approximation for the matrix function

$$H(B) = B\coth(B) = B\,\frac{\mathrm{e}^{2B} + I}{\mathrm{e}^{2B} - I}, \tag{277}$$

from which the exponential can be obtained as

$$\mathrm{e}^{2B} = \frac{H(B) + B}{H(B) - B}$$

and then iteratively square the result k times to recover the exponential of the original matrix C. From the continued fraction expansion of $H(B)$, it is possible to compute the first rational approximations as

$$H_2(B) = \frac{I + \frac{2}{5}B^2}{I + \frac{1}{15}B^2}, \qquad H_4(B) = \frac{I + \frac{4}{9}B^2 + \frac{1}{63}B^4}{I + \frac{1}{9}B^2 + \frac{1}{945}B^4}$$

and so on. Observe that the representation (277) can be regarded as a generalized Cayley transform of B and thus it also provides approximations in the group $\mathrm{O}_J(n)$. In [143] the authors report a saving of about 30% in the number of matrix multiplications with respect to diagonal Padé approximants when an optimal k and a rational approximation for $H(B)$ is used.

5.6.2. Chebyshev approximation

Another valid alternative is to use polynomial approximations to the exponential of C as a whole. Suppose, in particular, that C is a matrix of the form $C = -\mathrm{i}\tau H$, with H Hermitian and $\tau > 0$, as is the case in Quantum Mechanics. In the Chebyshev approach, the evolution operator $\exp(-\mathrm{i}\tau H)$ is expanded in a truncated series of Chebyshev polynomials, in analogy with the approximation of a scalar function [144]. As is well known, given a function $F(x)$ in the interval $[-1, 1]$, the Chebyshev polynomial approximations are optimal, in the sense that the maximum error in the approximation is minimal compared to almost all possible polynomial approximations [145]. To apply this procedure, one has to previously bound the extreme eigenvalues $E_{\min}$ and $E_{\max}$ of H. Then a truncated Chebyshev expansion of $\exp(-\mathrm{i}x)$ on the interval $[\tau E_{\min}, \tau E_{\max}]$ is considered:

$$\exp(-\mathrm{i}x) \approx \sum_{n=0}^{m} c_n P_n(x),$$

where

$$P_n(x) = T_n\left(\frac{2x - \tau E_{\max} - \tau E_{\min}}{\tau E_{\max} - \tau E_{\min}}\right)$$

with appropriately chosen coefficients c_n. Here $T_n(x)$ are the Chebyshev polynomials on the interval $[-1, 1]$ [45], which can be determined via the recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x); \qquad T_1(x) = x; \qquad T_0(x) = 1.$$

Finally, one uses the approximation

$$\exp(-\mathrm{i}\tau H) \approx \sum_{n=0}^{m} c_n P_n(\tau H). \tag{278}$$

This technique is frequently used in numerical quantum dynamics to compute $\exp(-\mathrm{i}\tau H)\psi_0$ over very long times. This can be done with m matrix-vector products if the approximation (278) is considered with a sufficiently large truncation index m. In fact, the degree m necessary for achieving a specific accuracy depends linearly on the step size τ and the spectral radius of H [146], and thus an increase of the step size reduces the computational work per unit step. In a practical implementation, m can be chosen such that the accuracy is dominated by the round-off error [147]. This approach has two main drawbacks: (i) it is not unitary, and therefore the norm is not conserved (although the deviation from unitarity is really small due to its extreme accuracy), and (ii) intermediate results are not obtained, since, typically, τ is very large.

5.6.3. Krylov space methods

As we have already pointed out, very often what is really required, rather than the exponential of the matrix C itself, is the computation of $\exp(C)$ applied to a vector. In this situation, evaluating e^C is somehow analogous to computing C^{-1} to get the solution of the linear system of equations $Cx = b$ for many different b's: other procedures are clearly far more desirable. The computation of $\mathrm{e}^C v$ can be efficiently done with Krylov subspace methods, in which approximations to the solution are obtained from the Krylov spaces spanned by the vectors $\{v, Cv, C^2v, \ldots, C^jv\}$ for some j that is typically small compared to the dimension of C [148,149]. The Lanczos method for solving iteratively symmetric eigenvalue problems is of this form [150]. If, as before, we let $C = -\mathrm{i}\tau H$, the symmetric Lanczos process recursively generates an orthonormal basis $V_m = [v_1 \cdots v_m]$ of the mth Krylov subspace $K_m(H, u) = \mathrm{span}(u, Hu, \ldots, H^{m-1}u)$ such that

$$HV_m = V_m L_m + [0 \cdots 0\, \beta_m\, v_{m+1}],$$

where the symmetric tridiagonal $m \times m$ matrix $L_m = V_m^{\mathrm{T}} H V_m$ is the orthogonal projection of H onto $K_m(H, u)$. Finally,

$$\exp(-\mathrm{i}\tau H)u \approx V_m \exp(-\mathrm{i}\tau L_m) V_m^{\mathrm{T}} u$$

and the matrix exponential $\exp(-\mathrm{i}\tau L_m)$ can be computed by diagonalizing L_m, $L_m = Q_m D_m Q_m^{\mathrm{T}}$, as

$$\exp(-\mathrm{i}\tau L_m) = Q_m \exp(-\mathrm{i}\tau D_m) Q_m^{\mathrm{T}},$$

with D_m a diagonal matrix. This iterative process is stopped when

$$\beta_m \left\| (\exp(-\mathrm{i}\tau L_m))_{m,m} \right\| < tol$$

for a fixed tolerance. Very good approximations are often obtained with relatively small values of m, and computable error bounds exist for the approximation. This class of schemes generally require $\mathcal{O}(N^2)$ floating point operations in the computation of $\mathrm{e}^C v$. More details are contained in the references [151–153].

5.6.4. Splitting methods

Frequently, one has to exponentiate a matrix which can be split into several parts which are either solvable or easy to deal with. Let us assume for simplicity that $C = A + B$, where the computation e^C is very expensive, but e^A and e^B are cheap and easy to evaluate. In such circumstances, it makes sense to approximate $\mathrm{e}^{\varepsilon C}$ with ε, a small parameter, by the following scheme:

$$\psi_h^{[p]} \equiv \mathrm{e}^{\varepsilon b_m B} \mathrm{e}^{\varepsilon a_m A} \cdots \mathrm{e}^{\varepsilon b_1 B} \mathrm{e}^{\varepsilon a_1 A} = \mathrm{e}^{\varepsilon(A+B)} + \mathcal{O}(\varepsilon^{p+1}) \tag{279}$$

with appropriate parameters a_i, b_i. This can be seen as the approximations to the solution at $t = \varepsilon$ of the equation $Y' = (A + B)Y$ by a composition of the exact solutions of the equations $Y' = AY$ and $Y' = BY$ at times $t = a_i\varepsilon$ and $t = b_i\varepsilon$, respectively. Two instances of this kind of approximation are given by the well known Lie–Trotter formula

$$\psi_h^{[1]} = \mathrm{e}^{\varepsilon A} \mathrm{e}^{\varepsilon B} \tag{280}$$

and the second order symmetric composition

$$\psi_h^{[2]} = \mathrm{e}^{\varepsilon A/2} \mathrm{e}^{\varepsilon B} \mathrm{e}^{\varepsilon A/2}, \tag{281}$$

referred to as Strang splitting, Störmer, Verlet and leap-frog, depending on the particular area where it is used.

Splitting methods have been considered in different contexts: in designing symplectic integrators, for constructing volume-preserving algorithms, in the numerical integration of partial differential equations etc. An extensive survey of the theory and practice of splitting methods can be found in [154,79,127–129] and references therein.

Splitting methods are particularly useful in geometric numerical integration. Suppose that the matrix $C = A + B$ resides in a Lie algebra $\mathfrak{g}$. Then, obviously, $\exp(C)$ belongs to the corresponding Lie group $\mathcal{G}$ and one is naturally interested in also getting approximations in $\mathcal{G}$. In this respect, notice that if $A, B \in \mathfrak{g}$, then the scheme (279) also provides an approximation in $\mathcal{G}$. It is worth noticing that other methods for the approximation of the matrix exponential (e.g., Padé approximants and Krylov subspace techniques) are not guaranteed to map elements from $\mathfrak{g}$ to $\mathcal{G}$. Although diagonal Padé approximants map the Lie algebra $\mathfrak{o}_J$ to the underlying group, it is possible to show that the only analytic function that maps $\mathfrak{sl}(n)$ into the special linear group $\mathrm{SL}(n)$ approximating the exponential function up to a given order, is the exponential itself [155]. In consequence, diagonal Padé approximants only provide results in $\mathrm{SL}(n)$ if the computation is accurate to machine precision.

In [156], Celledoni and Iserles devised a splitting technique for obtaining an approximation to $\exp(C)$ in the Lie group $\mathcal{G}$ based on a decomposition of $C \in \mathfrak{g}$ into low-rank matrices $C_i \in \mathfrak{g}$. Basically, given a $n \times n$ matrix $C \in \mathfrak{g}$, they proposed to split it in the form

$$C = \sum_{i=1}^{k} C_i,$$

such that

(1) $C_i \in \mathfrak{g}$, for $i = 1, \ldots, k$.
(2) Each $\exp(C_i)$ is easy to evaluate exactly.
(3) Products of such exponentials are computationally cheap.

For instance, for the Lie algebra $\mathfrak{so}(n)$, the choice

$$C_i = \frac{1}{2} c_i e_i^{\mathrm{T}} - \frac{1}{2} e_i b_i^{\mathrm{T}}, \quad i = 1, \ldots, n,$$

where $c_1, \ldots, c_n$ are the columns of C and e_i is the i-th vector of the canonical basis of $\mathbb{R}^n$, satisfies the above requirements (with $k = n$), whereas in the case of $\mathfrak{g} = \mathfrak{sl}(n)$ other (more involved) alternatives are possible [156].

Proceeding as in (280), with

$$\psi^{[1]} = \exp(\varepsilon C_1) \exp(\varepsilon C_2) \cdots \exp(\varepsilon C_k)$$

we get an order one approximation in ε to $\exp(\varepsilon C)$, whereas the symmetric composition

$$\psi^{[2]} = \mathrm{e}^{\frac{1}{2}\varepsilon C_k} \mathrm{e}^{\frac{1}{2}\varepsilon C_{k-1}} \cdots \mathrm{e}^{\frac{1}{2}\varepsilon C_2} \mathrm{e}^{\varepsilon C_1} \mathrm{e}^{\frac{1}{2}\varepsilon C_2} \cdots \mathrm{e}^{\frac{1}{2}\varepsilon C_{k-1}} \mathrm{e}^{\frac{1}{2}\varepsilon C_k} \tag{282}$$

provides an approximation of order two in ε, and this can be subsequently combined with different techniques for increasing the order.

With respect to the computational cost, the results reported in [156] show that, up to order four in ε, this class of splitting schemes is competitive with the Matlab built-in function `expm` when machine accuracy is not required in the final approximation. Running `expm` on randomly generated matrices, it is possible to verify that computing $\exp(C)$ to machine accuracy requires about $(20–30)n^3$ floating point operations, depending on the eigenvalues of C, whereas the 4th-order method constructed from (282) involves $(12–15)n^3$ operations when $C \in \mathfrak{so}(100)$ [156]. In the case of the approximation of $\exp(C)v$ and $v \in \mathbb{R}^n$, the cost of low-rank splitting methods drop down to Kn^2, where K is a constant, and thus they are comparable to those achieved by polynomial approximants [153].

Splitting methods of the above type are by no means the only way to express $\exp(C)$ as a product of exponentials of elements in $\mathfrak{g}$. For instance, the Wei–Norman approach (111) can also be implemented in this setting. Suppose that $\dim \mathfrak{g} = s$ and let $\{X_1, X_2, \ldots, X_s\}$ be a basis of $\mathfrak{g}$. In that case, as we have seen (Section 2.10.3), it is possible to represent $\exp(tC)$ for $C \in \mathfrak{g}$ and sufficiently small $|t|$ in canonical coordinates of the second kind,

$$\mathrm{e}^{tC} = \mathrm{e}^{g_1(t)X_1} \mathrm{e}^{g_2(t)X_2} \cdots \mathrm{e}^{g_s(t)X_s},$$

where the scalar functions g_k are analytic at $t = 0$. Although the g_ks are implicitly defined, they can be approximated by Taylor series. The cost of the procedure can be greatly reduced by adequately choosing the basis and exploiting the Lie-algebraic structure [157].

Yet another procedure to get approximations of $\exp(C)$ in a Lie-algebraic setting, which has received considerable attention during the last years, is based on generalized polar decompositions (GPD), an approach introduced in [158] and further elaborated in [159,160]. In particular, in [159], by bringing together GPD with techniques from numerical linear algebra, new algorithms are presented with complexity $\mathcal{O}(n^3)$, both when the exponential is applied to a vector and to a matrix. This is certainly not competitive with Krylov subspace methods in the first case, but represents at least a 50% improvement on the execution time, depending on the Lie algebra considered, in the latter. Another difference with respect to Krylov methods is that the algorithms based on generalized polar decompositions approximate the exponential to a given order of accuracy and, thus, they are well suited to exponential approximations within numerical integrators for ODEs, since the error is subsumed in that of the integration method. For a complete description of the procedure and its practical implementation, we refer the reader to [159].

5.7. Additional numerical examples

The purpose of Section 5.5 was to illustrate the main features of the numerical schemes based on the Magnus expansion in comparison with other standard integrators (such as Runge–Kutta schemes) and other Lie-group methods (e.g., Fer and Cayley) on a solvable system. For larger systems the efficiency analysis is more challenging since the (exact or approximate) computation of exponential matrices play an important role on the performance of the methods. It makes sense, then, to analyze from this point of view, more realistic problems, where one necessarily has to approximate the exponential in a consistent way.

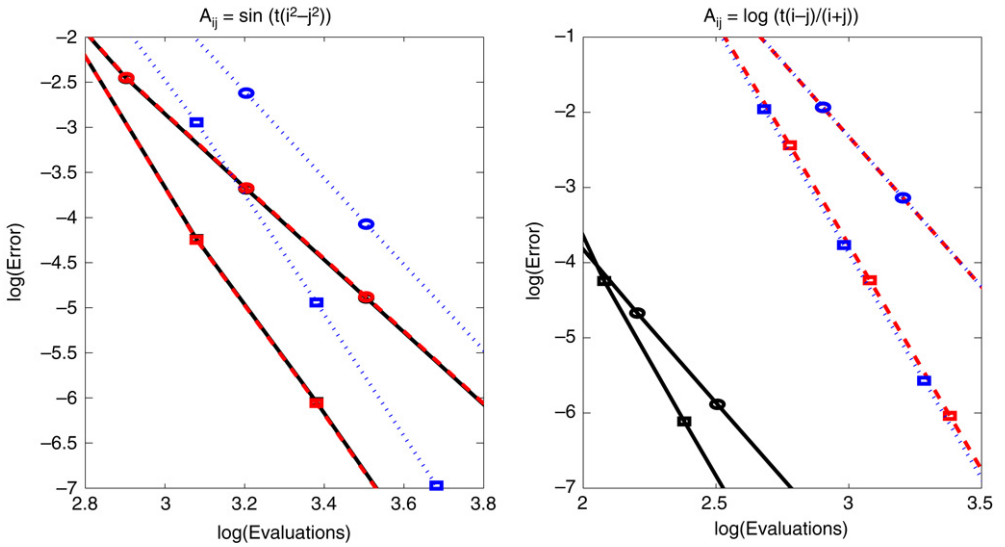


Fig. 18. Efficiency diagram corresponding to the optimized 4-th (circles) and 6-th (squares) order Lie-group solvers based on Magnus (solid lines) and Cayley (broken lines), and the standard Runge–Kutta methods (dashed lines).

As an illustration of this situation, we next consider two skew-symmetric matrices, $A(t)$ and $Y(0) = I$, so that the solution $Y(t)$ of $Y' = A(t)Y$ is orthogonal for all t. In particular, the upper triangular elements of the matrices $A(t)$ are as follows:

$$\text{(a)} \quad A_{ij} = \sin\left(t(j^2 - i^2)\right) \quad 1 \le i < j \le N \tag{283}$$

$$\text{(b)} \quad A_{ij} = \log\left(1 + t\,\frac{j-i}{j+i}\right) \tag{284}$$

with $N = 10$. In both cases $Y(t)$ oscillates with time, mainly due to the time-dependence of $A(t)$ (in (283)) or the norm of the eigenvalues (in (284)).

The integration is carried out in the interval $t \in [0, 10]$ and the approximate solutions are compared with the exact one at the final time $t_f = 10$ (obtained with very high accuracy by using a sufficiently small time step). The corresponding error at t_f is computed for different values of the time step h. The Lie-group solvers are implemented with Gauss–Legendre quadratures and constant step size.

First, we plot the accuracy of the different 4-th and 6-th order methods as a function of the number of $A(t)$ evaluations. In contrast to the previous examples, now there is no closed formula for the matrix exponentials appearing in the Magnus based integrators, so that some alternative procedure must be applied. In particular, the computation of e^C to machine accuracy is done by scaling–(diagonal Padé)–squaring, so that the result is correct up to round-off. Fig. 18 shows the results obtained for the problems (283) and (284) with fourth- and sixth-order numerical schemes based on Magnus and Cayley, and also explicit Runge–Kutta methods. In the first problem, Magnus and Cayley show a very similar performance, which happens to be only slightly better than that of RK methods.

The situation changes drastically, however, for the second problem. Here the behavior of Cayley and RK methods is essentially similar, whereas schemes based on Magnus are clearly more efficient. The reason seems to be that Cayley and RK methods give poor approximations to the exponential, which, on the other hand, has to be accurately approximated, since the eigenvalues of $A(t)$ may take large values.

With respect to symmetric Fer methods, their efficiency is quite similar to that of Magnus if the matrix exponentials are evaluated accurately up to machine precision. This is so for the matrix (283) even if Padé approximants of relatively low order are used to replace the exponentials.

On the other hand, the efficiency of "Magnus–Padé" methods (we denote by MPnm a Magnus method of order n where the exponential is approximated by a diagonal Padé of order m, and MPn if $n = m$) is highly deteriorated for the problem (284), although it is always better than that corresponding to Cayley schemes.

To better illustrate all these comments, in Fig. 19 we display the error in the solution corresponding to (283) and (284) as a function of time in the interval $t \in [0, 100]$ for $h = 1/20$, as is obtained by the previous methods. We should stress that all schemes require the same number of A evaluations.

In the right picture, the exponentials appearing in the Magnus method are computed using a Padé approximant of order six (MP6), of order eight (MP68) and to machine accuracy (M6). Observe the great importance of evaluating the exponential as accurately as possible for the matrix (284): by slightly increasing the computational cost per step in the computation of the matrix exponential, it is possible to dramatically improve the accuracy of the methods. On the contrary, for problem (283) the meaningful fact seems to be that the integration scheme provides a solution in the corresponding Lie group.

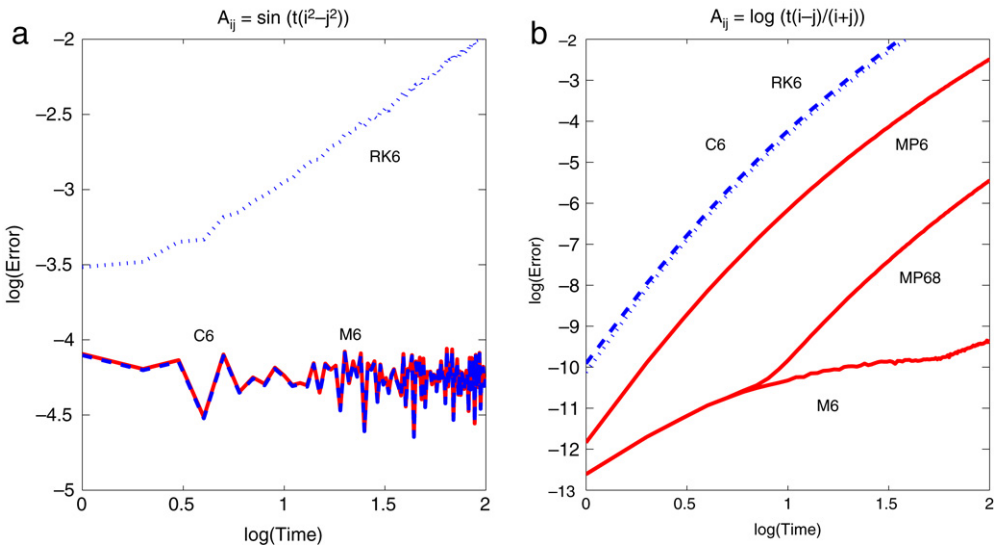


Fig. 19. Error as a function of time (in logarithmic scale) obtained with different 6-th order integrators for $h = 1/20$: (a) problem (283); (b) problem (284).

5.8. Modified Magnus integrators

5.8.1. Variations on the basic algorithm

The examples collected in Sections 5.5 and 5.7 show that the numerical methods constructed from the Magnus expansion can be, computationally, very efficient indeed. It is fair to say, however, that this efficiency can be seriously affected when dealing with certain types of problems.

Suppose, for instance, that one has to numerically integrate a problem defined in SU(n). Although Magnus integrators are unconditionally stable in this setting (since they preserve unitarity up to round-off, independently of the step size h), in practice only small values of h are used for achieving accurate results. Otherwise the convergence of the series is not assured. Of course, the use of small time steps may render the algorithm exceedingly costly.

In other applications, the problem depends on several parameters, so that the integration has to be carried out for different values of the parameters. In that case, the overall integration procedure can be computationally very expensive.

In view of all these difficulties, it is hardly surprising that several modifications of the standard algorithm of Magnus integrators had been developed to try to minimize these undesirable effects and get specially adapted integrators for particular problems.

One basic tool, used time and again in this context, is performing a preliminary linear transformation, similarly to those introduced in Section 2.9. These transformations can be carried out either for the whole integration interval or at each step in the process. Given an appropriately chosen transformation, $\tilde{Y}_0(t)$, one factorizes $Y(t) = \tilde{Y}_0(t)\tilde{Y}_1(t)$, where the unknown $\tilde{Y}_1(t)$ satisfies the equation

$$\tilde{Y}_1' = B(t)\tilde{Y}_1 \tag{285}$$

and $B(t)$ depends on $A(t)$ and $\tilde{Y}_0(t)$. This transformation makes sense, of course, if $\|B(t)\| < \|A(t)\|$ and thus typically $\tilde{Y}_0$ is chosen in such a way that the norm of B verifies the above inequality. As a consequence, Magnus integrators can be applied on (285) with larger time steps also providing more accurate results.

Alternatively, for problems where, in addition to the time step h there is another parameter (E, say), one may analyze the Magnus expansion in terms of h and E. This allows us to identify which terms at each order in the series expansion give the main contribution to the error, and to design methods which include these terms in their formulation. The resulting schemes should then provide more accurate results at a moderate computational cost without altering the convergence domain. As a general rule, it is always desirable to have, in advance, as much information about the equation and the properties of its solution as possible, and then to try to incorporate all this information into the algorithm.

Let us review some useful possibilities in this context. From (241) one has

$$A(t) = a_0 + \tau a_1 + \tau^2 a_2 + \cdots, \tag{286}$$

where $\tau = t - t_{1/2}$. The first term is exactly solvable ($a_0 = A(t_{1/2})$) and, for many problems, it just provides the main contribution to the evolution of the system. In that case, it makes sense to take

$$\tilde{Y}_0(t) = \mathrm{e}^{(t-t_n)a_0} = \mathrm{e}^{(t-t_n)A(t_{1/2})}$$

and subsequently integrate Eq. (285) with

$$B(t) = \mathrm{e}^{-(t-t_n)A(t_{1/2})}\left(A(t) - A(t_{1/2})\right)\mathrm{e}^{(t-t_n)A(t_{1/2})}.$$

This approach has been considered in [161,162], and shows an extraordinary improvement when the system is highly oscillatory and the main oscillatory part is integrated with $\tilde{Y}_0$. In those cases, the norm of $B(t)$ is considerably smaller than $\|A(t)\|$, but $B(t)$ is still highly oscillatory, so that specially adapted quadrature rules have to be used in conjunction with the Magnus expansion [163,164].

In some other problems, the contributions from the derivatives can also be significant, so that a more appropriate transformation is defined by

$$\tilde{Y}_0(t) = \exp\left(\int_{t_n}^{t} A(\tau)\mathrm{d}\tau\right). \tag{287}$$

The resulting methods can be considered, then, as a combination of the Fer or Wilcox expansions and the Magnus expansion. This approach has been pursued in [165].

On the other hand, it is known that several physically relevant systems evolve adiabatically or almost-adiabatically. In that case, it seems appropriate to consider the adiabatic picture which instantaneously diagonalizes $A(t)$ (Section 2.9). This analysis is carried out in [166–168]. In [167] the adiabatic picture is used perturbatively, whereas in [166] it is shown that Magnus in the new picture leads to significant improvements.

Alternatively, one can analyze the structure of the leading error terms in order to identify the main contribution to the error at each Ω_i in the Magnus series expansion. In most cases, they correspond to terms involving only α_1 and its nested commutators with α_2. Thus, in particular, the standard fourth-order method given by (243) can be easily improved by including the dominant error term at higher orders, i.e.,

$$\Omega^{[4]} = \alpha_1 - \frac{1}{12}[\alpha_1,\alpha_2] + \frac{1}{720}[\alpha_1,\alpha_1,\alpha_1,\alpha_2] - \frac{1}{30\,240}[\alpha_1,\alpha_1,\alpha_1,\alpha_1,\alpha_1,\alpha_2] + \cdots.$$

We recall that using the fourth-order Gauss–Legendre quadrature rule we can take $\alpha_1 = \frac{h}{2}(A_1 + A_2)$, $\alpha_2 = \frac{h\sqrt{3}}{12}(A_2 - A_1)$ with A_1, A_2 given in (252). The new method requires additional commutators but the accuracy can be improved a good deal. This procedure is analyzed in [70], where it is shown how to sum up all terms of the form $[\alpha_1,\alpha_1,\ldots,\alpha_1,\alpha_2]$. An error analysis in the limit of very large values of $\|\alpha_1\|$ is done in [169,170].

5.8.2. Commutator-free Magnus integrators

All numerical methods based on the Magnus expansion appearing in the preceding sections require the evaluation of a matrix exponential which contains several nested commutators. As we have repeatedly pointed out, computing the exponential is frequently the most consuming part of the algorithm. There are problems where the matrix $A(t)$ has a sufficiently simple structure which allows one to efficiently approximate the exponential $\exp(A(t_i))$, or the exponential of a linear combination of the matrix $A(t)$ evaluated at different points. In some sense, this is equivalent to having efficient methods to compute or to approximate $\tilde{Y}_0$ in (287). It may be, however, that the computation of the matrix exponential is a much more involved task due to the presence of commutators in the Magnus expansion. For this reason, it makes sense to look for approximations to the Magnus expansion which do not involve commutators, whilst still preserving the same qualitative properties. In other words, one may be interested in compositions of the form

$$\Psi_m^{[n]} \equiv \exp\left(\int_{t_n}^{t_n+h} p_m(s)A(s)\mathrm{d}s\right)\ \cdots\ \exp\left(\int_{t_n}^{t_n+h} p_1(s)A(s)\mathrm{d}s\right) \tag{288}$$

where $p_i(s)$ are scalar functions chosen in such a way that $\Psi_m^{[n]} = \mathrm{e}^{\Omega(t_n+h)} + \mathcal{O}(h^{n+1})$. Alternatively, instead of the functions $p_i(s)$, it is possible to find coefficients $\varrho_{i,j}$, $i = 1,\ldots,m$, $j = 1,\ldots,k$ such that

$$\Psi_m^{[n]} \equiv \mathrm{e}^{\tilde{A}_m}\cdots\mathrm{e}^{\tilde{A}_1}, \quad \text{with } \tilde{A}_i = h\sum_{j=1}^{k}\varrho_{i,j}A_j \tag{289}$$

is an approximation of the same order. This procedure requires one first to compute $A_j = A(t_n + c_j h)$, $j = 1,\ldots,k$ for some quadrature nodes, c_j, of order n or higher and, obviously, the coefficients $\varrho_{i,j}$ will depend on this choice. The process simplifies if one works in the associated graded free Lie algebra generated by $\{\alpha_i\}$, as in the sequel of Eq. (242). Thus, achieving fourth-order integrators reduces just to solve the equations for the new coefficients $a_{i,1}$, $a_{i,2}$ in

$$\Psi_m^{[4]} \equiv \exp\left(a_{m,1}\,\alpha_1 + a_{m,2}\,\alpha_2\right)\cdots\exp\left(a_{1,1}\,\alpha_1 + a_{1,1}\,\alpha_2\right) \tag{290}$$

with the requirement that $\Psi_m^{[4]} = \exp\left(\Omega^{[4]}\right) + O(h^5)$, where $\Omega^{[4]}$ is given by (243). Here, the dependence of $a_{i,1}$, $a_{i,2}$ on the coefficients $\varrho_{i,j}$ is determined through the existing relationship between the α_i and the A_j given in (250). The order conditions for the coefficients $a_{i,1}$, $a_{i,2}$ can be easily obtained from the Baker–Campbell–Hausdorff formula. As we have

already mentioned, time-symmetry is an important property to be preserved by the integrators, whereas, at the same time, it also simplifies the analysis. Scheme (290) is time-symmetric if

$$a_{m+1-i,1} = a_{i,1}, \qquad a_{m+1-i,2} = -a_{i,2}, \quad i = 1, 2, \ldots, m \tag{291}$$

in which case the order conditions at even order terms are automatically satisfied. As an illustration, the simple compositions

$$\Psi_2^{[4]} \equiv \exp\left(\frac{1}{2}\alpha_1 + \frac{1}{6}\alpha_2\right) \exp\left(\frac{1}{2}\alpha_1 - \frac{1}{6}\alpha_2\right) \tag{292}$$

$$\Psi_3^{[4]} \equiv \exp\left(\frac{1}{12}\alpha_2\right) \exp\left(\alpha_1\right) \exp\left(-\frac{1}{12}\alpha_2\right) \tag{293}$$

are in fact fourth-order (commutator-free) methods requiring two and three exponentials, respectively [171]. In particular, scheme (292), when α_1, α_2 are approximated using the fourth-order Gauss–Legendre quadrature as shown in (252) and (253) leads to the scheme

$$\Psi_2^{[4]} \equiv \exp\left(h(\varrho_{2,1}A_1 + \varrho_{2,2}A_2)\right) \exp\left(h(\varrho_{1,1}A_1 + \varrho_{1,2}A_2)\right) \tag{294}$$

with $\varrho_{1,1} = \varrho_{2,2} = \frac{1}{2} + \frac{\sqrt{3}}{72}$, $\varrho_{1,2} = \varrho_{2,1} = \frac{1}{2} - \frac{\sqrt{3}}{72}$. Methods closely related to the scheme (293) are presented in [171–173], where they are applied to the Schrödinger equation with a time-dependent potential. A method quite similar to (292) is analyzed in [174] through its application to parabolic initial boundary value problems. A detailed study of fourth and sixth order commutator-free methods is presented in [175].

On the other hand, very often the differential equation (223) can be split into two parts, so that one has, instead,

$$Y' = (A(t) + B(t))\, Y, \tag{295}$$

where each part can be trivially or very efficiently solved. For instance, the Schrödinger equation with a time-dependent potential and, possibly, a time-dependent kinetic energy belongs to this class. In principle, the following families of geometric integrators are specially tailored for this problem:

I- The commutator-free Magnus integrators (289), which in this case read

$$\Psi_m^{[n]} \equiv \mathrm{e}^{\tilde{A}_m + \tilde{B}_m} \cdots \mathrm{e}^{\tilde{A}_1 + \tilde{B}_1}, \quad \text{with } \tilde{A}_i = h\sum_{j=1}^{k} \varrho_{i,j} A_j,\ \tilde{B}_i = h\sum_{j=1}^{k} \varrho_{i,j} B_j. \tag{296}$$

Assuming that $\mathrm{e}^{\tilde{A}_i}$ and $\mathrm{e}^{\tilde{B}_i}$ are easily computed, then each exponential can be approximated by a conveniently chosen splitting method (279) [128]

$$\mathrm{e}^{\tilde{A}_i + \tilde{B}_i} \simeq \mathrm{e}^{b_s \tilde{B}_i} \mathrm{e}^{a_s \tilde{A}_i} \cdots \mathrm{e}^{b_1 \tilde{B}_i} \mathrm{e}^{a_1 \tilde{A}_i}. \tag{297}$$

II- If one takes the time variable in $A(t), B(t)$ as two new coordinates, one may use any splitting method as follows [176]:

$$\Psi_{l,h}^{[n]} \equiv \mathrm{e}^{b_l h B(w_l)} \mathrm{e}^{a_l h A(v_l)} \cdots \mathrm{e}^{b_1 h B(w_1)} \mathrm{e}^{a_1 h A(v_1)}, \tag{298}$$

with

$$v_i = \sum_{j=1}^{i-1} b_j, \qquad w_i = \sum_{j=1}^{i} a_j.$$

and $b_0 = 0,\ A(v_i) \equiv A(t_n + v_i h),\ B(w_i) \equiv B(t_n + w_i h)$.

Both approaches have pros and cons. By applying procedure I we may get methods of order $2n$ with only n evaluations of $A(t), B(t)$ using, for example, Gauss–Legendre quadratures, but if m in (296) is large, the number of matrix exponentials to be computed leads to exceedingly costly methods. The approach II, on the other hand, has the advantage of a smaller number of stages, but also presents two drawbacks: (i) many evaluations of $A(t), B(t)$ are required in general; (ii) for matrices A and B with a particular structure, there are specially designed splitting methods which are far more efficient, but these schemes are not easily adapted to this situation.

Next, we show how to combine splitting methods with techniques leading to commutator free Magnus schemes, to design efficient numerical algorithms possessing the advantages of approaches I and II, and at the same time generalizing the splitting idea (279) to this setting [177,178].

The starting point is similar as in previous schemes, i.e. we consider a composition of the form

$$\psi_{l,h}^{[n]} = \mathrm{e}^{\tilde{B}_l}\, \mathrm{e}^{\tilde{A}_l} \cdots \mathrm{e}^{\tilde{B}_1}\, \mathrm{e}^{\tilde{A}_1}, \tag{299}$$

Table 1
Splitting methods of order 4 for separable non-autonomous systems. GS_6-4 is intended for general separable problems, whereas MN_6-4 can be applied when $[B(t_i),[B(t_j),[B(t_k),A(t_l)]]]=0$. All the coefficients are given in terms of b_{11}, a_{11}, b_{21}, a_{21}, b_{31} for each method.

GS_6-4	MN_6-4
$b_{11}=0.0792036964311957$	$b_{11}=0.0829844064174052$
$a_{11}=0.209515106613362$	$a_{11}=0.245298957184271$
$b_{21}=0.353172906049774$	$b_{21}=0.396309801498368$
$a_{21}=-0.143851773179818$	$a_{21}=0.604872665711080$
$b_{31}=-0.0420650803577195$	$b_{31}=-0.0390563049223486$
$a_{31}=1/2-(a_{11}+a_{21})$	$b_{41}=1-2(b_{11}+b_{21}+b_{31})$
$a_{41}=a_{31}$	$b_{51}=b_{31}$
$a_{51}=a_{21}$	$b_{61}=b_{21}$
$a_{61}=a_{11}$	$b_{71}=b_{11}$
$a_{12}=(2a_{11}+2a_{21}+a_{31}-2b_{11}-2b_{21})/c$	$b_{12}=(2a_{11}+2a_{21}-2b_{11}-b_{21})/d$
$a_{22}=0$	$b_{22}=(-2a_{11}+b_{11})/d$
$a_{32}=-a_{11}/c$	$b_{32}=b_{42}=0$
$a_{42}=-a_{32}$	$b_{52}=-b_{32}$
$a_{52}=-a_{22}$	$b_{62}=-b_{22}$
$a_{62}=-a_{12}$	$b_{72}=-b_{12}$
$c=12(a_{11}+2a_{21}+a_{31}-2b_{11}+2a_{11}b_{11}-2b_{21}+2a_{11}b_{21})$	
$d=12(2a_{21}-b_{11}+2a_{11}b_{11}-2a_{21}b_{11}-b_{21}+2a_{11}b_{21})$	

where the matrices $\tilde{A}_i$ and $\tilde{B}_i$ are

$$\tilde{A}_i=h\sum_{j=1}^{k}\rho_{ij}A_j,\qquad \tilde{B}_i=h\sum_{j=1}^{k}\sigma_{ij}B_j, \tag{300}$$

with appropriately chosen real parameters ρ_{ij}, σ_{ij} depending on the coefficients of the chosen quadrature rule. Notice that $e^{\tilde{A}_i}$ can be seen as the solution of the initial value problem $Y'=\hat{A}_iY$, $Y(t_n)=I$ at t_{n+1}, where $\tilde{A}_i=h\hat{A}_i$. Of course, the same considerations apply to $e^{\tilde{B}_i}$.

In many cases it is convenient to write the coefficients ρ_{ij}, σ_{ij} explicitly in terms of the coefficients c_i. Following [177] they can be written as

$$\rho_{ij}=\sum_{l=1}^{s}a_{i,l}\left(R^{(s)}Q_X^{(s,k)}\right)_{lj},\qquad \sigma_{ij}=\sum_{l=1}^{s}b_{i,l}\left(R^{(s)}Q_X^{(s,k)}\right)_{ij}. \tag{301}$$

where the coefficients for the matrices $R^{(s)}$, $s=2,3$ are given in (249) and for $Q_X^{(s,k)}$ (whose elements depend on the coefficients b_i, c_i for the quadrature rule) as shown in (246).

In this way, the coefficients a_{ij} and b_{ij} are independent of the quadrature choice and can be obtained by solving some order conditions (see [177] for more details).

This procedure allows us to separately analyze particular cases for the matrices A, B in order to build efficient methods. For instance, in [177] the following particular cases are considered: (i) when the matrices $A(t)$, $B(t)$ have a general structure; (ii) when they satisfy the additional constraint $[B(t_i),[B(t_j),[B(t_k),A(t_l)]]]=0$ as it happens, for instance, if A corresponds to the kinetic energy and B to the potential energy (both in classical or quantum mechanics).

As an illustration, we consider the following 4th-order 6-stage BAB composition

$$\psi_{6,h}^{[4]}=e^{\tilde{B}_7}\,e^{\tilde{A}_6}\,e^{\tilde{B}_6}\cdots e^{\tilde{A}_1}\,e^{\tilde{B}_1}. \tag{302}$$

In Table 1 we collect the coefficients a_{ij}, b_{ij} to be used in (301) to obtain the coefficients ρ_{ij}, σ_{ij} to be used in the scheme (302) for two methods, denoted by GS_6-4 in the general case (whose coefficients a_{i1}, b_{i1} correspond to S_6 in [154]) and MN_6-4 when $[B(t_i),[B(t_j),[B(t_k),A(t_l)]]]=0$ (the coefficients a_{i1}, b_{i1} correspond to $SRKN_6^b$ in [154]).

Finally, one has to write the scheme in terms of the matrices A_i, B_i. For instance, the composition (302) with the 4th-order Gauss–Legendre quadrature (i.e. taking $Q^{(2,2)}$ in (247) and $R^{(2)}$ in (249) to obtain the coefficients ρ_{ij}, σ_{ij} in (301)) gives

$$\begin{aligned}\tilde{A}_i&=\left(\frac{1}{2}a_{i1}-\sqrt{3}a_{i2}\right)hA_1+\left(\frac{1}{2}a_{i1}+\sqrt{3}a_{i2}\right)hA_2\\ \tilde{B}_i&=\left(\frac{1}{2}b_{i1}-\sqrt{3}b_{i2}\right)hB_1+\left(\frac{1}{2}b_{i1}+\sqrt{3}b_{i2}\right)hB_2.\end{aligned} \tag{303}$$

5.9. Magnus integrators for nonlinear differential equations

The success of Magnus methods applied to the numerical integration of linear systems has motivated several attempts to adapt the schemes for solving time dependent nonlinear differential equations. For completeness we present some recently proposed generalizations of Magnus integrators. We consider two different problems: (i) a nonlinear matrix equation defined in a Lie group, and (ii) a general nonlinear equation to which the techniques of Section 3.4 can be applied.

5.9.1. Nonlinear matrix equations in Lie groups

As we have already mentioned, the strategy adopted by most Lie-group methods for solving the nonlinear matrix differential equation (139),

$$Y' = A(t, Y)Y, \qquad Y(0) = Y_0 \in \mathcal{G}$$

defined in a Lie group $\mathcal{G}$, whilst preserving its Lie group structure, is to lift $Y(t)$ from $\mathcal{G}$ to the underlying Lie algebra $\mathfrak{g}$ (usually with the exponential map), then formulate and numerically solve there an associated differential equation, and finally map the solution back to $\mathcal{G}$. In this way, the discretization procedure works in a linear space rather than in the Lie group. In particular, the idea of the so-called Runge–Kutta–Munthe-Kaas class of schemes is to approximate the solution of the associated differential equation in the Lie algebra $\mathfrak{g}$ by means of a classical Runge–Kutta method [59,179,180].

To generalize Magnus integrators when $A = A(t, Y)$, an important difference with respect to the linear case is that, now, multivariate integrals depend also on the value of the (unknown) variable Y at quadrature points. This leads to implicit methods and nonlinear algebraic equations in every step of the integration [181], which in general cannot compete in efficiency with other classes of geometric integrators such as splitting and composition methods.

An obvious alternative is just to replace the integrals appearing in the nonlinear Magnus expansion developed in Section 3.3 by affordable quadratures, depending on the particular problem. If, for instance, we use Euler's method to approximate the first term in (142), $\Omega^{[1]}(h) = hA(0, Y_0) + \mathcal{O}(h^2)$ and $\Omega^{[2]}$ is discretized with the midpoint rule, we get the second order scheme

$$\begin{aligned} v_2 &\equiv hA\left(\frac{h}{2}, \mathrm{e}^{\frac{h}{2}A(0,Y_0)}Y_0\right) = \Omega^{[2]}(h) + \mathcal{O}(h^3) \\ Y_1 &= \mathrm{e}^{v_2}Y_0. \end{aligned} \tag{304}$$

The same procedure can be carried out at higher orders, consistently discretizing the integrals appearing in $\Omega^{[m]}(h)$ for $m > 2$ [99].

5.9.2. The general nonlinear problem

In principle, it is possible to adapt all methods built for linear problems to the general nonlinear non-autonomous equation (159)

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}),$$

or equivalently, the operator differential equation (160),

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^t = \Phi^t L_{\mathbf{f}(t,\mathbf{y})}, \quad \mathbf{y} = \mathbf{x}_0.$$

As we pointed out in Section 3.4, there are two problematic aspects when designing practical numerical schemes based on Magnus expansion in the nonlinear case. The first one is how to compute or approximate the truncated Magnus expansion (or its action on the initial conditions). The second one is how to evaluate the required Lie transforms. For example, to compute the Lie transform $\exp(tL_{\mathbf{f}(\mathbf{y})})$ acting on $\mathbf{y}$ is equivalent to solving the autonomous differential equation $\mathbf{x}' = \mathbf{f}(\mathbf{x})$ at $t = h$ with $\mathbf{x}(0) = \mathbf{y}$, or $\mathbf{x}(t) = \exp(tL_{\mathbf{f}(\mathbf{x}_0)})\mathbf{x}_0$ where $\mathbf{x}_0 = \mathbf{y}$ can be considered as a set of coordinates.

Very often, the presence of Lie brackets in the exponent leads to fundamental difficulties, since the resulting vector fields usually have very complicated structures. Sometimes, however, this problem can be circumvented by using the same techniques, leading to commutator-free Magnus integrators in the linear case. In any case, one should bear in mind that the action of the exponentials in the methods designed for the linear case has to be replaced by their corresponding maps. Alternatively, if the method is formulated in terms of Lie transforms, the order of the exponentials has to be reversed, according to Eq. (155).

Next we illustrate how to numerically solve the problem

$$\mathbf{x}' = \mathbf{f}_1(t, \mathbf{x}) + \mathbf{f}_2(t, \mathbf{x}) \tag{305}$$

using the scheme (299) with (303) and the coefficients a_{ij}, b_{ij} taken from MN_6-4 in Table 1.

Let us consider the Duffing equation

$$q'' + \epsilon q' + q^3 - q = \delta\cos(\omega t) \tag{306}$$

Table 2
Algorithms for the numerical integration of (307) or (308): (Algorithm 1) with scheme (298), and (Algorithm 2) with scheme (302).

Algorithm 1: Standard split	**Algorithm 1: Magnus split**
$q_0 = q(t_n); \quad p_0 = p(t_n);$ $t_a = t_n; \quad t_b = t_n$ **do** $i = 1, m$ $p_i = p_{i-1} - ha_i V'(t_a, q_{i-1})$ $t_a = t_a + ha_i$ $q_i = q_{i-1} + hb_i T'(t_b, p_i)$ $t_b = t_b + hb_i$ **enddo**	$q_0 = q(t_n); \quad p_0 = p(t_n);$ **do** $i = 1, k$ $T_i'(p) = T'(t_n + c_i h, p); \quad V_i'(q) = V'(t_n + c_i h, q)$ **enddo** **do** $i = 1, m$ $\tilde{V}_i(q) = \sigma_{i1} V_1'(q) + \cdots + \sigma_{ik} V_k'(q)$ $\tilde{T}_i(p) = \rho_{i1} T_1'(p) + \cdots + \rho_{ik} T_k'(p)$ $p_i = p_{i-1} - h\tilde{V}_i(q_{i-1})$ $q_i = q_{i-1} + h\tilde{T}_i(p_i)$ **enddo**

which can be obtained from the time-dependent Hamiltonian

$$H(q,p,t) = T(p,t) + V(q,t) = \mathrm{e}^{-\epsilon t}\,\frac{1}{2}p^2 + \mathrm{e}^{\epsilon t}\left(\frac{1}{4}q^4 - \frac{1}{2}q^2 - \delta\cos(\omega t)q\right) \tag{307}$$

or equivalently from

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{Bmatrix} q \\ p \end{Bmatrix} = \begin{Bmatrix} T'(t,p) \\ -V'(t,q) \end{Bmatrix} = \begin{Bmatrix} \mathrm{e}^{-\epsilon t}p \\ 0 \end{Bmatrix} + \begin{Bmatrix} 0 \\ \mathrm{e}^{\epsilon t}\left(q - q^3 + \delta\cos(\omega t)\right) \end{Bmatrix}. \tag{308}$$

Notice that this system already has the form (305), each part being exactly solvable. In consequence, the splitting method shown in (298) can be used here. The procedure is described as Algorithm 1 in Table 2.

Observe that the leap-frog composition (281) corresponds to $m = 2$ and

$$a_1 = a_2 = \frac{1}{2}, \qquad b_1 = 1, \qquad b_2 = 0. \tag{309}$$

Since $b_2 = 0$ one stage can be saved (with a trivial modification of the algorithm) and the scheme is considered as a one stage method. An efficient symmetric 5-stage fourth order integrator is given by the coefficients ($m = 6$)

$$a_i = \frac{\gamma_i + \gamma_{i-1}}{2}, \qquad b_i = \gamma_i. \tag{310}$$

$i = 0, 1, \ldots, 6$ with $\gamma_0 = \gamma_6 = 0$ and $\gamma_1 = \gamma_2 = \gamma_4 = \gamma_5 = 1/(4 - 4^{1/3})$, $\gamma_3 = 1 - 4\gamma_1$.

Alternatively, we can use the Magnus integrator (302). Since the kinetic energy is quadratic in momenta, we can apply the fourth-order method MN_6-4. If we take the fourth-order Gauss-Legendre quadrature rule for the evaluation of the time-dependent function then we can consider (303), where the coefficients a_{ij}, b_{ij} are given in Table 1. Here, $A(t)$ plays the role of $T'(t,p)$ and $B(t)$ plays the role of $V'(t,q)$ (they are not interchangeable, otherwise the performance seriously deteriorates). The computation of one time step is shown as Algorithm 2 in Table 2.

We take $\epsilon = 1/20$, $\delta = 1/4$, $\omega = 1$ and initial conditions $q(0) = 1.75$, $p(0) = 0$. We integrate up to $t = 10\,\pi$ and measure the average error in phase space, in terms of the number of force evaluations for different time steps (in logarithmic scale). The results are shown in Fig. 20. The scheme MN_6-4 has 6 stages per step, but only two time-evaluations. For this reason, in the figure, we have considered, as the number of evaluations per step, both two and six (left and right curves connected by an arrow). The superiority of the new splitting Magnus integrators for this problem is evident. If the time-dependent functions dominate the cost of the algorithm, the superiority is even higher. Surprisingly, the method shows better stability than the leap-frog method, which attains the highest stability possible among the splitting methods for autonomous problems.

6. Some applications of the numerical integrators based on ME

In this section we collect several examples where the numerical integration methods based on the Magnus expansion have been applied in the recent literature. Special attention is dedicated to the numerical integration of the Schrödinger equation, since the Magnus series expansion has been extensively used in this setting almost since its very formulation. The time-independent Schrödinger equation can be considered as a particular example of a Sturm–Liouville problem, so we also review the applicability of Magnus based techniques in this context. Then we consider a particular nonlinear system (the differential Riccati equation) which can be, in some sense, linearized, so that at the end one may work with finite-dimensional matrices. Finally, we summarize a recent but noteworthy application: the design of new classes of numerical schemes for the integration of stochastic differential equations.

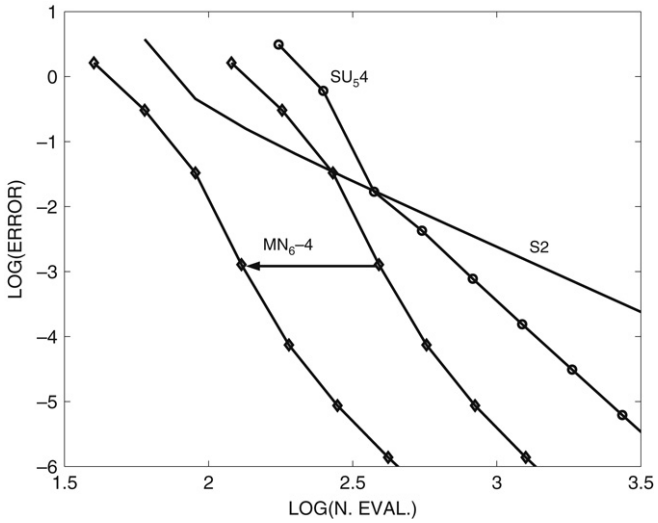


Fig. 20. Average error versus number of force evaluations in the numerical integration of (308) using second and fourth order symplectic integrators for general separable systems (S2 corresponds to the second order leapfrog method with coefficients (309) and $SU_5 4$ to the fourth order method with coefficients (310)) and the fourth order symplectic Runge–Kutta–Nyström method MN_6-4 with initial conditions $q(0) = 1.75$, $p(0) = 0$ and $\epsilon = 1/20$, $\delta = 1/4$, $\omega = 1$.

6.1. Case study: Numerical treatment of the Schrödinger equation

Before embarking on the use of numerical methods based on the Magnus expansion in the integration of the Schrödinger equation, let us establish first the theoretical framework which allows one to use numerical integrators in this setting for obtaining approximate solutions in time and space.

6.1.1. Time-dependent Schrödinger equation

To keep the treatment as simple as possible, we commence by considering the one-dimensional time-dependent Schrödinger equation ($\hbar = 1$)

$$\mathrm{i}\frac{\partial}{\partial t}\psi(t,x) = H\psi(t,x) \equiv -\frac{1}{2}\frac{\partial^2}{\partial x^2}\psi(t,x) + V(x)\psi(t,x), \tag{311}$$

with initial condition $\psi(0,x) = \psi_0(x)$. If we look for a solution of the form $\psi(t,x) = \phi(t)\varphi(x)$, it is clear that, by substituting into (311), one gets $\phi(t) = \mathrm{e}^{-\mathrm{i}tE}$, where E is a constant and $\varphi(x)$ is the solution of the second order differential equation

$$-\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} + V(x)\varphi = E\varphi. \tag{312}$$

If $E > V$ the solution is oscillatory, whereas if $E < V$ the solution is a linear combination of exponentially increasing and decreasing functions. For bounded problems, this last condition always takes place at the boundaries. Since

$$\int |\psi(x,t)|^2 \mathrm{d}x = \int |\varphi(x)|^2 \mathrm{d}x < \infty, \tag{313}$$

it is clear that the exponentially increasing solutions have to be canceled, and this can only occur for certain values of the constant E, which are precisely the eigenvalues of the problem.

Let us assume that the system has only a discrete spectrum, and denote by $\{E_n, \varphi_n\}_{n=0}^{\infty}$, with $E_i < E_j,\ i < j$, the complete set of eigenvalues and associated eigenvectors. It is well known that we can take $\{\varphi_n\}_{n=0}^{\infty}$ as an orthonormal basis and, since (311) is linear, any solution can be written as

$$\psi(t,x) = \sum_{n=0}^{\infty} c_n\, \mathrm{e}^{-\mathrm{i}tE_n}\varphi_n(x). \tag{314}$$

Using the standard notation for the inner product, one has

$$\langle\varphi_n(x)|\psi(t,x)\rangle = \int \varphi_n^*(x)\,\psi(t,x)\mathrm{d}x = c_n\,\mathrm{e}^{-\mathrm{i}tE_n} \tag{315}$$

and

$$|\langle\varphi_n(x)|\psi(t,x)\rangle|^2 = |c_n|^2$$

is the probability of finding the system in the eigenstate φ_n, so that $\sum_n |c_n|^2 = 1$. The energy is given by

$$\mathcal{E} = \langle\psi|H|\psi\rangle = \int \psi^*(t,x) H\,\psi(t,x)\mathrm{d}x = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} c_n^* c_m H_{n,m}, \tag{316}$$

where

$$H_{n,m} \equiv \langle m|H|n\rangle = \langle\varphi_m|H|\varphi_n\rangle.$$

In general, the coefficients c_n decrease very fast with n and, in some cases, the system allows only a finite number of states. In that situation, one may consider the Schrödinger equation as a finite dimensional linear system where the Hamiltonian is a matrix with elements $H_{n,m}$. This is precisely the case for the examples examined in Section 4.

When the Hamiltonian is explicitly time-dependent, this procedure is no longer valid. Instead, one may use some alternative techniques which we now briefly review.

(i) *Spectral decomposition*. Let us assume that the system is perturbed with a time-dependent potential, i.e., Eq. (311) takes the form

$$\mathrm{i}\frac{\partial}{\partial t}\psi(t,x) = \hat{H}(t)\psi(t,x) \equiv (\hat{T} + \hat{V}(t))\psi(t,x), \tag{317}$$

where

$$\hat{T}\psi \equiv -\frac{1}{2}\frac{\partial^2\psi}{\partial x^2}, \qquad \hat{V}(t)\psi \equiv (V(x) + \tilde{V}(t,x))\psi.$$

In this case, we cannot use separation of variables. However, since $\{\varphi_n\}$ is a complete basis we can still write the solution as

$$\psi(t,x) \simeq \sum_{n=0}^{d-1} c_n(t)\mathrm{e}^{-\mathrm{i}tE_n}\,\varphi_n(x), \tag{318}$$

where E_n and φ_n are the exact eigenvalues and eigenfunctions when $\tilde{V} = 0$, and the complex coefficients c_n give the probability amplitude to find the system in the state φ_n ($\sum_n |c_n(t)|^2 = 1$ for all t). Then, substituting (318) into (317) we obtain the matrix equation

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{c}(t) = \mathbf{H}(t)\mathbf{c}(t), \quad \mathbf{c}(0) = \mathbf{c}_0, \tag{319}$$

where $\mathbf{c} = (c_0, \ldots, c_{d-1})^{\mathrm{T}} \in \mathbb{C}^d$ and $\mathbf{H} \in \mathbb{C}^{d\times d}$ is a Hermitian matrix associated to the Hamiltonian

$$(\mathbf{H}(t))_{ij} = \langle\varphi_i|\hat{H}(t) - \hat{H}_0|\varphi_j\rangle\,\mathrm{e}^{\mathrm{i}(E_i - E_j)t}, \quad i,j = 1,\ldots,d$$

and $\hat{H}_0 = \hat{H}(t=0)$. Given the initial wave function $\psi(0,x)$, the components of $\mathbf{c}_0$ are determined by $c_{0,i} = \langle\varphi_i|\psi(0,x)\rangle$.

Obviously, any complete basis can be used in this case, although the norm of the matrix $\mathbf{H}(t)$ may depend on the choice. In addition, the number of basis elements (i.e. the minimum dimension d necessary to obtain a sufficiently accurate result) also depends on the chosen basis.

(ii) *Space discretization*. This procedure intends to take advantage of the structure of the Hamiltonian $\hat{H}$ in (317): $\hat{V}$ is diagonal in the coordinate space and $\hat{T}$ is diagonal in the momentum space. Let us assume that the system is defined in the interval $x \in [x_0, x_f]$ with periodic boundary conditions. We can then split this interval in d parts of length $\Delta x = (x_f - x_0)/d$ and consider $c_n = \psi(t, x_n)$ where $x_n = x_0 + n\Delta x$, $n = 0, 1, \ldots, d-1$. Then a finite dimensional linear equation similar (but with a different coefficient matrix $\mathbf{H}$) to Eq. (319) results. Since $\hat{V}$ is diagonal in the coordinate space and $\hat{T}$ is diagonal in momentum space, it is possible to use complex Fast Fourier Transforms (FFTs) for evaluating the products $\mathbf{Hc}$, where $\hat{T}\psi(t,x_n) = \mathcal{F}^{-1}D_T\mathcal{F}\psi(t,x_n)$, and D_T is a diagonal operator.

We thus see that whatever the procedure used (spectral decomposition or space discretization), one ends up with a linear equation of the form

$$\mathrm{i}\frac{\mathrm{d}\psi}{\mathrm{d}t}(t) = H(t)\psi(t), \qquad \psi(0) = \psi_0 \tag{320}$$

where $\psi(t)$ now represents a complex vector with d components which approximates the (continuous) wave function. The computational Hamiltonian $H(t)$ appearing in (320) is thus a space discretization (or other finite-dimensional model) of $\hat{H}(t) = \hat{T} + \hat{V}(t)$. Numerical difficulties come mainly from the unbounded nature of the Hamiltonian and the highly oscillatory behavior of the wave function.

It is at this point when numerical algorithms based on the Magnus expansion, such as they have been formulated in previous sections, come into play for integrating in time the linear system (320). To put them in perspective, let us first introduce some other numerical methods also used in this context. Our exposition is largely based on the reference [182].

6.1.1.1. The implicit midpoint rule. The approximation to the solution of (320) provided by this scheme is implicitly defined by

$$\mathrm{i}\frac{\psi_{n+1}-\psi_n}{\Delta t} = H(t_{n+1/2})\frac{1}{2}(\psi_{n+1}+\psi_n), \tag{321}$$

where $t_{n+1/2} = \frac{1}{2}(t_{n+1}+t_n)$. Here, and in the sequel, for clarity, we have denoted by Δt the time step size and $t_n = n\Delta t$. Alternatively,

$$\psi_{n+1} = r(-\mathrm{i}\Delta t H(t_{n+1/2}))\psi_n, \quad \text{with } r(z) = \frac{1+\frac{1}{2}z}{1-\frac{1}{2}z}. \tag{322}$$

Observe that, as r is nothing but the Cayley transform, the numerical propagator is unitary and consequently the Euclidean norm of the discrete wave function is preserved along the evolution: $\|\psi_{n+1}\| = \|\psi_n\|$. This is a crucial qualitative feature the method shares with the exact solution, contrarily to other standard numerical integrators, such as explicit Runge–Kutta methods. From a purely numerical point of view, the algorithm is stable for any step size Δt.

Another useful property of this numerical scheme is time-symmetry: exchanging in (322) n by $n+1$ and Δt by $-\Delta t$ we get the same numerical method again. Equivalently, $r(-z) = r(z)^{-1}$, exactly as the exponential e^z.

With respect to accuracy, it is not difficult to show that, if $H(t)$ is bounded and sufficiently smooth, the error verifies

$$\|\psi_n - \psi(t_n)\| = \mathcal{O}(\Delta t^2) \tag{323}$$

uniformly for $n\Delta t$ in a time interval $[0, t_f]$. In other words, the implicit midpoint rule is a second-order method. It happens, however, that the constant in the term $\mathcal{O}(\Delta t^2)$ depends on bounds of H' and H'' and on the norm of the third derivative of the solution ψ. Since, in general, the wave function is highly oscillatory in time, this time derivative can become large, and so the use of very small time steps is mandatory.

6.1.1.2. The exponential midpoint rule. Another possibility to get approximate solutions of (320) consists of replacing $r(z)$ by $\exp(z)$ in (322):

$$\psi_{n+1} = \exp(-\mathrm{i}\Delta t\, H(t_{n+1/2}))\psi_n. \tag{324}$$

Now, instead of solving systems of linear equations as previously, one has to compute the exponential of a large matrix times a vector at each integration step. In this respect, the techniques reviewed in Section 5.6 can be efficiently implemented. The exponential midpoint rule (324) also provides a unitary propagator and it is time-symmetric. In addition, the error satisfies the same condition (323), but now the constant in the $\mathcal{O}(\Delta t^2)$ term is independent of the time derivatives of ψ under certain assumptions on the commutator $[H(t), H(s)]$ [183]. As a consequence, much larger time steps can be taken to achieve the same accuracy as with the implicit midpoint rule.

6.1.1.3. Integrators based on the Magnus expansion. The method (324) is a particular instance of a second order Magnus method when the integral $\int_0^{\Delta t} H(s)\mathrm{d}s$ is replaced by the midpoint quadrature rule. In fact, we have already used it in (274). Obviously, if higher order approximations are considered, the accuracy can be enhanced a great deal. This claim has to be conveniently justified, however, since the order of the numerical methods based on Magnus has been deduced when $\|\Delta t H(t)\| \to 0$ and is obtained by studying the remainder of the truncated Magnus series. In the Schrödinger equation, on the other hand, one has to cope with discretizations of unbounded operators, so in principle it is not evident how the previous results on the order of accuracy apply in this context. In [183], Hochbruck and Lubich analyze, in detail, the application of the fourth-order Magnus integrator (254) to Eq. (320), showing that it works extremely well even with step sizes for which the corresponding $\|\Delta t H(t)\|$ is large. In particular, the scheme retains fourth order of accuracy in Δt *independently of the norm* of $H(t)$ when $H(t) = T + V(t)$, T is a discretization of $-\frac{1}{2}\frac{\partial^2}{\partial x^2}$ (with maximum eigenvalue $E_{\max} \sim (\Delta x)^{-2}$) and $V(t)$ is sufficiently smooth under the time step restriction $\Delta t\sqrt{E_{\max}} \leq \mathit{Const}$. This is so even when there is no guarantee that the Magnus series converges at all.

6.1.1.4. Symplectic perspective. The evolution operator corresponding to (317) is not only unitary, but also symplectic with canonical coordinates and momenta $\mathrm{Re}(\psi)$ and $\mathrm{Im}(\psi)$, respectively. If we carry out a discretization in space, this symplectic structure is inherited by the corresponding Eq. (319). It makes sense, then, to write $\mathbf{c} = \mathbf{q} + \mathrm{i}\mathbf{p}$ and consider the equations satisfied by $\mathbf{q}, \mathbf{p} \in \mathbb{R}^d$, namely

$$\mathbf{q}' = \mathbf{H}(t)\mathbf{p}, \qquad \mathbf{p}' = -\mathbf{H}(t)\mathbf{q}, \tag{325}$$

which can be interpreted as the canonical equations corresponding to the Hamiltonian [184]

$$\mathcal{H}(t, \mathbf{q}, \mathbf{p}) = \mathbf{p}^{\mathrm{T}}\mathbf{H}(t)\mathbf{p} + \mathbf{q}^{\mathrm{T}}\mathbf{H}(t)\mathbf{q}. \tag{326}$$

Denoting $\mathbf{z} = (\mathbf{q}, \mathbf{p})^{\mathrm{T}}$, it is clear that

$$\mathbf{z}' = (\mathbf{A}(t) + \mathbf{B}(t))\ \mathbf{z}$$

where

$$\mathbf{A}(t) = \begin{pmatrix} 0 & \mathbf{H}(t) \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{B}(t) = \begin{pmatrix} 0 & 0 \\ -\mathbf{H}(t) & 0 \end{pmatrix}. \tag{327}$$

For this system it is possible, therefore, to apply the commutator-free Magnus integrators constructed in Section 5.8.2. In addition, one has

$$[\mathbf{B}, [\mathbf{B}, [\mathbf{B}, \mathbf{A}]]] = [\mathbf{A}, [\mathbf{A}, [\mathbf{A}, \mathbf{B}]]] = \mathbf{0},$$

and this property allows us to use specially designed and highly efficient integration methods [178].

6.1.2. Time-independent Schrödinger equation

Restricting ourselves to the time-independent Schrödinger equation (312), we next illustrate how Magnus integrators can, in fact, be used to compute the discrete eigenvalues defined by the problem. Although only the Schrödinger equation in a finite domain is considered,

$$-\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} + V(x)\varphi = \lambda\varphi, \quad x \in (a, b) \tag{328}$$

the procedure can be easily adapted to other types of eigenvalue problems, in which one has to find both $\lambda \equiv E$ and φ. Here it is assumed that the potential is smooth, $V \in C^m(a, b)$ and, for simplicity, $\varphi(a) = \varphi(b) = 0$.

Under these assumptions, it is well known that the eigenvalues are real, distinct and bounded from below. The problem (328) can be formulated in the special linear group SL(2),

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}x} = \begin{pmatrix} 0 & 1 \\ V(x) - \lambda & 0 \end{pmatrix}\mathbf{y}, \quad x \in (a, b), \quad \text{where } \mathbf{y} = (\varphi, \mathrm{d}\varphi/\mathrm{d}x)^{\mathrm{T}}, \tag{329}$$

so that the Magnus expansion can be applied in a natural way. As usual, rather than approximating the fundamental solution of (329) in the entire interval (a, b) by $\exp(\Omega)$, the idea is to partition the interval into N small subintervals, and then apply a conveniently discretized version of the Magnus expansion. In this way, the convergence problem no longer restricts the size $(b - a)$ [70].

For the sake of simplicity, let us consider the fourth-order method (254). Writing

$$V_{n,1} = V\left(x_n + \left(\frac{1}{2} - \frac{\sqrt{3}}{6}\right)h\right), \qquad V_{n,2} = V\left(x_n + \left(\frac{1}{2} + \frac{\sqrt{3}}{6}\right)h\right),$$

where $h = (b - a)/N$ and $x_n = a + h\,n$, we form

$$\sigma_n(\lambda) = \begin{pmatrix} -\dfrac{\sqrt{3}}{12}h^2(V_{n,1} - V_{n,2}) & h \\ \dfrac{1}{2}h(V_{n,1} + V_{n,2}) - h\lambda & \dfrac{\sqrt{3}}{12}h^2(V_{n,1} - V_{n,2}) \end{pmatrix},$$

for $n = 0, 1, \ldots, N - 1$. Then, the fourth-order approximation to the solution of (329) at $x = b$ is

$$\mathbf{y}(b) = \mathrm{e}^{\sigma_{N-1}(\lambda)} \cdots \mathrm{e}^{\sigma_1(\lambda)}\,\mathrm{e}^{\sigma_0(\lambda)}\mathbf{y}(a) \tag{330}$$

and the values of λ are obtained from (330) by repeatedly using the expression of the exponential of a traceless matrix, Eq. (24), and requiring that $\varphi(a) = \varphi(b) = 0$. The resulting nonlinear equation in λ can be solved, for instance, by Newton–Raphson iteration, which provides quadratic convergence for starting values sufficiently near the solution [70].

Although, by construction this procedure leads to a global order of approximation $\mathcal{O}(h^p)$ if a pth-order Magnus method is applied, it turns out that the error also depends on the magnitude of the eigenvalue. Specifically, the error in a pth-order method grows as $\mathcal{O}(h^{p+1}\lambda^{p/2-1})$ [70], and thus one expects poor approximations for large eigenvalues. This difficulty can be overcome, up to a point, by analyzing the dependence on λ of each term in the Magnus series and considering partial sums of the terms carrying the most significant dependence on λ. For instance, it is possible to design a sixth-order Magnus integrator for this problem with error $\mathcal{O}(h^7\lambda)$, which therefore behaves like a fourth-order method when $h^2\lambda \approx 1$, whereas the standard sixth-order Magnus scheme, carrying an error of $\mathcal{O}(h^7\lambda^2)$, reduces to an order-two method [70]. In any case, getting accurate approximations when $|\lambda| \to \infty$ is more problematic [185].

6.2. Sturm–Liouville problems

The system defined by (328) with boundary conditions $\varphi(a) = \varphi(b) = 0$ is just one particular example of a second order Sturm–Liouville problem [186,187]. It is thus quite natural to try to apply Magnus integrators to more general problems within this class.

A second order Sturm–Liouville eigenvalue problem has the form

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}y}{\mathrm{d}x}(x)\right) + q(x)y(x) = \lambda\, r(x)y(x) \quad \text{on } (a, b) \tag{331}$$

with separated boundary conditions which commonly have the form

$$A_1 y(a) + A_2 p(a)y'(a) = 0 \qquad B_1 y(b) + B_2 p(b)y'(b) = 0 \tag{332}$$

for given constants A_i, B_i and functions $p(x)$, $q(x)$ and $r(x)$. Solving this problem means, of course, determining the values λ_n of λ for which Eq. (331) has a nontrivial (continuously differentiable square integrable) solution $y_n(x)$ satisfying Eq. (332) [187,188].

These and other higher order Sturm–Liouville problems can be recast as a linear matrix system of the form

$$Y' = (\lambda B + C(x))Y \tag{333}$$

by transforming to the so-called *compound matrix* or modified Riccati variables [189,185]. Here B is a constant matrix. When generalizing the above treatment based on the Magnus expansion to this problem, there is one elementary but important remark worth stating explicitly: *unless the differential equation (333) has the same large λ-asymptotics as some differential equation with x-independent coefficients, then it will be impossible to develop a Magnus method which accurately approximates its solutions for large λ* [185]. The reason is that a Magnus method approximates the solution by a discrete solution, calculated using a formula of the form $Y(x_{n+1}) = \exp(\sigma_n(\lambda))Y(x_n)$; in particular, on the first step (x_0, x_1), the differential equation is approximated by one in which the coefficient matrix is replaced by the x-independent matrix $\sigma_0(\lambda)/(x_1 - x_0)$.

In consequence, the attention should be restricted to systems for which it is known that a suitable constant-coefficient system provides the correct asymptotics. This is the case, in particular, for Eq. (328), and more generally for linear equations of order $2n$ in which the $(2n-1)$st derivative is zero, such as

$$(-1)^n y^{(2n)} + \sum_{j=0}^{2n-2} q_j(x)y^{(j)} = \lambda y.$$

Here the asymptotics are determined by the equation $(-1)^n y^{(2n)} = \lambda y$ [190]. Even then, the methods developed in [70] for Eq. (328) and implemented for systems with matrices of general size in [185] require a λ-dependent step size restriction of the form $h \leq \mathcal{O}(|\lambda|^{-1/4})$ in order to be defined. Nevertheless, the analysis carried out in [185] shows that the fourth order Magnus integrator based on a two-point Gaussian quadrature appears to offer significant advantages over conventional methods based on power series and library routines.

Magnus integrators have also been successfully applied in the somewhat related problem of computing the Evans function for spectral problems arising in the analysis of the linear stability of traveling wave solutions to reaction–diffusion PDEs [169]. In this setting, Magnus integrators possess some appealing features in comparison, for instance, with Runge–Kutta schemes: (1) they are unconditionally stable; (2) their performance is superior in highly oscillatory regimes and (3) their step size can be controlled in advance. Items (2) and (3) are due to the fact that error bounds for Magnus methods depend only on low order derivatives of the coefficient matrix, not (as for Runge–Kutta schemes) on derivatives of the solution. Therefore, performance and, correspondingly, the choice of optimal step size remain uniform over any bounded region of parameter space [169].

6.3. The differential Riccati equation

Let us now consider the two-point boundary value problem in the t variable defined by the linear differential equation

$$\mathbf{y}' \equiv \begin{pmatrix}\mathbf{y}_1' \\ \mathbf{y}_2'\end{pmatrix} = \begin{pmatrix}A(t) & B(t) \\ C(t) & D(t)\end{pmatrix}\begin{pmatrix}\mathbf{y}_1 \\ \mathbf{y}_2\end{pmatrix}, \quad 0 < t < T \tag{334}$$

with separated boundary conditions

$$(K_{11}\ \ K_{12})\begin{pmatrix}\mathbf{y}_1 \\ \mathbf{y}_2\end{pmatrix}_{t=0} = \gamma_1, \qquad (K_{21}\ \ K_{22})\begin{pmatrix}\mathbf{y}_1 \\ \mathbf{y}_2\end{pmatrix}_{t=T} = \gamma_2. \tag{335}$$

Here $A \in \mathbb{C}^{q\times q}$, $B \in \mathbb{C}^{q\times p}$, $C \in \mathbb{C}^{p\times q}$, $D \in \mathbb{C}^{p\times p}$, whereas $\mathbf{y}_1, \gamma_2 \in \mathbb{C}^p$, $\mathbf{y}_2, \gamma_1 \in \mathbb{C}^q$ and the matrices K_{ij} have appropriate dimensions. We next introduce the time-dependent change of variables (or picture) $\mathbf{y} = Y_0(t)\mathbf{w}$, with

$$Y_0(t) = \begin{pmatrix}I_p & 0 \\ X(t) & I_q\end{pmatrix} \tag{336}$$

and choose the matrix $X \in \mathbb{C}^{p\times q}$ so as to ensure that in the new variables $\mathbf{w} = Y_0^{-1}(t)\mathbf{y}$ the system assumes the partly decoupled structure [191]

$$\mathbf{w}' \equiv \begin{pmatrix}\mathbf{w}_1' \\ \mathbf{w}_2'\end{pmatrix} = \begin{pmatrix}A + BX & B \\ O & D - XB\end{pmatrix}\begin{pmatrix}\mathbf{w}_1 \\ \mathbf{w}_2\end{pmatrix}, \tag{337}$$

together with the corresponding boundary conditions for $\mathbf{w}$. It turns out that this is possible if, and only if, $X(t)$ satisfies the so-called differential Riccati equation [192]

$$X' = C(t) + D(t)X - XA(t) - XB(t)X, \qquad X(0) = X_0 \tag{338}$$

for some X_0. By requiring

$$X_0 = -K_{12}^{-1}K_{11}, \tag{339}$$

then the boundary conditions (335) also decouple as

$$(O \;\; K_{12})\mathbf{w}(0) = \gamma_1, \qquad (K_{21} + K_{22}X(T) \quad K_{22})\,\mathbf{w}(T) = \gamma_2. \tag{340}$$

Here we assume, without loss of generality, that K_{12} is invertible. In this way, the original boundary value problem can be solved as follows [191,192]: (i) solve Eq. (338) with initial condition (339) from $t = 0$ to $t = T$; (ii) solve the $\mathbf{w}_2$-equation in (337) and (340), also from zero to T; (iii) solve the $\mathbf{w}_1$-equation in (337) from $t = T$ to $t = 0$ and recover $\mathbf{y} = Y_0(t)\mathbf{w}$. In other words, the solution of the original two-point boundary value problem can be obtained by solving a sequence of three different initial value problems, one of which involves the nonlinear equation (338). Obviously, steps (i) and (iii) can be solved using numerical integrators based on the Magnus expansion. It could be perhaps more surprising that these algorithms can, indeed, be used to integrate the Riccati equation in step (ii).

Although the boundary value problem (334) is a convenient way to introduce the differential Riccati equation (338), this equation arises in many fields of science and engineering, such as linear quadratic optimal control, stability theory, stochastic control, differential games etc. Accordingly, it has received considerable attention in the literature, both focused on its theoretical aspects [193,194] and its numerical treatment [191,195,196].

In order to apply Magnus methods to numerically solve the Riccati equation, we first apply the transformation

$$X(t) = V(t)W^{-1}(t), \tag{341}$$

with $V \in \mathbb{C}^{p\times q}$, $W \in \mathbb{C}^{q\times q}$ and $V(0) = X_0$, $W(0) = I_q$, in the region where $W(t)$ is invertible. Then Eq. (338) is equivalent to the linear system

$$Y' = S(t)\,Y(t), \qquad Y(0) = \begin{pmatrix} I_q \\ X_0 \end{pmatrix} \tag{342}$$

with

$$Y(t) = \begin{pmatrix} W(t) \\ V(t) \end{pmatrix}, \qquad S(t) = \begin{pmatrix} A(t) & B(t) \\ C(t) & D(t) \end{pmatrix} \tag{343}$$

so that the previous Magnus integrators for linear problems can be applied here. Apparently, this system is similar to (334), but now we are dealing with an initial value problem and Y is a matrix instead of a vector.

When dealing in general with the differential Riccati equation (338), it is meaningful to distinguish the following three cases:

(i) The so-called symmetric Riccati equation, which corresponds to $q = p$, $D(t) = -A(t)^{\mathrm{T}}$ real, and $B(t)$, $C(t)$ real and symmetric matrices. In this case, the solution satisfies $X^{\mathrm{T}} = X$. It is straightforward to show that this problem is equivalent to the treatment of the generalized time-dependent harmonic oscillator, described by the Hamiltonian function

$$H = \frac{1}{2}\mathbf{p}^{\mathrm{T}}B(t)\mathbf{p} + \mathbf{p}^{\mathrm{T}}A(t)\mathbf{q} - \frac{1}{2}\mathbf{q}^{\mathrm{T}}C(t)\mathbf{q}.$$

The approximate solution attained by Magnus integrators, when applied to (342)–(343), can be seen as the exact solution corresponding to a perturbed symplectic matrix $\tilde{S}(t) \simeq S(t)$. In other words, we are solving, exactly, a perturbed Hamiltonian system so that the approximate solution, $\tilde{X}$, will share several properties of the exact solution, in particular $\tilde{X}^{\mathrm{T}} = \tilde{X}$.

(ii) The linear non-homogeneous problem

$$X' = D(t)X + C(t) \tag{344}$$

corresponds to the particular case $A = 0$ and $B = 0$ in (338).

(iii) The problem

$$X' = D(t)X + XA(t) \tag{345}$$

is recovered from (338) by taking $C = 0$ and $B = 0$. It has been treated in [36] by developing an *ad hoc* Magnus-type expansion. Notice that the case $p = q$, $D = -A$ corresponds to the linear isospectral system (238).

6.4. Stochastic differential equations

In recent years, the use of stochastic differential equations (SDEs) has become widespread in the simulation of random phenomena appearing in physics, engineering, economics etc, such as turbulent diffusion, polymer dynamics and investment finance [197]. Although models based on SDEs can offer a more realistic representation of the system than ordinary differential equations, the design of effective numerical schemes for solving SDEs is, in comparison with ODEs, a less developed field of research. This fact notwithstanding, it is true that, recently, new classes of integration methods have been constructed which automatically incorporate conservation properties the SDE possesses. Since some of the methods are based precisely on the Magnus expansion, we briefly review their main features here, and refer the reader to the more advanced literature on the subject [197–199].

A SDE in its general form is usually written as

$$\mathrm{d}y(t) = g_0(t, y(t))\mathrm{d}t + \sum_{j=1}^{d} g_j(t, y(t))\mathrm{d}W_j(t), \qquad y(0) = y_0, \quad y \in \mathbb{R}^m, \tag{346}$$

where g_j, ($j \geq 0$), are m-vector-valued functions. The function g_0 is the deterministic continuous component (called the *drift coefficient*), the g_j, ($j \geq 1$), represent the stochastic continuous components (the *diffusion coefficients*) and W_j are d independent Wiener processes. A Wiener process W (also called Brownian motion) is a stochastic process [197] satisfying

$$W(0) = 0, \qquad E[W(t)] = 0, \qquad \mathrm{Var}[W(t) - W(s)] = t - s, \quad t > s$$

which has independent increments on non-overlapping intervals. In other words, a Wiener process is normally distributed with mean or expectation value E equal to zero and variance t.

Eq. (346) can be written in integral form as

$$y(t) = y_0 + \int_0^t g_0(s, y(s))\mathrm{d}s + \sum_{j=1}^{d} \int_0^t g_j(s, y(s))\mathrm{d}W_j(s). \tag{347}$$

The d integrals in (347) cannot be considered as Riemann–Stieltjes integrals, since the sample paths of a Wiener process are not of bounded variation. In fact, if different choices are made for the point τ_i (in the subintervals $[t_{i-1}, t_i]$ of a given partition) where the function is evaluated, then the approximating sums for each g_j,

$$\sum_{i=1}^{N} g_j(\tau_i, y(\tau_i))(W_j(t_i) - W_j(t_{i-1})), \quad \tau_i = \theta t_i + (1-\theta)t_{i-1}, \tag{348}$$

converge (in the mean-square sense) to different values of the integral, depending on the value of θ [35]. Thus, for instance,

$$\int_a^b W(t)\mathrm{d}W(t) = \frac{1}{2}(W^2(b) - W^2(a)) + \left(\theta - \frac{1}{2}\right)(b-a).$$

If $\theta = 0$, then $\tau_i = t_{i-1}$ (the left-hand point of each subinterval) and the resulting integral is called an Itô integral; if $\theta = 1/2$ (so that the midpoint is used instead), one has a Stratonovich integral. These are the two main choices and, although they are related, the particular election depends ultimately on the nature of the process to be modeled [35]. It can be shown that the Stratonovich calculus satisfies the Riemann–Stieltjes rules of calculus, and thus it is the natural choice here.

When dealing with numerical methods for solving (346), there are two ways of measuring accuracy [197]. The first is *strong convergence*, essential when the aim is to get numerical approximations to the trajectories which are close to the exact solution. The second is *weak convergence*, when only certain moments of the solution are of interest. Thus, if $\hat{y}_n$ denotes the numerical approximation to $y(t_n)$ after n steps with constant step size $h = (t_n - t_0)/n$, then the numerical solution $\hat{y}$ converges strongly to the exact solution y with strong global order p if there exist $C > 0$ (independent of h) and $\delta > 0$ such that

$$E[\|\hat{y}_n - y(t_n)\|] \leq Ch^p, \quad h \in (0, \delta).$$

It is worth noticing that p can be fractional, since the root mean-square order of the Wiener process is $h^{1/2}$. One of the simplest procedures for solving (346) numerically is the so-called Euler–Maruyama method [200],

$$y_{n+1} = y_n + \sum_{j=0}^{d} J_j g_j(t_n, y_n), \tag{349}$$

where

$$h = t_{n+1} - t_n, \quad J_0 = h, \qquad J_j = W_j(t_{n+1}) - W_j(t_n), \quad j = 1, \ldots, d.$$

This scheme turns out to be of strong order 1/2. Here the J_j can be computed as $\sqrt{h}N_j$, where the N_j are $N(0, 1)$ normally distributed independent random variables [35].

For the general non-autonomous linear Stratonovich problem defined by

$$\mathrm{d}y = G_0(t)y\,\mathrm{d}t + \sum_{j=1}^{d} G_j(t)y\mathrm{d}W_j, \qquad y(0) = y_0 \in \mathbb{R}^m \tag{350}$$

the Magnus expansion for the deterministic case can be extended in quite a straightforward way. It is well worth noting that in Eq. (350), even when the functions G_j are constant, there is no explicit solution [198], unless all the G_j, $j \geq 0$, commute with one another, in which case it holds that

$$y(t) = \exp\left(G_0\, t + \sum_{j=1}^{d} G_j W_j(t)\right) y_0. \tag{351}$$

In many modeling situations, however, there is no reason to expect that the functions G_j associated with the Wiener processes commute. If, for simplicity, we only consider the autonomous case and write

$$G(t) \equiv G_0\,\mathrm{d}t + \sum_{j=1}^{d} G_j \mathrm{d}W_j(t),$$

then (350) can be expressed as

$$\mathrm{d}y = G(t)y\,\mathrm{d}t, \qquad y(0) = y_0$$

and thus one can formally apply the Magnus expansion to this equation to get $y(t) = \exp(\Omega(t))y_0$. The first term in the series reads, in this case,

$$\int_0^t G(s)\mathrm{d}s \equiv \int_0^t G_0 \mathrm{d}s + \sum_{j=1}^{d} \int_0^t G_j \mathrm{d}W_j(s) = G_0\, t + \sum_{j=1}^{d} G_j J_j,$$

where now

$$J_j = \int_0^t \mathrm{d}W(s) = W_j(t) - W_j(0).$$

By inserting these expressions into the recurrence associated with the Magnus series, Burrage and Burrage [35] show that

$$\begin{aligned}\Omega(t) = &\sum_{j=0}^{d} G_j J_j + \frac{1}{2}\sum_{i=0}^{d}\sum_{j=i+1}^{d} [G_i, G_j](J_{ji} - J_{ij}) \\ &+ \sum_{i=0}^{d}\sum_{k=0}^{d}\sum_{j=k+1}^{d} [G_i, [G_j, G_k]]\left(\frac{1}{3}(J_{kji} - J_{jki}) + \frac{1}{12}J_i(J_{jk} - J_{kj})\right) + \cdots,\end{aligned} \tag{352}$$

where the multiple Stratonovich integrals are defined by

$$J_{j_1 j_2 \cdots j_l}(t) = \int_0^t \int_0^{s_l} \cdots \int_0^{s_2} \mathrm{d}W_{j_1}(s_1) \cdots \mathrm{d}W_{j_l}(s_l), \quad j_i \in \{0, 1, \ldots, d\}. \tag{353}$$

Since not all the Stratonovich integrals are independent, one has to compute only $d(d+1)(d+5)/6$ stochastic integral evaluations to achieve strong order 1.5 with the expression (352) [35]. If, on the other hand, $\Omega(t)$ is truncated after the first set of terms, then the resulting numerical approximation

$$y(t) = \exp\left(\sum_{j=0}^{d} G_j J_j\right) y_0$$

has strong order 1/2, but leads to smaller error coefficients than the Euler–Maruyama method (349) [35]. Furthermore, the error becomes smaller as the $G_i G_j$ terms get closer to commuting and the scheme preserves the underlying structure of the problem.

One should notice, at this point, that Eq. (346) (or, in the linear case, Eq. (350)), has formally the same structure as the nonlinear ODE (171) appearing in control theory. Therefore, the formalism developed there to get the Chen–Fliess series can be applied here with the alphabet $I = \{0, 1, \ldots, d\}$ and the integrals

$$\left(\int_0 \mu\right)(t) = \int_0^t \mu(s)\mathrm{d}s, \qquad \left(\int_i \mu\right)(t) = \int_0^t \mu(s)\mathrm{d}W_i(s), \quad i \geq 1,$$

since the Stratonovich integrals satisfy the integration by parts rule. In other words, one can obtain the corresponding Magnus expansion for arbitrary (linear or nonlinear) stochastic differential equations simply by following the same procedure as for deterministic ODEs.

With respect to nonlinear Stratonovich stochastic differential equations, it should be remarked that the use of Lie algebraic techniques, as well as the design of Lie group methods for obtaining strong approximations when the solution evolves on a smooth manifold, has received considerable attention in the recent literature [201–203].

7. Physical applications

From previous sections it should be clear that ME has a strong bearing on both Classical and Quantum Mechanics. As far as Classical Mechanics is concerned, this has been most explicitly shown in Section 3.4. On its turn Quantum Mechanics has been repeatedly invoked as a source of applications, among others, in Sections 2.9 and 4. In this section we present in a very schematic way, and with no aim at completeness, some applications of ME in different areas of the physical sciences. This will show that, over the years, ME has been one of the preferred options to deal with Eq. (4) which, under different appearances, pervades the entire field of Physics. In the works mentioned here, almost exclusively analytical methods are used and, in general, one must recognize that in most, if not all, cases listed, only the first two orders of the expansion have been considered. In very specially simple applications, due to particular algebraic properties of the operators involved, this happens to be exact. Only with the more recent advent of numerical applications, as has been emphasized in Sections 5 and 6, has the expansion been carried in a more systematic way to higher orders.

7.1. Nuclear, atomic and molecular physics

As far as we know, the first physical application of ME dates back to 1963. Robinson [18] published a brand new formalism to investigate multiple Coulomb excitations of deformed nuclei. As a matter of fact, he states explicitly that only after completion of his work did he discover the ME. His derivation of ME formulas is certainly worth reading.

The Coulomb excitation process yields information about the low lying nuclear states. Prior to Robinson's work, the theory was essentially based on perturbation expansions, which requires that the bombarding energy is kept so low that no nuclear reaction takes place. Even worse, if heavier ions are used as projectiles, the electric field exerted on the target nucleus is so strong that perturbation methods fail.

The work by Robinson improved the so-called at that time *sudden approximation*, which is equivalent to the assumption that all nuclear energy levels are degenerate. Results are reported, in that reference, for rotational and vibrational nuclei.

As representatives of the applications of ME in the field of Atomic Physics, we mention several types of atomic collisions. The ME is used in [204] to derive the transition amplitude and the cross section for K-shell ionization of atoms by heavy-ion impact. This is an important process in heavy-ion physics. The theoretical investigation of these reactions always assumed that the projectile is a relatively light ion such as a proton or an α particle. The use of ME allowed one to extend the studies to the ionization of light target atoms by much heavier projectile ions.

In [205,206] the ME is applied to study the time-evolution of rotationally induced inner-shell excitation in atomic collisions. In this context, the internuclear motion can be treated classically, and the remaining quantum-mechanical problem for the electronic motion is then time-dependent. In particular, in [205] explicit results for $Ne^{+}Ne$ collisions are given, as well as a study of the convergence properties of ME with respect to the impact parameter.

The ME is applied in [207] to the theoretical study of electron–atom collisions, involving many channels coupled by strong, long-range forces. Then, as a test case, the theory is applied to electron-impact excitation of the resonance transitions of Li, Na and K. Computations up to second order are carried out, and the cross sections found are in good agreement with experimental data for the intermediate-energy range.

The following examples illustrate the use of ME in Molecular Physics. In [208] it is applied for the first time to the theory of the pressure broadening of rotational spectra. Unlike the previous approaches to the problem, the S-matrix obtained is unitary. As a consequence of it, the relative contributions on the linewidth of the attractive and repulsive anisotropy terms in the interaction potential may be calculated.

Floquet theory is applied in [209] to systems, periodic in time in the semiclassical approximation of the radiation–quantum-molecule interaction in an intense field. The paper contains an interesting discussion about the appropriateness of the Schrödinger and Interaction pictures. One and two-photon probability transitions are obtained up to second order in ME. Noteworthy, formulas through fifth order in ME are given, in a less symmetrical form.

In [23] the applicability of ME to the multiphoton excitation of a sparse level system for which the rotating wave function approximation is not applicable, is explored. This reference provides a method of treating the time-evolution of a pumped molecular system in the low energy region, which is characterized by a sparse distribution of bound vibrational states.

7.2. Nuclear magnetic resonance: Average Hamiltonian theory

This is certainly the field where ME has been most systematically used and so we consider it separately. From elementary quantum mechanics, it is known that a constant magnetic field breaks the degeneracy of the energy levels of an atomic nucleus with spin. If the nuclear spin is s then $2s + 1$ sub-levels appear. In a sample, these states are occupied, according to Boltzman distribution, with an exponentially distributed population. When a time-dependent radio-frequency electromagnetic field of appropriate frequency is applied, energy can be absorbed by certain nuclei, which are consequently promoted to higher levels. This is the physical phenomenon of Nuclear Magnetic Resonance (NMR).

It was Evans [210] and Haeberlen and Waugh [211] who first applied the ME to NMR. Since that time, the ME has been instrumental in the development of improved techniques in NMR spectroscopy [56].

The major advantage of NMR is the possibility of modifying the nuclear spin Hamiltonian almost at will, and to adapt it to the needs of the problem to be solved [24]. This manipulation requires an external perturbation of the system that can

be either time-independent (changes of temperature, pressure, solvents, etc.) or time-dependent (sample spinning, pulsed radio-frequency fields). In the later context, the concept of average Hamiltonian provides an elegant description of the effects of a time-dependent perturbation applied to the system. It was originally introduced in NMR by Waugh [24,25] to explain the effects of multiple-pulse sequences.

The basic idea of average Hamiltonian theory, for a system governed by $H(t)$, consists of describing the effective evolution within a fixed time interval by an average Hamiltonian $\overline{H}$. The theory states that this is always possible provided $H(t)$ is periodic. The average Hamiltonian depends, however, on the beginning and the end of the time interval observed. It is precisely the average Hamiltonian $\overline{H}$ which is obtained by means of the ME.

When the total Hamiltonian splits in a time-independent and a time-dependent piece, $H(t) = H_0 + H_1(t)$, with $H_1(t)$ periodic, an interesting new picture is used, labeled *toggling frame*. It certainly recalls the Interaction Picture defined in Eq. (91) but is rather different. In (89) the operator $G(t)$ associated to the toggling frame is given by the time-ordered expression

$$G(t) = \mathcal{T}\left(\exp \int_0^t \tilde{H}_1(s)\mathrm{d}s\right) \tag{354}$$

and the key point here is whether the formal time-ordering is solvable.

As already mentioned, the interplay between NMR and ME has been fruitful along the years and acted in both directions. To prove that it is still alive, we quote two recent papers directly dealing with that mutual interaction. In [212] the relevance of ME through NMR for the new field of quantum information processing and computing is envisaged. The authors of [213] have recently explored the fourth and sixth order of ME to design a software package for the simulation of NMR experiments. Although their results are not yet conclusive their work shows the vitality of the ME.

7.3. Quantum field theory and high energy physics

The starting point of any quantum field theory (QFT) calculation is again Eq. (4) which is conventionally treated by time-dependent perturbation theory. So the first question which arises is the connection between ME and Dyson-type series. This has already been dealt with in Section 2.4. The main advantage of the first one is, as has already been repeatedly pointed out, that the unitary character of the evolution operator is preserved at all orders of approximation. In the historical development of QFT it was, however, the Dyson approach which was followed. The lost of unitarity was not thought to be of great relevance, considering the problems presented by the infinities appearing all over the place. Once the renormalization idea was introduced, this awful aspect of the theory was also put under control. The results were, from the point of view of the calculation of observable magnitudes, an unprecedented success: the agreement between experimental results and their theoretical counterparts was impressive.

So no wonder that alternatives to Dyson series, such as ME, did not see popular acceptance. However, during the years, there have been interesting developments involving ME in the context of field theory. In particular, its use has been shown to imply a re-ordering of terms in the calculations in such a way that some infinities do not appear, and so make the introduction of counterterms in the Hamiltonian unnecessary. This is what happens, for example, in [214] where models are built in which ultraviolet divergence appears neither in the Hamiltonian nor in the S-matrix. In principle, the results are valid for relativistic field theories with any particle content, and with minimal assumptions about the form of the interaction.

ME as an alternative to conventional perturbation theory for quantum fields, has also been studied in [215] where normal products, Wick theorem and the like, are used to deduce graphical rules *à la* Feynman for the terms Ω_i for any value of i. This has proved helpful in the treatment of infrared divergences for some QED processes such as the scattering of an electron on an external potential, or the bremsstrahlung of one hard photon, both cases accompanied by the emission of an arbitrary number of soft photons [216]. An interesting feature of the ME based approach is that the theory is free from infrared and mass divergences, as a consequence of the unitary character of the approximate time-evolution operator [26, 216]. The method is simpler than previous techniques based on re-summation of the perturbation series, to get rid of those divergences. Furthermore, in contrast with the usual treatment, the resolution of the detector is not an infrared regularization parameter. An application to Bhabha scattering (elastic electron–positron scattering) is developed in [217]. The difficulties of extending the results to Quantum Chromodynamics are commented on in [215].

Recently, an extension of the Magnus expansion has also been used in the context of Connes–Kreimer's Hopf algebra approach to perturbative renormalization of quantum field theory [218,219]. In particular, in [220], it is shown that this generalized ME allows one to solve the Bogoliubov–Atkinson recursion in this setting.

In the field of high energy physics, ME has also found applications. Next we just quote two instances: one referring to heavy ion collisions, and the other to elementary particle physics.

In collision problems, the unitarity of the time evolution operator imposes some bound on the experimentally observable cross sections. When these magnitudes are theoretically calculated, one usually keeps only the lowest orders in conventional perturbation theory. This may be harmless at relatively low energies, but it may lead to unitarity bounds violation as the energy increases. The use of a manifestly unitary approximation scheme is then necessary. ME provides such a scheme. In heavy ion collision at sufficiently high energy and given kinematic configuration (small impact parameter), that violation is produced for e^+e^- when analyzed in the lowest-order time-dependent perturbation theory. In [221] a remedy for this situation was advanced by the use of first order ME. It is discussed how most theoretical approaches are based on either lowest-order time-dependent perturbation theory or on the Fermi–Weizsäker–Williams method of virtual photons. These

approaches violate unitarity bounds for sufficiently high collision energies, and thus the probability for single-pair creation exceeds unity. With some additional assumptions, a restricted class of diagrams associated with electron–positron loops can be summed to infinite order in the external charge. The electron–positron transitions amplitudes and production probabilities obtained, are manifestly unitary and gauge invariant.

In recent years, there has been a great interest in neutrino oscillations and its closely related solar neutrino problem. The known three families of neutrinos with different flavors (electron, muon and tau) were experimentally shown to be able to convert into each other. The experiments were carried out with neutrinos of different origins: solar, atmospheric, produced in nuclear reactors and in particle accelerators. Here, oscillation means that neutrinos of a given flavor can, after propagation, change their flavor. The accepted explanation for this phenomenon is that neutrinos with a definite flavor do not have a definite mass, and vice-versa. Let us denote by $|\nu_\alpha\rangle$ the neutrinos of definite flavor with α the flavor index (i.e., electron, muon, tau) and by $|\nu_i\rangle$ the neutrinos with well defined distinct masses m_i, $i = 1, 2, 3$. Then the previous assertion means that $|\nu_\alpha\rangle$ will be a linear combination of the different $|\nu_i\rangle$.

As neutrinos with different masses propagate with different velocities, this mixing allows for flavor conversion, i.e. for neutrinos oscillations. ME enters the game in the solution of the evolution operator in one basis. If one neutrino "decouples" from the other two, then the problem reduces to one with only two effective generations. Mathematically, it is similar to the two level system studied in Section 3. The reader is referred to [27,222–224] for details of the calculations for two and three generations.

7.4. Electromagnetism

The Maxwell equations govern the evolution of electromagnetic waves. The equations are linear in the usual form and, although they have been extensively studied, the complexity of obtaining approximate solutions is rather significant. When reformulating the equations for a given problem, where the geometry, the boundary conditions etc, are considered, and appropriate discretizations are taken into account, it is common to end with linear non-autonomous equations, so the Magnus expansion can be of interest here.

To illustrate some possible applications, let us consider the Maxwell equations

$$\begin{cases} \dfrac{\partial \mathbf{H}}{\partial t} = -\dfrac{1}{\mu}\,\nabla \times \mathbf{E} \\ \dfrac{\partial \mathbf{E}}{\partial t} = \dfrac{1}{\varepsilon}\,\nabla \times \mathbf{H} - \dfrac{1}{\varepsilon}\mathbf{J}(t) \end{cases} \tag{355}$$

where $\mathbf{H}$, $\mathbf{E}$, $\mathbf{J}$ are the magnetic and electric field intensities, and the current density, respectively, μ is the permeability and ε is the permittivity. After space discretization, these equations turn into a large linear system of non-homogeneous equations and Magnus integrators can, in principle, be applied. Time dependent contributions can also appear from boundary conditions or external interactions. In some cases $\mathbf{J} = \sigma\,\mathbf{E}$ [225,226], so that, if the conductivity σ is not constant, Magnus integrators can be useful.

Let us now consider the frequency domain Maxwell equations (with $\mathbf{J} = \mathbf{0}$)

$$\begin{cases} \nabla \times \mathbf{E} = \mathrm{i}w\mu\mathbf{H} \\ \nabla \times \mathbf{H} = -\mathrm{i}w\varepsilon\mathbf{E} \end{cases} \tag{356}$$

where w is the angular frequency. These equations are of interest for time-harmonic light waves, propagating in a wave-guiding structure composed of linear isotropic materials. If one is interested in the x and y components of $\mathbf{H}$ and $\mathbf{E}$, and how they propagate in the z direction, the equations to solve, after appropriate discretization, take the form [227]

$$-\mathrm{i}w\varepsilon\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}z} = A(z)\mathbf{v}, \qquad -\mathrm{i}w\mu\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}z} = B(z)\mathbf{u}. \tag{357}$$

Here $\mathbf{u}$, $\mathbf{v}$ are vectors and A, B matrices depending on z, which, in this case, play the role of evolution parameter. A fourth-order Magnus integrator has been used in [227]. From Section 5 we observe that higher order Magnus integrators, combined with splitting methods, could also lead to efficient algorithms to obtain accurate numerical results with preserved qualitative properties of the exact solution.

7.5. Optics

In the review paper [228] one can find references to some early applications of ME to Optics. For example, to Hamiltonians involving the generators of SU(2), SU(1, 1) and Heisenberg–Weyl groups, with applications to laser-plasma scattering and pulse propagation in free-electron lasers. Here, as representatives of the more modern interest of ME in Optics, we quote two applications referring to Helmholtz equation, and to the study of Stokes parameters.

Helmholtz equation in one spatial dimension with a variable refractive index $n(x)$ reads

$$\psi''(x) + k^2 n^2(x)\psi(x) = 0, \tag{358}$$

where k is the wavenumber in vacuum.

Recently, this time-honored wave equation has been treated in two different ways, both using ME. From a more formal point of view, in [229] Helmholtz equation is analyzed following the well known procedure followed by Feshbach and Villars to convert the second order relativistic quantum Klein–Gordon differential equation for spin-0 particles in a first order differential equation involving two components wave functions (the original wave function and its time derivative). The evolution operator for Helmholtz equation is then a 2×2 matrix which evolves according to the fundamental equation (4), with the only difference that now the evolution parameter is x instead of t. In [229] the whole procedure is explained and the main physical consequence, which amounts to the addition of correcting terms to the Hamiltonian, is discussed in the case of an axially symmetric graded-index medium, i.e. one in which the refractive index is a polynomial.

Helmholtz equation has also been investigated with the help of ME in [227,230,231]. Here, the propagation in a slowly varying waveguide is considered, and the boundary value problem is converted into an initial value problem by the introduction of appropriate operators which are shown to satisfy Eq. (4). Numerical methods to fourth order borrowed from [37] are then used.

Since the mid 19th century, the polarization state of light, and in general electromagnetic radiation or any other transverse waves, is described by the so-called Stokes parameters, which constitute a four-dimensional vector $\mathbf{S}(\omega)$ depending on the frequency ω. When the light traverses an optical element which acts on its polarization state, the in and out Stokes vectors are related by

$$\mathbf{S}_{\text{out}}(\omega) = M(\omega)\mathbf{S}_{\text{in}}(\omega), \tag{359}$$

where the 4×4 matrix $M(\omega)$ is called the Mueller matrix. It can be proved [232,233] that it satisfies the equation

$$M'(\omega) = H(\omega)M(\omega), \qquad M(\omega_0) = M_0, \tag{360}$$

where, now, the prime denotes derivative with respect to the real independent variable ω. For systems with zero polarization-dependent loss (PDL) and no polarization mode dispersion (PMD), $H(\omega)$ is constant, whereas with PDL and PDM, the previous equation is just our Eq. (4) and the appropriateness of ME is apparent. The matrix $H(\omega)$ in this application has a Hermitian and a non-Hermitian component. ME has allowed a recursive calculation of successive orders of the frequency variation of the Mueller matrix. This yields PMD and PDL compensators that counteract the effects of PMD and PDL with increased accuracy.

Also related to the use of Stokes vector, one can mention the so-called radiative transfer equation for polarized light. It is relevant in Astrophysics to measure the magnetic fields in the Sun and stars. That equation gives the variation of $\mathbf{S}(z)$ with the light path z

$$\frac{\mathrm{d}}{\mathrm{d}z}\mathbf{S}(z) = -K(z)\mathbf{S}(z) + \mathbf{J},$$

where $\mathbf{S}$ is the Stokes vector, K is a 4×4 matrix which describes absorption in the presence of Zeeman effect and $\mathbf{J}$ stands for the emission term. In [234–236] ME is used to obtain an exponential solution.

7.6. General Relativity

To illustrate, once more, the pervasive presence of the linear differential equation (4), let us mention reference [237] in which the aim is to determine the time elapsed between two events when the space-time is treated as in General Relativity. Then it turns out to be necessary to solve a two-point boundary value problem for null geodesics. In so doing, one needs to know a Jacobian whose expression involves a 8×8 matrix function obeying the basic Eq. (4). In [237] an eighth order numerical method from [39] is used, which is proved to be an efficient scheme.

7.7. Search of periodic orbits

The search of periodic orbits for some non-linear differential autonomous equations, $\mathbf{x}' = \mathbf{f}(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$ is of interest in Celestial Mechanics (periodic orbits of the N-body problem) as well as in the general theory of dynamical systems. Due to the complexity of this process, it is important to have efficient numerical algorithms.

The Lindstedt–Poincaré technique is frequently used to calculate periodic orbits. An iterative process proposed in [238] consists of starting with a guessed periodic orbit, and this guess is subsequently improved by solving a correlation non-autonomous linear differential equation. The numerical integration of this equation is carried out by means of Magnus integrators.

7.8. Geometric control of mechanical systems

Many mechanical systems studied in control theory can be modeled by an ordinary differential equation of the form [90,239]

$$\mathbf{x}'(t) = \mathbf{f}_0(\mathbf{x}(t)) + \sum_{i=1}^{m} u_i(t)\ \mathbf{f}_i(\mathbf{x}(t)), \tag{361}$$

initialized at $\mathbf{x}(0) = \mathbf{p}$. Here $\mathbf{x} \in \mathbb{R}^d$ represents all the possible states of the system, $\mathbf{f}_i$ are (real) analytic vector fields and the function $\mathbf{u} = (u_1, \ldots, u_m)$ (the *controls*) is assumed to be integrable with respect to time and taking values in a compact subset $U \subset \mathbb{R}^m$. The vector field $\mathbf{f}_0$ is called the *drift* vector field, whereas $\mathbf{f}_i$, $i \geq 1$, are referred to as the *control* vector fields. When $\mathbf{f}_0 \equiv 0$, the system (361) is called 'without drift', and its analysis is typically easier.

For a given set of controls $\{u_i\}$, Eq. (361) with initial value $\mathbf{x}(0)$ is nothing but a dynamical system, which can be analyzed and (approximately) solved by standard techniques. In control theory, however, one is interested typically in the *inverse problem*: given a target $\mathbf{x}(T)$, find controls $\{u_i\}$ that steer from $\mathbf{x}(0)$ to $\mathbf{x}(T)$ [90], perhaps by following a prescribed path. Just to illustrate these abstract considerations, a typical problem could be the determination of a set of controls that drive the actions of a robot during a task.

The first step is to guarantee that there exists a solution. This is the problem of controllability. To characterize the controllability of linear systems of the form

$$\mathbf{x}'(t) = A(t)\mathbf{x}(t) + B(t)\mathbf{u}(t) \tag{362}$$

is a relatively simple task thanks to an algebraic criterion known as the Kalman rank condition [239,240]. This issue, however, is much more involved for the nonlinear system (361) [90].

The interest of ME in control theory, as has been already discussed in Section 3.5, stems from the approximate ansatz it provides connecting the states $\mathbf{x}(0)$ and $\mathbf{x}(T)$. Thus, the Magnus expansion can be used either to predict a state $\mathbf{x}(T)$ for a given control $\mathbf{u}$ or to find reasonable controls $\mathbf{u}$ which made reachable, the target $\mathbf{x}(T)$ from $\mathbf{x}(0)$. Of course, many sets of controls may exist and it raises questions concerning the *cost* of every scheme, and consequently the search for the optimal choice. For instance, the ME has been used in non-holonomic motion planning of systems without drift [241,242]. Among non-holonomic systems there are free-floating robots, mobile robots and underwater vehicles [242,243].

In the particular case of linear quadratic optimal control problems (appearing in engineering problems as well as in differential games) a given cost functional has to achieve a minimum. When this happens, Eq. (362) can be written as [244,194] is

$$\mathbf{x}' = M(t, K(t))\mathbf{x}, \tag{363}$$

where $K(t) \in \mathbb{R}^{d\times d}$ has to solve a Riccati differential equation similar to (338) with final condition $K(T) = K_f$. In other words, the Riccati equation has to be integrated backward in time, and then to use it as an input in (363). As mentioned, the Riccati differential matrix equation has received much attention [191,194–196,244–246], but an efficient implementation to this problem requires further investigation, and methods from ME can play an important role.

8. Conclusions

In this report, we have thoroughly reviewed the abiding work on Magnus expansion carried out over more than fifty years, from very different perspectives.

As a result of a real interdisciplinary activity, some aspects of the original formulation have been refined. This applies, for example, to the convergence properties of the expansion which have been much sharpened.

In other features, much practical progress has been made. This is the case of the calculation of the terms of the series, both explicitly and recurrently. New techniques, like the ones borrowed from graph theory, have also profitably entered the play.

Although originally formulated for linear systems of ordinary differential equations, the domain of usage of ME has enlarged to include other types of problems with differential equations: stochastic equations, nonlinear equations or Sturm-Liouville problems.

In parallel with these developments in the mathematical structure of the ME, the realm of its applications has also widened over the years. It is worth stressing, in this respect, the versatility of the expansion to cope with new applications in old fields, like NMR for instance, and at the same time its capability to generate new contributions, like the generation of efficient numerical algorithms for geometric integrators.

All these facts, historical and current, presented and discussed in this report, strongly support the idea that ME can be a very useful tool for physicists.

Acknowledgments

One of us (J.A.O) was introduced to the field of Magnus expansion by Prof. Silvio Klarsfeld (Orsay) to whom we also acknowledge his continuous interest in our work. Encouraging comments by Prof. A.J. Dragt (Maryland) in early stages of our involvement with ME are deeply acknowledged. During the last decade, we have benefited from enlightening discussions with, and useful comments by, Prof. Arieh Iserles. We would like to thank him for that, as much as for the kind hospitality extended to us during our visits to his group in DAMTP in Cambridge. We also thank Prof. Iserles for providing us with his Latex macros for the graphs in Section 2.5, and Dr. Ander Murua for reading various parts of the manuscript and providing useful feedback. Friendly collaboration with Dr. Per Christian Moan is also acknowledged.

The work of SB and FC has been partially supported by Ministerio de Ciencia e Innovación (Spain) under project MTM2007-61572 (co-financed by the ERDF of the European Union), that of JAO and JR by contracts MCyT/FEDER, Spain (Grant No. FIS2004-0912) and Generalitat Valenciana, Spain (Grant No. ACOMP07/03). SB also acknowledges the support of the UPV through the project 20070307.

References

[1] W. Abikoff, J.S. Birman, K. Kuiken (Eds.), The Mathematical Legacy of Wilhelm Magnus. Groups, Geometry and Special Functions, American Mathematical Society, 1994.
[2] W. Magnus, A. Karrass, D. Solitar, Combinatorial Group Theory, Dover, 2004.
[3] A. Erdelyi, W. Magnus, F. Oberhettinger, F.G. Tricomi, Higher Transcendental Functions, Vols. I, II, III, McGraw Hill, 1953.
[4] W. Magnus, On the exponential solution of differential equations for a linear operator, Commun. Pure Appl. Math. VII (1954) 649–673.
[5] K.O. Friedrichs, Mathematical aspects of the quantum theory of fields, part V, Commun. Pure Appl. Math. VI (1953) 1–72.
[6] R.P. Feynman, An operator calculus having applications in Quantum Electrodynamics, Phys. Rev. 84 (1951) 108–128.
[7] H.F. Baker, On the integration of linear differential equations, Proc. London Math. Soc. 35 (1902) 334–378.
[8] E.L. Ince, Ordinary Differential Equations, Dover, 1956.
[9] H.F. Baker, Alternant and continuous groups, Proc. London Math. Soc. (Second series) 3 (1905) 24–47.
[10] J.E. Campbell, On a law of combination of operators, Proc. London Math. Soc. 29 (1898) 14–32.
[11] F. Hausdorff, Die symbolische Exponential Formel in der Gruppen theorie, Leipziger Ber. 58 (1906) 19–48.
[12] E.B. Dynkin, Evaluation of the coefficients of the Campbell–Hausdorff formula, Dokl. Akad. Nauk. SSSR 57 (1947) 323–326.
[13] F.J. Dyson, The radiation theorem of Tomonaga, Swinger and Feynman, Phys. Rev. 75 (1949) 486–502.
[14] W. Magnus, Algebraic aspects in the theory of systems of linear differential equations, Technical Report BR-3, New York University, June 1953.
[15] P. Pechukas, J.C. Light, On the exponential form of time-displacement operators in quantum mechanics, J. Chem. Phys. 44 (1966) 3897–3912.
[16] J. Wei, Note on global validity of the Baker–Hausdorff and Magnus theorems, J. Math. Phys. 4 (1963) 1337–1341.
[17] J. Wei, E. Norman, Lie algebraic solution of linear differential equations, J. Math. Phys. 4 (1963) 575–581.
[18] D.W. Robinson, Multiple Coulomb excitations of deformed nuclei, Helv. Phys. Acta 36 (1963) 140–154.
[19] R.M. Wilcox, Exponential operators and parameter differentiation in quantum physics, J. Math. Phys. 8 (1967) 962–982.
[20] F. Fer, Résolution de l'equation matricielle $\dot{U} = pU$ par produit infini d'exponentielles matricielles, Bull. Cl. Sci. Acad. R. Belg. 44 (1958) 818–829.
[21] D. Chang, J.C. Light, Exponential solution of the Schrödinger equation: Potential scattering, J. Chem. Phys. 50 (1969) 2517–2525.
[22] D. Baye, P.H. Heenen, A theoretical study of fast proton–atomic hydrogen scattering, J. Phys. B 6 (1973) 105–113.
[23] I. Schek, J. Jortner, M.L. Sage, Application of the Magnus expansion for high-order multiphoton excitation, Chem. Phys. 59 (1981) 11–27.
[24] R.R. Ernst, G. Bodenhausen, A. Wokaun, Principles of NMR in One and Two Dimensions, Clarendon, 1986.
[25] J.S. Waugh, Theory of broadband spin decoupling, J. Magn. Reson. 50 (1982) 30–49.
[26] H.D. Dahmen, B. Scholz, F. Steiner, Infrared dynamics of QED and the asymptotic behavior of the electron form factor, Nuclear Phys. B 202 (1982) 365–381.
[27] J.C. D'olivo, J.A. Oteo, Magnus expansion and the two-neutrino oscillations in matter, Phys. Rev. D 42 (1990) 256–259.
[28] S. Klarsfeld, J.A. Oteo, Analytic properties for the Magnus operator for two solvable Hamiltonians, Phys. Lett. A 142 (1989) 393–397.
[29] S. Klarsfeld, J.A. Oteo, Recursive generation of higher-order terms in the Magnus expansion, Phys. Rev. A 39 (1989) 3270–3273.
[30] S. Blanes, F. Casas, J.A. Oteo, J. Ros, Magnus and Fer expansions for matrix differential equations: The convergence problem, J. Phys. A: Math. Gen. 22 (1998) 259–268.
[31] F. Casas, Sufficient conditions for the convergence of the Magnus expansion, J. Phys. A: Math. Theor. 40 (2007) 15001–15017.
[32] P.C. Moan, J.A. Oteo, J. Ros, On the existence of the exponential solution of linear differential systems, J. Phys. A: Math. Gen. 32 (1999) 5133–5139.
[33] P.C. Moan, J.A. Oteo, Convergence of the exponential Lie series, J. Math. Phys. 42 (2001) 501–508.
[34] F. Casas, J.A. Oteo, J. Ros, Floquet theory: Exponential perturbative treatment, J. Phys. A: Math. Gen. 34 (2001) 3379–3388.
[35] K. Burrage, P.M. Burrage, High strong order methods for non-commutative-stochastic ordinary differential equation systems and the Magnus formula, Physica D 133 (1999) 34–48.
[36] A. Iserles, A Magnus expansion for the equation $Y' = AY - YB$, J. Comput. Math. 19 (2001) 15–26.
[37] A. Iserles, S.P. Nørsett, On the solution of linear differential equations in Lie groups, Phil. Trans. R. Soc. A 357 (1999) 983–1019.
[38] C.J. Budd, A. Iserles, Geometric integration: Numerical solution of differential equations on manifolds, Phil. Trans. R. Soc. A 357 (1999) 945–956.
[39] S. Blanes, F. Casas, J. Ros, Improved high order integrators based on the Magnus expansion, BIT 40 (2000) 434–450.
[40] S. Blanes, F. Casas, J. Ros, High order optimized geometric integrators for linear differential equations, BIT 42 (2002) 262–284.
[41] M. Postnikov, Lie Groups and Lie Algebras. Semester V of Lectures in Geometry, URSS Publishers, 1994.
[42] G.H. Golub, C.F. Van Loan, Matrix Computations, Johns Hopkins University Press, 1996.
[43] E.E. Tyrtyshnikov, A Brief Introduction to Numerical Analysis, Birkhäuser, 1997.
[44] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.
[45] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, Dover, 1965.
[46] W. Rossmann, Lie Groups. An Introduction Through Linear Groups, Oxford University Press, 2002.
[47] S. Sternberg, Lie algebras, Book online: www.math.harvard.edu/~shlomo/, 2004.
[48] F. Schur, Neue Begründung der Theorie der endlichen Transformationsgruppen, Math. Ann. 35 (1890) 161–197.
[49] D. Prato, P. Lamberti, A note on Magnus formula, J. Chem. Phys. 106 (1997) 4640–4643.
[50] I. Bialynicki-Birula, B. Mielnik, J. Plebański, Explicit solution of the continuous Baker–Campbell–Hausdorff problem and a new expression for the phase operator, Ann. Phys. 51 (1969) 187–200.
[51] B. Mielnik, J. Plebański, Combinatorial approach to Baker–Campbell–Hausdorff exponents, Ann. Inst. Henri Poincaré XII (1970) 215–254.
[52] L. Sáenz, R. Suárez, Combinatorial approach to the generalized Baker–Campbell–Hausdorff–Dynkin formula, Systems Control Lett. 45 (2002) 357–370.
[53] R.S. Strichartz, The Campbell–Baker–Hausdorff–Dynkin formula and solutions of differential equations, J. Funct. Anal. 72 (1987) 320–345.
[54] R. Suárez, L. Sáenz, Lie algebras associated with the exponential solutions of nonautonomous linear differential equations, J. Math. Phys. 42 (2001) 4582–4605.
[55] W.R. Salzman, New criterion for convergence of exponential perturbation theory in the Schrödinger representation, Phys. Rev. A 36 (1987) 5074–5076.
[56] D.P. Burum, Magnus expansion generator, Phys. Rev. B 24 (1981) 3684–3692.
[57] C.S. Lam, Decomposition of time-ordered products and path-ordered exponentials, J. Math. Phys. 39 (1998) 5543–5558.
[58] J.A. Oteo, J. Ros, From time-ordered products to Magnus expansion, J. Math. Phys. 41 (2000) 3268–3277.
[59] A. Iserles, H.Z. Munthe-Kaas, S.P. Nørsett, A. Zanna, Lie-group methods, Acta Numer. 9 (2000) 215–365.
[60] K. Ebrahimi-Fard, D. Manchon, A Magnus- and Fer-type formula in dendriform algebras. Found. Comput. Math. (2008), doi:10.1007/s10208-008-9023-3.
[61] A. Iserles, S.P. Nørsett, A.F. Rasmussen, Time symmetry and high-order Magnus methods, Appl. Numer. Math. 39 (2001) 379–401.
[62] H. Munthe-Kaas, B. Owren, Computations in a free Lie algebra, Phil. Trans. R. Soc. A 357 (1999) 957–981.
[63] E.A. Coddington, N. Levinson, Theory of Ordinary Differential Equations, McGraw Hill, 1955.
[64] F.R. Gantmacher, The Theory of Matrices, Chelsea Publishing, 1959.
[65] J. Dixmier, L'application exponentielle dans les groupes de Lie résolubles, Bull. Soc. Math. France 85 (1957) 113–121.
[66] M. Saito, Sur certain groupes de Lie résolubles, I; II, Sci. Papers Coll. Gen. Educ. Univ. Tokyo 7 (1957) 1–11.
[67] P.C. Moan, On backward error analysis and Nekhoroshev stability in the numerical analysis of conservative systems of ODEs, Ph.D. Thesis, University of Cambridge, 2002.
[68] M.V. Karasev, M.V. Mosolova, Infinite products and T products of exponentials, J. Theor. Math. Phys. 28 (1977) 721–730.

[69] R.V. Chacon, A.T. Fomenko, Recursion formulas for the Lie integral, Adv. Math. 88 (1991) 200–257.
[70] P.C. Moan, Efficient approximation of Sturm–Liouville problems using Lie-group methods, Technical Report 1998/NA11, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England, 1998.
[71] P.C. Moan, J. Niesen, Convergence of the Magnus series, Found. Comput. Math. 8 (2008) 291–301.
[72] B.S. Mityagin, Unpublished notes, 1990.
[73] T.H. Gronwall, Note on the derivatives with respect to a parameter of the solutions of a system of differential equation, Ann. of Math. 20 (1919) 292–296.
[74] J.K. Hunter, B. Nachtergaele, Applied Analysis, World Scientific, 2001.
[75] N. Dunford, J.T. Schwartz, Linear Operators. Part I: General Theory, Wiley-Interscience, 1958.
[76] V.A. Yakubovich, V.M. Starzhinskii, Linear Differential Equations with Periodic Coefficients, John Wiley & Sons, 1975.
[77] M.M. Veshtort, Numerical simulations in Nuclear Magnetic Resonance: Theory and applications, Ph.D. Thesis, Massachusetts Institute of Technology, 2003.
[78] V.V. Gorbatsevich, A.L. Onishchik, E.B. Vinberg, Foundations of Lie Theory and Lie Transformation Groups, Springer, 1997.
[79] E. Hairer, Ch. Lubich, G. Wanner, Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations, second ed., Springer-Verlag, 2006.
[80] K. Kumar, On expanding the exponential, J. Math. Phys. 6 (1965) 1928–1934.
[81] G.H. Weiss, A.A. Maradudin, The Baker–Hausdorff formula and a problem in Crystal Physics, J. Math. Phys. 3 (1962) 771–777.
[82] M.V. Berry, Quantal phase factors accompanying adiabatic changes, Proc. Roy. Soc. Lond. Ser. A 392 (1984) 45–57.
[83] L.M. Garrido, Generalized adiabatic invariance, J. Math. Phys. 5 (1964) 355–362.
[84] M.V. Berry, Quantum phase corrections from adiabatic iteration, Proc. Roy. Soc. Lond. Ser. A 414 (1987) 31–46.
[85] R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, 1960.
[86] S. Klarsfeld, J.A. Oteo, Exponential infinite-product representations of the time displacement operator, J. Phys. A: Math. Gen. 22 (1989) 2687–2694.
[87] J. Wei, E. Norman, On global representations of the solutions of linear differential equations as a product of exponentials, Proc. Amer. Math. Soc. 15 (1964) 327–334.
[88] J.G.F. Belinfante, B. Kolman, A Survey of Lie Groups and Lie Algebras with Applications and Computational Methods, SIAM, 1972.
[89] F. Casas, Solution of linear partial differential equations by Lie algebraic methods, J. Comput. Appl. Math. 76 (1996) 159–170.
[90] M. Kawski, The combinatorics of nonlinear controllability and noncommuting flows, in: Mathematical Control Theory, in: ICTP Lect. Notes, vol. VIII, Abdus Salam Int. Cent. Theoret. Phys., Trieste, 2002, pp. 223–311.
[91] A. Murua, The Hopf algebra of rooted trees, free Lie algebras, and Lie series, Found. Comput. Math. 6 (2006) 387–426.
[92] D. Voslamber, On exponential approximations for the evolution operator, J. Math. Anal. Appl. 37 (1972) 403–411.
[93] J.A. Oteo, J. Ros, Iterative approach to the exponential representation of the time-displacement operator, J. Phys. A: Math. Gen. 38 (2005) 1–11.
[94] F. Bloch, Über die Quantenmechanik der Electronen in Kristallgittern, Z. Physik 52 (1928) 555–600.
[95] A. Galindo, P. Pascual, Quantum Mechanics, Springer, 1990.
[96] J.K. Hale, Ordinary Differential Equations, Krieger Publishing Company, 1980.
[97] T.O. Levante, M. Baldus, B.H. Meier, R.R. Ernst, Formalized quantum mechanical Floquet theory and its applications to sample spinning in nuclear magnetic resonance, Mol. Phys. 86 (1995) 1195–1212.
[98] H. Munthe-Kaas, A. Zanna, Numerical integration of differential equations on homogeneous manifolds, in: F. Cucker, M. Shub (Eds.), Foundations of Computational Mathematics, Springer Verlag, 1997, pp. 305–315.
[99] F. Casas, A. Iserles, Explicit Magnus expansions for nonlinear equations, J. Phys. A: Math. Gen. 39 (2006) 5445–5461.
[100] M.P. Calvo, A. Iserles, A. Zanna, Numerical solution of isospectral flows, Math. Comput. 66 (1997) 1461–1486.
[101] R. Brockett, Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems, Linear Algebra Appl. 146 (1991) 79–91.
[102] M.T. Chu, K.R. Driessel, The projected gradient method for least squares matrix approximations with spectral constraints, SIAM J. Numer. Anal. 27 (1990) 1050–1060.
[103] J.B. Moore, R.E. Mahony, U. Helmke, Numerical gradient algorithms for eigenvalue and singular value calculations, SIAM J. Matrix Anal. Appl. 15 (1994) 881–902.
[104] A. Iserles, On the discretization of double-bracket flows, Found. Comput. Math. 2 (2002) 305–329.
[105] F. Casas, Numerical integration methods for the double-bracket flow, J. Comput. Appl. Math. 166 (2004) 477–495.
[106] A.A. Agrachev, R.V. Gamkrelidze, Chronological algebras and nonstationary vector fields, J. Sov. Math. 17 (1979) 1650–1675.
[107] A.A. Agrachev, R.V. Gamkrelidze, The exponential representation of flows and the chronological calculus, Math. USSR-Sb. 35 (1979) 727–785.
[108] R.V. Gamkrelidze, Exponential representation of solutions of ordinary differential equations, in: Lecture Notes in Math., vol. 703, Springer-Verlag, 1979, pp. 118–129.
[109] J.A. Oteo, J. Ros, The Magnus expansion for classical Hamiltonian systems, J. Phys. A: Math. Gen. 24 (1991) 5751–5762.
[110] S. Blanes, P.C. Moan, Splitting methods for non-autonomous Hamiltonian equations, J. Comput. Phys. 170 (2001) 205–230.
[111] P.J. Olver, Applications of Lie Groups to Differential Equations, second ed., Springer-Verlag, 1993.
[112] V.I. Arnold, Mathematical Methods of Classical Mechanics, second ed., Springer-Verlag, 1989.
[113] W. Thirring, Classical Dynamical Systems, Springer-Verlag, 1992.
[114] R.A. Marcus, High-order time-dependent perturbation theory for classical mechanics and for other systems of first-order ordinary differential equations, J. Chem. Phys. 52 (1970) 4803–4807.
[115] F. Spirig, Algebraic aspects of perturbation theories, Celest. Mech. 20 (1979) 343–354.
[116] S. Tani, Canonical transformation governing temporal development in classical mechanics, Amer. J. Phys. 36 (1968) 29–32.
[117] K.-T. Chen, Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula, Ann. of Math. 65 (1957) 163–178.
[118] M. Fliess, Fonctionnelles causales non linéaires et indéterminées non commutatives, Bull. Soc. Math. France 109 (1981) 3–40.
[119] M. Kawski, Calculating the logarithm of the Chen–Fliess series, Proc. MTNS, Perpignan, 2000.
[120] S. Klarsfeld, J.A. Oteo, Driven harmonic oscillators in the adiabatic Magnus approximation, Phys. Rev. A 47 (1993) 1620–1624.
[121] N. Rosen, C. Zener, Double Stern–Gerlach experiment and related collision phenomena, Phys. Rev. 40 (1932) 502–507.
[122] S. Klarsfeld, J.A. Oteo, Collisional models in a nonperturbative approach, Phys. Lett. A 164 (1992) 38–42.
[123] S. Klarsfeld, J.A. Oteo, Magnus approximation in the adiabatic picture, Phys. Rev. A 45 (1992) 3229–3332.
[124] R. Lefebvre, A. Palma, The harmonic Floquet oscillator, J. Mol. Struct. (Theochem) 390 (1997) 23–32.
[125] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations, John Wiley, 1987.
[126] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I, Nonstiff Problems, Second revised edition, Springer-Verlag, 1993.
[127] B. Leimkuhler, S. Reich, Simulating Hamiltonian Dynamics, Cambridge University Press, 2004.
[128] R.I. Mclachlan, R. Quispel, Splitting methods, Acta Numer. 11 (2002) 341–434.
[129] R.I. Mclachlan, R. Quispel, Geometric integrators for ODEs, J. Phys. A: Math. Gen. 39 (2006) 5251–5285.
[130] S.-K. Chan, J.C. Light, J.-L. Lin, Inelastic molecular collisions: Exponential solution of coupled equations for vibration–translation energy transfer, J. Chem. Phys. 49 (1968) 86–97.
[131] P.L. Devries, A high order method for the solution of the single channel Schrödinger equation, Mol. Phys. 54 (1985) 733–740.
[132] S. Blanes, F. Casas, Optimization of Lie group methods for differential equations, Future Gener. Comput. Syst. 19 (2003) 331–339.
[133] A. Iserles, A. Marthinsen, S.P. Nørsett, On the implementation of the method of Magnus series for linear differential equations, BIT 39 (1999) 281–304.
[134] A. Zanna, The Fer expansion and time symmetry: A Strang-type approach, Appl. Numer. Math. 39 (2001) 435–459.
[135] A. Iserles, On Cayley-transform methods for the discretization of Lie-group equations, Found. Comput. Math. 1 (2001) 129–160.
[136] F. Diele, L. Lopez, R. Peluso, The Cayley transform in the numerical solution of unitary differential systems, Adv. Comput. Math. 8 (1998) 317–334.

[137] A. Marthinsen, B. Owren, Quadrature methods based on the Cayley transform, Appl. Numer. Math. 39 (2001) 403–413.
[138] C.B. Moler, C.F. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, SIAM Review 20 (1978) 801–836.
[139] C.B. Moler, C.F. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later, SIAM Review 45 (2003) 3–49.
[140] R.B. Sidje, Expokit software, http://www.maths.uq.edu.au/expokit.
[141] R.B. Sidje, Expokit: Software package for computing matrix exponentials, ACM Trans. Math. Software 24 (1998) 130–156.
[142] A.J. Dragt, Computation of maps for particle and light optics by scaling, splitting, and squaring, Phys. Rev. Lett. 75 (1995) 1946–1948.
[143] I. Najfeld, T.F. Havel, Derivatives of the matrix exponential and their computation, Adv. Appl. Math. 16 (1995) 321–375.
[144] H. Tal-Ezer, R. Kosloff, An accurate and efficient scheme for propagating the time dependent Schrödinger equation, J. Chem. Phys. 81 (1984) 3967–3971.
[145] E. Süli, D. Mayers, An Introduction to Numerical Analysis, Cambridge University Press, 2003.
[146] P. Nettesheim, Mixed quantum-classical dynamics: A unified approach to mathematical modelling and numerical simulation, Ph.D. Thesis, Freie Universität Berlin, 2000.
[147] C. Leforestier, R.H. Bisseling, C. Cerjan, M.D. Feit, R. Friesner, A. Guldberg, A. Hammerich, G. Jolicard, W. Karrlein, H.-D. Meyer, N. Lipkin, O. Roncero, R. Kosloff, A comparison of different propagation schemes for the time dependent Schrödinger equation, J. Comput. Phys. 94 (1991) 59–80.
[148] V.L. Druskin, L.A. Knizhnerman, Krylov subspace approximation of eigenpairs and matrix functions in exact and computer arithmetics, Numer. Linear Algebra 2 (1995) 205–217.
[149] T.J. Park, J.C. Light, Unitary quantum time evolution by iterative lanczos reduction, J. Chem. Phys. 85 (1986) 5870–5876.
[150] D.S. Watkins, Fundamentals of Matrix Computations, 2nd ed., John Wiley & Sons, 2002.
[151] M. Hochbruck, C. Lubich, Exponential integrators for quantum-classical molecular dynamics, BIT 39 (1999) 620–645.
[152] M. Hochbruck, C. Lubich, H. Selhofer, Exponential integrators for large systems of differential equations, SIAM J. Sci. Comput. 19 (1998) 1552–1574.
[153] M. Hochbruck, Ch. Lubich, On Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal. 34 (1997) 1911–1925.
[154] S. Blanes, P.C. Moan, Practical symplectic partitioned Runge–Kutta and Runge–Kutta–Nyström methods, J. Comput. Appl. Math 142 (2002) 313–330.
[155] K. Feng, Z.-J. Shang, Volume-preserving algorithms for source-free dynamical systems, Numer. Math. 71 (1995) 451–463.
[156] E. Celledoni, A. Iserles, Approximating the exponential from a Lie algebra to a Lie group, Math. Comput. 69 (2000) 1457–1480.
[157] E. Celledoni, A. Iserles, Methods for the approximation of the matrix exponential in a Lie-algebraic setting, IMA J. Numer. Anal. 21 (2001) 463–488.
[158] H. Munthe-Kaas, R.G.W. Quispel, A. Zanna, Generalized polar decompositions on Lie groups with involutive automorphisms, Found. Comput. Math. 1 (2001) 297–324.
[159] A. Iserles, A. Zanna, Efficient computation of the matrix exponential by generalized polar decomposition, SIAM J. Numer. Anal. 42 (2005) 2218–2256.
[160] A. Zanna, H.Z. Munthe-Kaas, Generalized polar decompositions for the approximation of the matrix exponential, SIAM J. Matrix Anal. Appl. 23 (2001) 840–862.
[161] A. Iserles, On the global error of discretization methods for highly-oscillatory ordinary differential equations, BIT 42 (2002) 561–599.
[162] A. Iserles, Think globally, act locally: Solving highly-oscillatory ordinary differential equations, Appl. Numer. Math. 43 (2002) 145–160.
[163] A. Iserles, On the numerical quadrature of highly-oscillating integrals I: Fourier transform, IMA J. Numer. Anal. 24 (2004) 365–391.
[164] A. Iserles, S.P. Nørsett, On quadrature methods for highly oscillatory integrals and their implementation, BIT 44 (2004) 755–772.
[165] I. Degani, J. Schiff, RCMS: Right correction Magnus series approach for oscillatory ODEs, J. Comput. Appl. Math 193 (2006) 413–436.
[166] T. Jahnke, Long-time-step integrators for almost-adiabatic quantum dynamics, SIAM J. Sci. Comput. 25 (2004) 2145–2164.
[167] T. Jahnke, C. Lubich, Numerical integrators for quantum dynamics close to the adiabatic limit, Numer. Math. 94 (2003) 289–314.
[168] K. Lorenz, T. Jahnke, C. Lubich, Adiabatic integrators for highly oscillatory second-order linear differential equations with time-varying eigendecomposition, BIT 45 (2005) 91–115.
[169] N.D. Aparicio, S.J.A. Malham, M. Oliver, Numerical evaluation of the Evans function by Magnus integration, BIT 45 (2005) 219–258.
[170] S.J.A. Malham, J. Niesen, Evaluating the Evans function: Order reduction in numerical methods, Math. Comput. 77 (2008) 159–179.
[171] S. Blanes, P.C. Moan, Splitting methods for the time-dependent Schrödinger equation, Phys. Lett. A 265 (2000) 35–42.
[172] D. Baye, G. Goldstein, P. Capel, Fourth-order factorization of the evolution operator for time-dependent potentials, Phys. Lett. A 317 (2003) 337–342.
[173] Y.Y. Lu, Fourth order conservative exponential methods for linear evolution equations, BIT 40 (2000) 762–774.
[174] M. Thalhammer, A fourth-order commutator-free exponential integrator for nonautonomous differential equations, SIAM J. Numer. Anal. 44 (2006) 851–864.
[175] S. Blanes, P.C. Moan, Fourth- and sixth-order commutator-free Magnus integrators for linear and non-linear dynamical systems, Appl. Numer. Math. 56 (2006) 1519–1537.
[176] J.M. Sanz-Serna, A. Portillo, Classical numerical integrators for wave-packet dynamics, J. Chem. Phys. 104 (1996) 2349–2355.
[177] S. Blanes, F. Casas, Splitting methods for non-autonomous separable dynamical systems, J. Phys. A: Math. Gen. 39 (2006) 5405–5423.
[178] S. Blanes, F. Casas, A. Murua, Splitting methods for non-autonomous linear systems, Int. J. Comput. Math. 6 (2007) 713–727.
[179] H. Munthe-Kaas, Runge–Kutta methods on Lie groups, BIT 38 (1998) 92–111.
[180] H. Munthe-Kaas, High order Runge–Kutta methods on manifolds, Appl. Numer. Math. 29 (1999) 115–127.
[181] A. Zanna, Collocation and relaxed collocation for the Fer and the Magnus expansions, SIAM J. Numer. Anal. 36 (1999) 1145–1182.
[182] C. Lubich, Integrators for quantum dynamics: A numerical analyst's brief review, in: J. Grotendorst, D. Marx, A. Muramatsu (Eds.), Quantum Simulations of Many-Body Systems: From Theory to Algorithms, in: NIC Series, vol. 10, Jülich, 2002, pp. 459–466.
[183] M. Hochbruck, C. Lubich, On Magnus integrators for time-dependent Schrödinger equations, SIAM J. Numer. Anal. 41 (2003) 945–963.
[184] S. Gray, J.M. Verosky, Classical Hamiltonian structures in wave packet dynamics, J. Chem. Phys. 100 (1994) 5011–5022.
[185] L. Jódar, M. Marletta, Solving ODEs arising from non-selfadjoint Hamiltonian eigenproblems, Adv. Comput. Math. 13 (2000) 231–256.
[186] J.D. Pryce, Numerical Solution of Sturm–Liouville Problems, Oxford University Press, 1993.
[187] A. Zettl, Sturm–Liouville Theory, American Mathematical Society, 2005.
[188] P.B. Bailey, M.K. Gordon, L.F. Shampine, Automatic solution of the Sturm–Liouville problem, ACM Trans. Math. Software 4 (1978) 193–208.
[189] L. Greenberg, M. Marletta, Oscillation theory and numerical solution of sixth order Sturm–Liouville problems, SIAM J. Numer. Anal. 35 (1998) 2070–2098.
[190] M.A. Naimark, Linear Differential Operators, Part I, Ungar, 1968.
[191] L. Dieci, Numerical integration of the differential Riccati equation and some related issues, SIAM J. Numer. Anal. 29 (1992) 781–815.
[192] L. Dieci, M.R. Osborne, R.D. Russell, A Riccati transformation method for solving linear BVP I: Theoretical aspects, II: Computational aspects, SIAM J. Numer. Anal. 25 (1988) 1055–1092.
[193] S. Bittanti, A.J. Laub, J.C. Willems (Eds.), The Riccati Equation, Springer, 1991.
[194] W.T. Reid, Riccati Differential Equations, Academic Press, 1972.
[195] C.S. Kenney, R.B. Leipnik, Numerical integration of the differential Riccati equation, IEEE Trans. Automat. Control AC-30 (1985) 962–970.
[196] J. Schiff, S. Shnider, A natural approach to the numerical integration of Riccati differential equations, SIAM J. Numer. Anal. 36 (1999) 1392–1413.
[197] K. Burrage, P.M. Burrage, T. Tian, Numerical methods for strong solutions of stochastic differential equations: An overview, Proc. Roy. Soc. Lond. Ser. A 460 (2004) 373–402.
[198] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer Verlag, 1992.
[199] E. Platen, An introduction to numerical methods for stochastic differential equations, Acta Numer. 8 (1999) 197–246.
[200] G. Maruyama, Continuous Markov processes and stochastic equations, Rend. Circ. Mat. Palermo 4 (1955) 48–90.
[201] G. Lord, S.J.A. Malham, A. Wiese, Efficient strong integrators for linear stochastic systems, SIAM J. Numer. Anal. (2006).
[202] S.J.A. Malham, A. Wiese, Stochastic Lie group integrators, SIAM J. Sci. Comput. 30 (2008) 597–617.
[203] T. Misawa, A Lie algebraic approach to numerical integration of stochastic differential equations, SIAM J. Sci. Comput. 23 (2001) 866–890.
[204] J. Eichler, Magnus approximation for K-shell ionization by heavy-ion impact, Phys. Rev. A 15 (1977) 1856–1862.

[205] U. Wille, Magnus expansion for rotationally induced inner-shell excitation, Phys. Lett. A 82 (1981) 389–392.
[206] U. Wille, R. Hippler, Mechanisms of inner-shell vacancy production in slow ion-atom collisions, Phys. Rep. 132 (1986) 129–260.
[207] H.A. Hyman, Dipole Magnus approximation for electron–atom collisions: Excitation of the resonance transitions of Li, Na and K, Phys. Rev. A 31 (1985) 2142–2148.
[208] W.A. Cady, Rotational spectral line broadening of OCS by noble gases, J. Chem. Phys. 60 (1974) 3318–3323.
[209] K.F. Milfeld, R.E. Wyatt, Study, extension, and application of Floquet theory for quantum molecular systems in an oscillating field, Phys. Rev. A 27 (1983) 72–94.
[210] W. Evans, On some applications of Magnus expansion in Nuclear Magnetic Resonance, Ann. Phys. 48 (1968) 72–93.
[211] U. Haeberlen, J.S. Waugh, Coherent averaging effects in Magnetic resonance, Phys. Rev. 175 (1968) 453–467.
[212] L.M.K. Vandersypen, I.L. Chuang, NMR techniques for quantum control and computation, Rev. Mod. Phys. 76 (2004) 1037–1069.
[213] M. Veshtort, R.G. Griffin, Spin evolution: A powerful tool for the simulation of solid and liquid state NMR experiments, J. Magn. Reson. 178 (2006) 248–282.
[214] E.V. Stefanovich, Quantum field theory without infinities, Ann. Phys. 292 (2001) 139–156.
[215] H.D. Dahmen, W. Krzyzanowski, M.L. Larsen, Graphical rules for functions of the time-evolution operator, Phys. Rev. D 33 (1986) 1726–1744.
[216] H.D. Dahmen, T. Mannel, P. Manakos, Practical application of the unitary approximation to infrared singularities, Phys. Rev. D 38 (1988) 1176–1182.
[217] H.D. Dahmen, P. Manakos, T. Mannel, T. Ohl, Unitary approximation for radiative high energy scattering processes: Application to Bhabha scattering, Z. Phys. C 50 (1991) 75–84.
[218] A. Connes, D. Kreimer, Renormalization in quantum field theory and the Riemann–Hilbert problem I. The Hopf algebra structure of graphs and the main theorem, Comm. Math. Phys. 210 (2000) 249–273.
[219] A. Connes, D. Kreimer, Renormalization in quantum field theory and the Riemann–Hilbert problem II. The β-function, diffeomorphisms and the renormalization group, Comm. Math. Phys. 216 (2001) 215–241.
[220] K. Ebrahimi-Fard, D. Manchon, F. Patras, A noncommutative Bohnenblust–Spitzer identity for Rota–Baxter algebras solves Bogoliubov's recursion, Technical Report, arXiv:0705.1265v2, 19 August 2008.
[221] D.C. Ionescu, Unitary and electron–positron pairs created by strong external fields, Phys. Rev. A 49 (1994) 3188–3195.
[222] J.C. D'Olivo, Magnus approximation for nonadiabatic transitions and neutrino oscillations in matter, Phys. Rev. D 45 (1992) 924–929.
[223] J.C. D'Olivo, J.A. Oteo, Nonadiabatic three-neutrino oscillations in matter, Phys. Rev. D 54 (1996) 1187–1193.
[224] A.D. Supanitsky, J.C. D'Olivo, G. Medina-Tanco, Perturbative exponential expansion and matter neutrino oscillations, Phys. Rev. D 78 (2008) 045024.
[225] M.A. Botchev, J.G. Verwer, Numerical integration of damped Maxwell equations, Technical Report, CWI Amsterdam, April 2008.
[226] L.-L. Jiang, J.-F. Mao, X.-L. Wu, Symplectic finite difference time-domain method for Maxwell equations, IEEE Trans. Mag. 42 (2006) 1991–1995.
[227] Y.Y. Lu, Some techniques for computing propagation in optical waveguides, Comm. Comput. Phys. 1 (2006) 1056–1075.
[228] G. Dattoli, J.C. Gallardo, A. Torre, An algebraic view to the operatorial ordering and its applications to Optics, Riv. Nuovo Cimento 11 (1988) 1–79.
[229] S.A. Khan, Wavelength-dependent modifications in Helmholtz Optics, Int. J. Theor. Phys. 44 (2005) 96–125.
[230] Y.Y. Lu, A fourth-order Magnus scheme for Helmholtz equation, J. Comput. Appl. Math. 173 (2005) 247–258.
[231] Y.Y. Lu, A fourth order derivative-free operator marching method for Helmholtz equation in waveguides, J. Comput. Math. 25 (2007) 719–729.
[232] M. Reimer, D. Yevick, Clifford analysis of polarization-mode dispersion and polarization-dependent loss, IEEE Photon. Technol. Lett. 18 (2006) 734–736.
[233] M. Reimer, D. Yevick, Mueller matrix description of polarization mode dispersion and polarization-dependent loss, J. Opt. Soc. Amer. A 23 (2006) 1503–1508.
[234] A. López-Ariste, On the asymmetry of Stokes profiles, Astrophys. J. 564 (2002) 379–384.
[235] A. López-Ariste, M. Semel, DIAGONAL: A numerical solution of the Stokes transfer equation, Astron. Astrophys. Suppl. Ser. 139 (1999) 417–424.
[236] M. Semel, A. López Ariste, Integration of the radiative transfer equation for polarized light: The exponential solution, Astron. Astrophys. 342 (1999) 201–211.
[237] A. San Miguel, Numerical determination of time transfer in general relativity, Gen. Relativ. Gravit. 39 (2007) 2025–2033.
[238] D. Viswanath, Symbolic dynamics and periodic orbits of the Lorenz attractor, Nonlinearity 16 (2003) 1035–1056.
[239] F. Bullo, A.D. Lewis, Geometric Control of Mechanical Systems, Springer, 2005.
[240] R.E. Kalman, Y.-C. ho, K.S. Narendra, Controllability of linear dynamical systems, Contributions to Differential Equations 1 (1963) 189–213.
[241] I. Duleba, Locally optimal motion planning of nonholonomic systems, Int. J. Robot. Syst. 14 (1997) 767–788.
[242] I. Duleba, On a computationally simple form of the generalized Campbell–Baker–Hausdorff–Dynkin formula, Systems Control Lett. 34 (1998) 191–202.
[243] R.M. Murray, Z. Li, S.S. Satry, A Mathematical Introduction to Robotic Manipulation, CRC Press, 1994.
[244] J.C. Engwerda, LQ Dynamic Optimization and Differential Games, Wiley, 2005.
[245] S. Blanes, L. Jódar, E. Ponsoda, Approximate solutions with a priori error bounds for continuous coefficient matrix Riccati equations, Math. Comput. Modelling 31 (2000) 1–15.
[246] L. Jódar, H. Abou-Kandil, A resolution method for Riccati differential systems coupled in their quadratic terms, SIAM J. Matrix Anal. Appl. 19 (1988) 1425–1430.