IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Sufficient conditions for the convergence of the Magnus expansion

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2007 J. Phys. A: Math. Theor. 40 15001

(http://iopscience.iop.org/1751-8121/40/50/006)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.147
The article was downloaded on 03/06/2010 at 06:30

Please note that terms and conditions apply.

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) 15001–15017 doi:10.1088/1751-8113/40/50/006

Sufficient conditions for the convergence of the Magnus expansion

Fernando Casas

Departament de Matemàtiques, Universitat Jaume I, E-12071 Castellón, Spain

E-mail: Fernando.Casas@uji.es

Received 30 July 2007, in final form 30 October 2007
Published 28 November 2007
Online at stacks.iop.org/JPhysA/40/15001

Abstract
Two different sufficient conditions are given for the convergence of the Magnus expansion arising in the study of the linear differential equation $Y' = A(t)Y$. The first one provides a bound on the convergence domain based on the norm of the operator $A(t)$. The second condition links the convergence of the expansion with the structure of the spectrum of $Y(t)$, thus yielding a more precise characterization. Several examples are proposed to illustrate the main issues involved and the information on the convergence domain provided by both conditions.

PACS numbers: 02.30.Hq, 02.30.Lt, 02.70.−c

1. Introduction

The approach followed by Magnus in [21] to solve the non-autonomous linear differential equation

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = A(t)Y, \qquad Y(0) = I, \tag{1.1}$$

where $Y(t)$ and $A(t)$ are (sufficiently smooth real or complex) $n \times n$ matrices, is to express the solution $Y(t)$ as the exponential of a certain matrix $\Omega(t)$,

$$Y(t) = \exp \Omega(t). \tag{1.2}$$

By substituting (1.2) into (1.1), one can derive the differential equation satisfied by the exponent Ω [15]:

$$\Omega' = \mathrm{d}\exp_\Omega^{-1}(A(t)), \qquad \Omega(0) = O, \tag{1.3}$$

where $\mathrm{d}\exp_\Omega^{-1}$ is the inverse operator of the power series

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

Specifically, its expression is given by

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

Here $\{B_k\}_{k\in\mathbb{Z}_+}$ are the Bernoulli numbers [1], ad^k is a shorthand for an iterated commutator,

$$\mathrm{ad}_{\Omega}^{0} A = A, \qquad \mathrm{ad}_{\Omega}^{k+1} A = \left[\Omega, \mathrm{ad}_{\Omega}^{k} A\right],$$

and $[\Omega, A] = \Omega A - A\Omega$. By applying Picard's iteration on (1.3), one gets an infinite series for $\Omega(t)$,

$$\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t), \tag{1.4}$$

whose first terms read

$$\begin{aligned}
\Omega_1(t) &= \int_0^t A(t_1)\mathrm{d}t_1, \\
\Omega_2(t) &= \frac{1}{2}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\, [A(t_1), A(t_2)] \\
\Omega_3(t) &= \frac{1}{6}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3([A(t_1), [A(t_2), A(t_3)]] + [A(t_3), [A(t_2), A(t_1)]]).
\end{aligned} \tag{1.5}$$

Explicit formulae for Ω_m of all orders have been given in [16], whereas in [19] a recursive procedure for the generation of any Ω_m was proposed, which presents some advantages from a computational point of view. When this recursion is worked out explicitly, it is possible to express Ω_m as a linear combination of m-fold integrals of $m-1$ nested commutators containing m operators A,

$$\Omega_m(t) = \sum_{j=1}^{m-1} \frac{B_j}{j!} \sum_{\substack{k_1+\cdots+k_j=m-1 \\ k_1\geqslant 1,\ldots,k_j\geqslant 1}} \int_0^t \mathrm{ad}_{\Omega_{k_1}(s)} \mathrm{ad}_{\Omega_{k_2}(s)} \cdots \mathrm{ad}_{\Omega_{k_j}(s)} A(s)\, \mathrm{d}s \qquad m \geqslant 2, \tag{1.6}$$

an expression, however, that becomes increasingly intricate with m, as it should be already evident from the first terms (1.5).

Equations (1.2) and (1.4) constitute the so-called *Magnus expansion* for the solution of (1.1), whereas the infinite series (1.4) with (1.6) is known as the *Magnus series*.

Since the 1960s, the Magnus expansion has been successfully applied as a perturbative tool in numerous areas of physics and chemistry, from atomic and molecular physics to nuclear magnetic resonance and quantum electrodynamics (see [2, 3] for a review and a list of references). Also, since the work by Iserles and Nørsett [16], it has been used as a tool to construct practical algorithms for the numerical integration of equation (1.1), while preserving the main qualitative properties of the exact solution. In this sense, the corresponding schemes are prototypical examples of *geometric numerical integrators* [13].

To be more specific, suppose that $A(t)$ belongs to some matrix Lie (sub) algebra $\mathfrak{g}$ for all t. Then the exact solution of (1.1) evolves in the matrix Lie group $\mathcal{G}$ having $\mathfrak{g}$ as its corresponding Lie algebra (the tangent space at the identity of $\mathcal{G}$). Observe now that all terms in the Magnus series are constructed as sums of multiple integrals of nested commutators, so that Ω and indeed any approximation to it obtained by truncation will also be in the same Lie algebra. Finally, its exponential will be in $\mathcal{G}$. By truncating appropriately the series, approximating efficiently the multivariate integrals by quadratures and reducing the number of required commutators, it is possible to design new integrators based on the Magnus expansion which have proved

to be highly competitive with other, more conventional schemes with respect to accuracy and computational effort in the numerical integration of (1.1) on matrix Lie groups [4, 5, 15].

Although the Magnus expansion has been formulated here only for $n \times n$ matrices, the same result is also valid (as least formally) in a more general setting. As a matter of fact, it was originally established assuming only that $A(t)$ is a known function of t in an associative ring [21]. On the other hand, in order to apply this approach in quantum mechanics, it is tacitly assumed that the expansion is also valid when $A(t)$ is a linear operator in a Hilbert space.

From a mathematical point of view, it is clear that there are at least two different issues of paramount importance at the very basis of the Magnus expansion:

- First, for what values of t and for what operators A does equation (1.1) admit a true exponential solution in the form (1.2) with a certain $\Omega(t)$? This could be called the *existence* problem.
- Second, given a certain operator $A(t)$, for what values of t does the Magnus series (1.4) converge? In other words, when $\Omega(t)$ in (1.2) can be obtained as the sum of the series (1.4)? This we describe as the *convergence* problem.

Of course, given the relevance of the expansion, both problems have been extensively treated in the literature since Magnus proposed this formalism in 1954 [21]. In section 2, we review some of the most relevant contributions already available regarding both aspects, whereas in the rest of the paper we will concentrate ourselves on the convergence issue. Thus, in section 3 we provide a general result on the convergence of the Magnus series which is valid for bounded linear operators $A(t)$ in a Hilbert space, and not only for real matrices. Then, in section 4, we analyze the problem from a different point of view, characterizing the convergence (or divergence) of the series in terms of the eigenvalues of the matrix $Y(t)$. This allows us, in some cases, to obtain more accurate estimates and at the same time gives us more insight into the convergence problem. Several examples are also considered to illustrate the main issues involved. Finally, section 5 contains a discussion of the results obtained.

2. Existence and convergence of $\Omega(t)$: previous results

2.1. *On the existence of* $\Omega(t)$

In most cases, one is interested in the case where $A(t)$ belongs to a Lie algebra $\mathfrak{g}$ under the commutator product. In this rather general setting, Magnus result can be formulated as four statements concerning the solution of $\mathrm{d}Y/\mathrm{d}t = A(t)Y$, each one more stringent than the preceding [31]. Specifically,

(a) The differential equation $\mathrm{d}Y/\mathrm{d}t = A(t)Y$ 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 $\mathrm{d}\Omega/\mathrm{d}t = \mathrm{d}\exp_{\Omega}^{-1}(A(t))$.
(d) The operator $\Omega(t)$ can be computed by a series

$$\Omega(t) = \Omega_1(t) + \Omega_2(t) + \cdots,$$

where every term is a multivariate integral involving a linear combination of nested commutators of A evaluated at different times (i.e., the Magnus series (1.4) with (1.6)).

We proceed now to analyze in detail the conditions under which statements (A)–(C) hold, whereas the validity of (D) will be established by examining the convergence problem in the rest of the paper.

(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 (1.1) 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 [7]. 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, any matrix Y can be written in the form $\exp\Omega$ if and only if $\det Y \neq 0$ [11, 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 $\Omega(t) \in \mathfrak{g}$ [8, 29]: the exponential map of a complex Lie algebra is globally one to one if and only if the algebra is nilpotent. 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 $\mathfrak{g}$ [25, 26].

(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. For instance, the logarithm of $Y(t)$ may be complex even for real $A(t)$ [31]. In such a situation, the solution of (1.1) 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 (D), 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 [31].

It is therefore interesting to determine for which range of t a real matrix $A(t)$ in (1.1) leads to a real logarithm. This issue has been tackled in [26] in the context of a complete normed (Banach) algebra, proving that if

$$\int_0^t \|A(s)\|_2\,\mathrm{d}s < \pi \tag{2.1}$$

then the solution of (1.1) can be written indeed as $Y(t) = \exp\Omega(t)$, where $\Omega(t)$ is in the Banach algebra. In (2.1), $\|\cdot\|_2$ stands specifically for the 2-norm (or spectral norm) of A.

(C) In his original paper [21], 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 existence of the right-hand side of equation (1.3) and gave an implicit condition for an arbitrary A. More specifically, he proved the following result for the case of $n \times n$ matrices (theorem 5 in [21]):

Theorem 2.1. *The equation $A(t) = \mathrm{d}\exp_\Omega(\Omega')$ can be solved by $\Omega' = \mathrm{d}\exp_\Omega^{-1} A(t)$ for an arbitrary A if and only if none of the differences between any two of the eigenvalues of Ω equals $2\pi\,\mathrm{i}m$, where $m = \pm 1, \pm 2, \ldots,$ $(m \neq 0)$.*

Unfortunately, such a result has not very much practical application unless we can easily determine the eigenvalues of Ω from those of $A(t)$.

2.2. *Convergence of the Magnus series*

Let us analyze now in some detail statement (D). Magnus considered the question of when the series (1.4) terminates at some finite index m, thus giving a globally valid $\Omega = \Omega_1 + \cdots + \Omega_m$. This happens, 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 $k > l$ is that

$$[A(s_1), [A(s_2), [A(s_3), \ldots, [A(s_l), A(s_{l+1})] \cdots]]] = 0$$

for any choice of $s_1, \ldots, s_{l+1}$. 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 [21].

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 along the years. Most of these results can be stated as follows. If $\Omega_m(t)$ denotes the homogeneous element with m commutators in the Magnus series as given by (1.6), then $\Omega(t) = \sum_{m=1}^{\infty} \Omega_m(t)$ is absolutely convergent for $0 \leqslant t < T$, with

$$T = \max\left\{ t \geqslant 0 : \int_0^t \|A(s)\|_2 \, \mathrm{d}s < r_c \right\}. \tag{2.2}$$

Thus, Pechukas and Light [28] and Karasev and Mosolova [17] both obtained $r_c = \log 2 = 0.693\,147\ldots$, whereas Chacon and Fomenko [6] got $r_c = 0.57745\ldots$. In 1998, Blanes *et al* [2] and Moan [24] obtained independently the improved bound

$$r_c = \frac{1}{2} \int_0^{2\pi} \frac{1}{2 + \frac{x}{2}\left(1 - \cot \frac{x}{2}\right)} \, \mathrm{d}x \equiv \xi = 1.086\,868\,69\ldots.$$

Based on the analysis of some selected examples, Moan [26] concluded that, in order to get convergence *for all real matrices* $A(t)$, necessarily $r_c \leqslant \pi$, and more recently Moan and Niesen [27] have been able to prove rigorously that indeed $r_c = \pi$.

This result shows, in particular, 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).

3. A generic result on the convergence of the Magnus series

3.1. General formulation

As we have mentioned before, if $A(t)$ is a real $n \times n$ matrix, then (2.1) gives a condition for $Y(t)$ to have a real logarithm. Moreover, it has been shown that, under the same condition, the Magnus series (1.4) converges precisely to this logarithm, i.e., its sum $\Omega(t)$ satisfies $\mathrm{e}^{\Omega(t)} = Y(t)$ [27]. Our purpose in this section is provide a different proof of this property which in fact is also valid in the more general setting of linear operators in a Hilbert space of arbitrary dimension.

To begin with, let $A(t)$ be a bounded linear operator in a Hilbert space $\mathcal{H}$, with $2 \leqslant \dim \mathcal{H} \leqslant \infty$. Let us introduce a new parameter $\varepsilon \in \mathbb{C}$ and denote by $Y(t; \varepsilon)$ the solution of the initial-value problem

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \varepsilon A(t) Y, \qquad Y(0) = I, \tag{3.1}$$

where now I 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 $\|\cdot\|$ stands for the norm defined by the inner product on $\mathcal{H}$. Our first statement is that, 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 2.2: if, in particular, $|\varepsilon| \int_0^t \|A(s)\|\,\mathrm{d}s < \log 2$, the Magnus series corresponding to (3.1) converges and its sum $\Omega(t;\varepsilon)$ satisfies $\mathrm{e}^{\Omega(t;\varepsilon)} = Y(t;\varepsilon)$. 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 $\varepsilon = 0$.

The following theorem shows that, indeed, $\gamma = \pi$.

Theorem 3.1. *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.

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 $\mathcal{H}$, $\mathrm{Ang}\{x, y\} = \alpha$, $0 \leqslant \alpha \leqslant \pi$, 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 there. A trivial property which will be used in the sequel is the following: if $\|x\| = 1$ and $\|u\| \leqslant 1/2$, then $\mathrm{Ang}\{x + u, x\} \leqslant \|u\|(1 + \|u\|^2)$.

The second basic property of the angle we need is given by the following lemma, whose proof (due to Moan [26]) is included here for completeness.

Lemma 3.2. *(Moan). For any $x \neq 0$ in $\mathcal{H}$, $\mathrm{Ang}\{Y(t;\varepsilon)x, x\} \leqslant |\varepsilon| \int_0^t \|A(s)\|\,\mathrm{d}s$*

Proof of lemma 3.2. Let $y_0 \equiv x$ and consider the vector $y(t) = Y(t;\varepsilon)y_0$ satisfying the initial-value problem $y' = \varepsilon A(t)y$, $y(0) = y_0$. Then, clearly, $\|y'\| \leqslant |\varepsilon|\|A(t)\|\|y\|$. Let $\hat{y}(t) = \frac{y(t)}{\|y(t)\|}$ denote the unit vector in the direction of $y(t)$, so that $y' = \frac{\mathrm{d}\|y\|}{\mathrm{d}t}\hat{y} + \|y\|\hat{y}'$. On the other hand, since $\langle\hat{y},\hat{y}\rangle = 1$, then $\langle\hat{y}',\hat{y}\rangle + \langle\hat{y},\hat{y}'\rangle = \mathrm{Re}\langle\hat{y},\hat{y}'\rangle = 0$, i.e., $\hat{y}$ and $\hat{y}'$ are orthogonal. In consequence,

$$\langle y', y'\rangle = \|y'\|^2 = \left(\frac{\mathrm{d}\|y\|}{\mathrm{d}t}\right)^2 + \|y\|^2\|\hat{y}'\|^2,$$

whence, by discarding the $(\|y\|')^2$ term, $\|\hat{y}'\|^2\|y\|^2 \leqslant \|y'\|^2$, and thus $\|\hat{y}'\|\|y\| \leqslant \|y'\| \leqslant |\varepsilon|\|A(t)\|\|y\|$ or simply $\|\hat{y}'\| \leqslant |\varepsilon|\|A(t)\|$. Integrating this last inequality we get

$$\int_0^t \|\hat{y}'(s)\|\mathrm{d}s \leqslant |\varepsilon| \int_0^t \|A(s)\|\,\mathrm{d}s,$$

but

$$\int_0^t \|\hat{y}'(s)\|\,\mathrm{d}s = \int_0^t \sqrt{\langle\hat{y}'(s), \hat{y}'(s)\rangle}\,\mathrm{d}s$$

is the length (defined through the metric given by the inner product) of the curve traced by the unit vector $\hat{y}(s)$ when $s \in [0, t]$ on the unit sphere S^1, which is greater than or equal to $\mathrm{Ang}\{y(t), y_0\}$, and this proves the result. □

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

The following lemma provides useful information on the location of the eigenvalues of a given bounded linear operator in $\mathcal{H}$ [23].

Lemma 3.3. *(Mityagin). Let T be a (bounded) operator on $\mathcal{H}$. If $\mathrm{Ang}\{Tx, x\} \leqslant \gamma$ and $\mathrm{Ang}\{T^*x, x\} \leqslant \gamma$ for any $x \neq 0$, $x \in \mathcal{H}$, where T^* denotes the adjoint operator of T, then the spectrum of T, $\sigma(T)$, is contained in the set*

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

Proof of lemma 3.3. Without loss of generality, we may assume $\gamma < \pi$ (if $\gamma \geqslant \pi$, there is no statement here). If $\dim \mathcal{H} < \infty$, only the first requirement on T, $\mathrm{Ang}\{Tx, x\} \leqslant \gamma$ is sufficient, since in that case, if $\lambda = |\lambda|\mathrm{e}^{\mathrm{i}\omega} \neq 0$, $-\pi < \omega \leqslant \pi$, is in $\sigma(T)$, then there exists some $f \neq 0$ such that $Tf = \lambda f$ and

$$\mathrm{Ang}\{Tf, f\} = \mathrm{Ang}\{\lambda f, f\} = |\omega| \leqslant \gamma.$$

If, on the other hand, $\dim \mathcal{H} = \infty$ and $\lambda \in \sigma(T)$, $\lambda \neq 0$, then, as is well known, either (i) λ belongs to the approximate spectrum of T, $\sigma_{ap}(T)$, or (ii) λ is in the residual spectrum, $\sigma_r(T)$ [14].

(i) In the first case, there is a sequence $\{f_n\}$ in $\mathcal{H}$ such that $\|f_n\| = 1$ for all n and $\lim_{n\to\infty} \|(T - \lambda I)f_n\| = 0$. Equivalently, $Tf_n = \lambda f_n + \varepsilon_n$, with $\|\varepsilon_n\| \to 0$ when $n \to \infty$. Then we have

$$\gamma \geqslant \mathrm{Ang}\{Tf_n, f_n\} = \mathrm{Ang}\{\lambda f_n + \varepsilon_n, f_n\} \geqslant \mathrm{Ang}\{\lambda f_n, f_n\} - \mathrm{Ang}\{\lambda f_n, \lambda f_n + \varepsilon_n\}, \tag{3.2}$$

since the angle is a metric in S^1. Now, as $\lambda = |\lambda|\mathrm{e}^{\mathrm{i}\omega} \neq 0$, $-\pi < \omega \leqslant \pi$, it is clear from (3.2) that

$$\gamma \geqslant |\omega| - \mathrm{Ang}\left\{f_n, f_n + \frac{1}{\lambda}\varepsilon_n\right\} \geqslant |\omega| - \frac{\|\varepsilon_n\|}{|\lambda|}\left(1 + \frac{\|\varepsilon_n\|^2}{|\lambda|^2}\right), \tag{3.3}$$

where the last inequality holds when $\frac{\|\varepsilon_n\|}{|\lambda|} \leqslant \frac{1}{2}$, i.e., for sufficiently large n. Taking the limit $n \to \infty$ in (3.3) leads to $\gamma \geqslant |\omega|$.

(ii) If $\lambda \in \sigma_r(T)$, then $\bar{\lambda}$ is an eigenvalue of T^*, i.e., $\mathrm{Ker}(T^* - \bar{\lambda}I) \neq \{0\}$ [14]. Since, by assumption, $\mathrm{Ang}\{T^*x, x\} \leqslant \gamma$ for all $x \neq 0$, we can apply again the argument in (i) to T^*, $\bar{\lambda}$ and conclude that $|\omega| \leqslant \gamma$. □

Now we are ready to prove the main theorem.

Proof of theorem 3.1. Let us introduce the operator $T \equiv Y(t;\epsilon)$, with $\varepsilon \in B_\gamma$, $\gamma < \pi$. Then by lemma 3.2, $\mathrm{Ang}\{Tx, x\} \leqslant \gamma$ for all $x \neq 0$, and thus, by lemma 3.3,

$$\sigma(T) \subset \Delta_\gamma. \tag{3.4}$$

If $\dim \mathcal{H} = \infty$ and we assume that $Y(t;\epsilon)$ is a normal operator, then (3.4) also holds.

From equation (3.1) in integral form,

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

one gets $\|Y\| \leqslant 1 + |\varepsilon| \int_0^t \|A(s)\|\|Y\|\,\mathrm{d}s$, and application of Gronwall's lemma [12] leads to

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

An analogous reasoning for the inverse operator also proves that

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

In consequence,

$$\|T\| \leqslant \mathrm{e}^{\gamma} \qquad \text{and} \qquad \|T^{-1}\| \leqslant \mathrm{e}^{\gamma}.$$

If $\lambda \neq 0 \in \sigma(T)$, then $|\lambda| \leqslant \|T\|$ [14] and therefore $|\lambda| \leqslant \mathrm{e}^{\gamma}$. In addition, $\frac{1}{\lambda} \in \sigma(T^{-1})$, so that $|\lambda| \geqslant \mathrm{e}^{-\gamma}$. Equivalently,

$$\sigma(T) \subset \{z \in \mathbb{C} : \mathrm{e}^{-\gamma} \leqslant |z| \leqslant \mathrm{e}^{\gamma}\} \equiv G_{\gamma}. \tag{3.5}$$

Putting together (3.4) and (3.5), 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 $\Gamma = \partial\Lambda_{\gamma'}$. Note that the curve Γ encloses $\sigma(T)$ in its interior, so that it is possible to define the function $\varphi(\varepsilon) = \log Y(t;\varepsilon)$ by the equation [9]

$$\varphi(\epsilon) = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma} \log z (zI - Y(t;\epsilon))^{-1}\,\mathrm{d}z, \tag{3.6}$$

where the integration along Γ is performed in the counterclockwise direction. As is well known, (3.6) defines an analytic function of ε in $B_{\gamma'}$ [9] and the result of the theorem follows. □

Theorem 3.4. *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 linear operator on $\mathcal{H}$. Then, the Magnus series $\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t)$, with Ω_k given by (1.6) converges in the interval $t \in [0, T)$ such that*

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

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 3.1 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 (1.6), 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 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. □

3.2. Examples

Theorem 3.4 thus 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 $\sum_{k\geqslant 1}\Omega_k(t)$ for $0 \leqslant t < T$ when the terms $\Omega_k(t)$ are computed with (1.6). A natural question arising here is the following: is the bound estimate provided by theorem 3.4 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 [27] consider the initial-value problem (1.1) with

$$A(t) = \begin{pmatrix} 2 & t \\ 0 & -1 \end{pmatrix}. \tag{3.7}$$

If we introduce, as before, the complex parameter ε in the problem, the corresponding exact solution $Y(t;\varepsilon)$ of (3.1) is given by

$$Y(t;\varepsilon) = \begin{pmatrix} \mathrm{e}^{2\varepsilon t} & \frac{1}{9\varepsilon}\,\mathrm{e}^{2\varepsilon t} - \left(\frac{1}{9\varepsilon} + \frac{1}{3}t\right)\mathrm{e}^{-\varepsilon t} \\ 0 & \mathrm{e}^{-\varepsilon t} \end{pmatrix} \tag{3.8}$$

and therefore

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

The Magnus series can be obtained by computing the Taylor expansion of $\log Y(t;\varepsilon)$ around $\varepsilon = 0$. Note 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, the 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 3.4.

Example 2. Let us introduce the matrices

$$X_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad X_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{3.9}$$

and define

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

with α, β complex constants. Then, the solution of equation (1.1) at $t = 2$ is $Y(2) = \mathrm{e}^{\alpha X_1}\,\mathrm{e}^{\beta X_2}$, so that

$$\Omega(2) = \log(\mathrm{e}^{\alpha X_1}\,\mathrm{e}^{\beta X_2}) = \alpha X_1 + \frac{2\alpha\beta}{1-\mathrm{e}^{-2\alpha}} X_2, \tag{3.10}$$

an analytic function if $|\alpha| < \pi$ with first singularities at $\alpha = \pm\,\mathrm{i}\pi$.

On the other hand, a simple calculation with the recurrence (1.6) shows that

$$\Omega(2) = \sum_{k=1}^{\infty}\Omega_k(2) = \alpha X_1 + \beta X_2 + \sum_{n=2}^{\infty}(-1)^{n-1}\frac{2^{n-1}B_{n-1}}{(n-1)!}\alpha^{n-1}\beta X_2. \tag{3.11}$$

Comparing with expression (3.10), it is clear that the Magnus series cannot converge at $t = 2$ if $|\alpha| \geqslant \pi$, independently of $\beta \neq 0$.

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 3.4 is $|\alpha| + |\beta| < \pi$ for this example.

From the analysis of examples 1 and 2 we can conclude the following. First, the convergence domain of the Magnus series provided by theorem 3.4 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 $A(t)$, 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 $r_c = \pi$ is still rather conservative: *the Magnus series converges indeed for a larger time interval than that given by theorem* 3.4. This is particularly evident if one considers equation (3.1) with $A(t)$ a diagonal matrix,

$$A(t) = \begin{pmatrix} a_1(t) & 0 \\ 0 & a_2(t) \end{pmatrix}. \tag{3.12}$$

Then, the exact solution $Y(t;\varepsilon)$ of (3.1) 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.

4. Another characterization of the convergence of the Magnus series

4.1. Main result on convergence

The examples collected in the preceding section (and many more one can build) clearly show that, although the condition $\int_0^t \|A(s)\| \, \mathrm{d}s < \pi$ is sharp (in the sense that the constant π is the largest number for which theorem 3.4 holds in general), it is certainly not necessary for the convergence of the Magnus series. Thus, it would be highly desirable to have a more realistic criterion which give both necessary and sufficient conditions for convergence.

In [27], a conjecture is formulated, relating the convergence of the Magnus series with the eigenvalues of the exact solution $Y(t;\varepsilon)$. Here we state a theorem which, on the one hand, explains the phenomena observed by Moan and Niesen [27] and, on the other hand, provides a new tool to determine the actual convergence domain of the Magnus series in some physically relevant examples and applications.

The main result in this section (theorem 4.2) is valid for complex $n \times n$ matrices and is based on the theory of analytic matrix functions, in particular, in the logarithm of an analytic matrix function. In fact, it is a direct consequence of the analysis done in ([32], chapter 1, section 3). Here we shall summarize the most relevant aspects of the formalism and refer the reader to [32] for a more detailed treatment (including proofs).

Our starting point is again the initial-value problem $Y' = \varepsilon A(t) Y$, $Y(0) = I$, where now $A(t)$ and Y are (complex) $n \times n$ matrices and $\varepsilon \in \mathbb{C}$. If we denote by $Y_t(\varepsilon)$ the exact solution for a fixed value of t, $Y_t(\varepsilon) \equiv Y(t;\varepsilon)$, it is clear that $Y_t(\varepsilon)$ is an analytic function of ε [7], since the Neumann series

$$Y_t(\varepsilon) = I + \sum_{k=1}^{\infty} \varepsilon^k \int_0^t \mathrm{d}t_1 A(t_1) \int_0^{t_1} \mathrm{d}t_2 A(t_2) \cdots \int_0^{t_{k-1}} \mathrm{d}t_k A(t_k)$$

converges provided that $\int_0^t \|A(s)\| \, \mathrm{d}s < \infty$. In addition, $\det Y_t(\varepsilon) \neq 0$ for all ε. Under these conditions, it has been shown that the matrix $\Omega_t(\varepsilon) = \log Y_t(\varepsilon)$ is also an analytic function

of ε at $\varepsilon = 0$. In other words, the series $\Omega_t(\varepsilon) = \sum_{k\geqslant 1} \varepsilon^k \Omega_{t,k}$ (i.e., the Magnus series) is convergent for sufficiently small ε. The goal is then to determine the actual radius of convergence r of this series.

Let us denote by $\rho_1(\varepsilon), \ldots, \rho_n(\varepsilon)$ the eigenvalues of the matrix $Y_t(\varepsilon)$. Note that $Y_t(0) = I$, so that $\rho_1(0) = \cdots = \rho_n(0) = 1$. It is therefore natural to take the principal values of the logarithm, $\log \rho_1(0) = \cdots = \log \rho_n(0) = 0$, as this choice is consistent with the series $\Omega_t(\varepsilon)$.

Let L be a curve on the ε plane in the disk $|\varepsilon| < r_0 < \infty$ issuing from the origin. Recall that the matrix $Y_t(\varepsilon)$ is analytic in the disk $|\varepsilon| < r_0$. On the curve L it is possible to define a unique function $\log \rho_j(\varepsilon)$, $j = 1, \ldots, n$, by continuity, given the values $\log \rho_j(0) = 0$.

Let ρ_0 be a multiple eigenvalue of $Y_t(\varepsilon_0)$ for some ε_0 with $|\varepsilon_0| < r_0$ with multiplicity l. If we reorder the eigenvalues of $Y_t(\varepsilon_0)$ in such a way that the first l are precisely ρ_0, it is clear that the numbers $\log \rho_1(\varepsilon_0), \log \rho_2(\varepsilon_0), \ldots, \log \rho_l(\varepsilon_0)$, $1 < l \leqslant n$, are congruent modulo 2π i and are such that $\rho_1(\varepsilon_0) = \cdots = \rho_l(\varepsilon_0) = \rho_0$. Associated with this multiple eigenvalue ρ_0 there is a pair of integers (p, q) defined as follows.

The integer p is the greatest number of equal terms in the set of numbers $\log \rho_1(\varepsilon_0), \log \rho_2(\varepsilon_0), \ldots, \log \rho_l(\varepsilon_0)$ such that $\rho_k(\varepsilon_0) = \rho_0$, $k = 1, \ldots, l$.

The integer q is the maximum degree of the elementary divisors $(\rho - \rho_0)^k$ of $Y_t(\varepsilon_0)$, i.e., the maximum dimension of the elementary Jordan block corresponding to ρ_0.

Note that the numbers l and q depend only on the particular eigenvalue ρ_0, whereas the integer p depends on ρ_0 and the curve L.

Under these conditions, it is possible to prove the following lemma [32, p 64] on the convergence of the series $\Omega_t(\varepsilon)$.

Lemma 4.1. *(Yakubovich–Starzhinskii). Suppose that the series* $\Omega_t(\varepsilon) = \sum_{k\geqslant 1} \varepsilon^k \Omega_{t,k}$ *satisfies that* $\exp \Omega_t(\varepsilon) = Y_t(\varepsilon)$ *for sufficiently small* $|\varepsilon|$. *Then*

(a) If $r < r_0$ is the radius of convergence of the series $\Omega_t(\varepsilon)$, the eigenvalues $\lambda_1(\varepsilon), \ldots, \lambda_n(\varepsilon)$ of the matrix $\Omega_t(\varepsilon)$, defined for $|\varepsilon| < r$, can be defined by continuity on the circle $|\varepsilon| = r$, and there exists a point ε_0 such that for some $j, k = 1, \ldots, n$

$$\lambda_j(\varepsilon_0) - \lambda_k(\varepsilon_0) = 2\pi\, \mathrm{i} m,$$

where $m \neq 0$ is an integer.

(b) Suppose that ε_0 is the value of ε of smallest absolute value ($\varepsilon_0 \neq 0$, $|\varepsilon_0| < r_0$) such that the matrix $Y_t(\varepsilon_0)$ has an eigenvalue ρ_0 of multiplicity $l > 1$. Suppose that there is at least one such an eigenvalue ρ_0 and at least one curve in the disk $|\varepsilon| < |\varepsilon_0|$ joining the origin $\varepsilon = 0$ with the point $\varepsilon = \varepsilon_0$ such that $p < q$, where the integers p and q have been defined before. Then $r = |\varepsilon_0|$ is the radius of convergence of the series $\Omega_t(\varepsilon) = \sum_{k\geqslant 1} \varepsilon^k \Omega_{t,k}$.

In order to apply this result one first has to solve the equation

$$\Delta(\varepsilon) = 0, \tag{4.1}$$

where $\Delta(\varepsilon)$ denotes the discriminant of the characteristic polynomial $\det(Y_t(\varepsilon) - \rho I)$. We recall here that the discriminant of a polynomial

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

is given by

$$a_n^{2n-2} \prod_{i<j} (r_i - r_j)^2,$$

with $r_1, \ldots, r_n$ complex roots of $p(x)$, so that it vanishes if and only if $p(x)$ has one or more multiple roots [20]. Thus, it can be used to test for the presence of multiple roots, without having to actually compute the roots of $p(x)$.

We write the solutions of equation (4.1) in order of non-decreasing absolute value,

$$\varepsilon_0^{(1)}, \varepsilon_0^{(2)}, \varepsilon_0^{(3)}, \ldots \tag{4.2}$$

and consider the circle $|\varepsilon| = \left|\varepsilon_0^{(1)}\right|$ in the complex ε-plane. Let $\rho_0^{(1)}$ denotes an eigenvalue of $Y_t(\varepsilon_0^{(1)})$ with multiplicity $l_1 > 1$. Let ε move along some fixed curve L from $\varepsilon = 0$ to $\varepsilon = \varepsilon_0^{(1)}$ in the circle $|\varepsilon| \leqslant |\varepsilon_0^{(1)}|$. Then it is clear that l_1 eigenvalues $\rho_j(\varepsilon)$ will tend to $\rho_0^{(1)}$ at $\varepsilon = \varepsilon_0^{(1)}$. If these points lie at $\varepsilon = \varepsilon_0^{(1)}$ on the same sheet of the Riemann surface of the function $\log z$, and this is true for all (possible) multiple eigenvalues of $Y_t(\varepsilon)$ at $\varepsilon = \varepsilon_0^{(1)}$, then $\varepsilon_0^{(1)}$ is called a *extraneous root* of equation (4.1). Otherwise, $\varepsilon_0^{(1)}$ is called a *non-extraneous root*.

Now, by lemma 4.1, when $|\varepsilon| < \left|\varepsilon_0^{(1)}\right|$, the series for $\Omega_t(\varepsilon)$ is convergent, so that the numbers $\log \rho_j(\varepsilon)$ are uniquely determined up to multiplicity as eigenvalues of the matrix $\Omega_t(\varepsilon)$.

If $\varepsilon_0^{(1)}$ is an extraneous root, there is no obstacle to the convergence of the series and thus we proceed to the next value in the sequence (4.2). We continue this classification until a non-extraneous root is obtained. Assume, for simplicity, that $\varepsilon_0^{(2)}$ is the first non-extraneous root.

The root $\varepsilon_0^{(2)}$ will generally correspond to some multiple eigenvalue ρ_0 of $Y_t\left(\varepsilon_0^{(2)}\right)$, with integers (p, q) as before. Then the statement of lemma 4.1 can be formulated as follows.

Theorem 4.2. *If $r \neq \infty$ is the radius of convergence of the series*

$$\Omega_t(\varepsilon) = \sum_{k=1}^{\infty} \varepsilon^k \Omega_{t,k}, \tag{4.3}$$

there is at least one non-extraneous root ε_0 of the equation $\Delta(\varepsilon) = 0$ on the circle $|\varepsilon| = r$. If for this root one has $p < q$ for some corresponding eigenvalue ρ_0 of multiplicity $l > 1$, then $r = |\varepsilon_0|$, i.e., the radius of convergence of the series $\Omega_t(\varepsilon)$ is precisely $|\varepsilon_0|$.

We should remark here that in some cases with $p \geqslant q$, the series (4.3) may well converge at $\varepsilon = \varepsilon_0$ and the radius of convergence r is indeed greater than $|\varepsilon_0|$. This occurs, for instance, when $A(t)$ is diagonal. To illustrate this phenomenon, consider again the matrix (3.12) with $a_1(t) \equiv a_2(t)$. Then, clearly, $\rho_1(\varepsilon) = \rho_2(\varepsilon)$ for all ε, so that $l = 2$ and $q = 1$. If we choose $\log \rho_1(\varepsilon) = \log \rho_2(\varepsilon)$, then $p = 2 > q$.

Although these cases are in a certain sense exceptional, as explained in [32, p 66], theorem 4.2 is not yet, strictly speaking, a necessary condition for the convergence of the series (4.3). In any case, the convergence in the diagonal case is compatible with its formulation, as we have seen.

4.2. Examples

We next illustrate theorem 4.2 on three different examples. We first consider those analyzed in subsection 3.2 and then we treat in some detail the Magnus expansion applied to the evolution operator describing a two-level quantum system.

Example 3 (*revisited*)**.** Given the exact solution (3.8) of example 1 in subsection 3.2, the corresponding discriminant is given by $\Delta(\varepsilon) = (\mathrm{e}^{2\varepsilon t} + \mathrm{e}^{-\varepsilon t})^2 - 4\mathrm{e}^{\varepsilon t}$, whose roots are

$$\varepsilon_0^{(1)} = 0, \qquad \text{and} \qquad \varepsilon_0^{(2)} = \mathrm{i}\frac{2\pi}{3t}.$$

The first value, $\varepsilon_0^{(1)} = 0$, is clearly an extraneous root, so we analyze $\varepsilon_0^{(2)}$. As ε varies along the imaginary axis from $\varepsilon = 0$ to $\varepsilon = \varepsilon_0^{(2)}$, the eigenvalues of the matrix $Y_t(\varepsilon)$,

$$\rho_1(\varepsilon) = \mathrm{e}^{2\varepsilon t}, \qquad \rho_2(\varepsilon) = \mathrm{e}^{-\varepsilon t}$$

move along the unit circle, one clockwise and the other counterclockwise from

$$\rho_{1,2}(0) = 1 \quad \text{to} \quad \rho_{1,2}\big(\varepsilon_0^{(2)}\big) = \mathrm{e}^{\mathrm{i}4\pi/3} = \mathrm{e}^{-i2\pi/3} = \mathrm{e}^{\mathrm{i}(4\pi/3-2\pi)}.$$

Thus, $\rho_1\big(\varepsilon_0^{(2)}\big)$ and $\rho_2\big(\varepsilon_0^{(2)}\big)$ lie on different sheets of the Riemann surface of the function $\log z$ and therefore $\varepsilon_0^{(2)}$ is a non-extraneous root, with $p = 1$. Since $Y_t\big(\varepsilon_0^{(2)}\big) \neq \rho I$, we have $q = 2$, so that, according to theorem 4.2, the radius of convergence of the series (4.3) is precisely

$$r = \big|\varepsilon_0^{(2)}\big| = \frac{2\pi}{3t}. \tag{4.4}$$

To get the actual convergence domain of the corresponding Magnus expansion we have to take $\varepsilon = 1$, and so, from (4.4), we get $2\pi/(3t) = 1$, or equivalently $t = 2\pi/3$, i.e., the result achieved from the analysis of the exact solution in subsection 3.2.

Example 2 (revisited)**.** Let us obtain the convergence domain for the Magnus expansion of the solution to the initial-value problem $Y' = \varepsilon A(t)Y$, $Y(0) = I$, when $A(t)$ is the piece-wise continuous matrix defined in example 2 (subsection 3.2). The exact solution for $t \geqslant 1$ is given by

$$Y(t;\varepsilon) = \begin{pmatrix} \mathrm{e}^{\varepsilon w} & \varepsilon\beta\,\mathrm{e}^{\varepsilon w} \\ 0 & \mathrm{e}^{-\varepsilon w} \end{pmatrix},$$

where $w \equiv \alpha(t-1)$. Equation (4.1) leads in this case to $\cosh^2(\varepsilon w) - 1 = 0$, with first solutions

$$\varepsilon = 0, \qquad \varepsilon = \pm\mathrm{i}\frac{\pi}{w}.$$

Again, $\varepsilon = 0$ is an extraneous root, whereas the eigenvalues of the matrix $Y_t(\varepsilon)$ move along the unit circle, one clockwise and the other counterclockwise from

$$\rho_{1,2}(0) = 1 \quad \text{to} \quad \rho_{1,2}(\mathrm{i}\pi/w) = -1,$$

when ε varies along the imaginary axis from $\varepsilon = 0$ to $\varepsilon = \mathrm{i}\pi/w$ (the same considerations apply to the case $\varepsilon = -\mathrm{i}\pi/w$). Then, obviously, $p = 1$ and $q = 2$, so that the radius of convergence of the series (4.3) is

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

If we now fix $\varepsilon = 1$, we get the actual t-domain of convergence of the Magnus series (1.4) as

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

Observe that, when $t = 2$, we get $|\alpha| = \pi$ and the result of subsection 3.2 is recovered: the Magnus series converges only for $|\alpha| < \pi$.

Example 3. Our final illustration corresponds to the quantum-mechanical treatment of a two-level system in a rotating field. It is described by the Hamiltonian

$$H(t) = \tfrac{1}{2}\hbar\omega_0\sigma_z + \beta(\sigma_x\cos\omega t + \sigma_y\sin\omega t), \tag{4.5}$$

where $\sigma_x, \sigma_y, \sigma_z$ are Pauli matrices and β is a coupling constant. In fact, this system constitutes a truncation in state space of a more general one, namely an atom or freely rotating molecule in a circularly polarized radiation field [18, 30].

It has been previously established that when $t = 2\pi/\omega$ the Magnus expansion of the corresponding evolution operator $U(t)$, solution of the Schrödinger equation

$$\mathrm{i}\hbar\frac{\mathrm{d}U}{\mathrm{d}t} = H(t)U, \qquad U(0) = I \tag{4.6}$$

converges for $\omega > \omega_0$ and diverges otherwise [10, 22, 30]. Several different arguments have been offered trying to explain this phenomenon [18]. Here we show that this bound can be directly provided by theorem 4.2.

The exact time-evolution operator can be obtained in closed form by transforming into a rotating frame. Replacing, as usual, H by εH in (4.6) one has

$$U(t) = \exp\left(-\frac{1}{2}\,\mathrm{i}\omega t\sigma_z\right)\exp\left(-\mathrm{i}t\left(\frac{1}{2}(\varepsilon\omega_0 - \omega)\sigma_z + \varepsilon\frac{\beta}{\hbar}\sigma_x\right)\right). \tag{4.7}$$

From (4.7), a lengthy but straightforward calculation allows us to write the corresponding matrix $Y(t;\varepsilon) \equiv U(t)$ in the form

$$Y_t(\varepsilon) = \begin{pmatrix} \mathrm{e}^{-\frac{1}{2}\mathrm{i}t\omega}\left(\cos\frac{\tilde\omega t}{2} - \mathrm{i}\frac{\delta}{\tilde\omega}\sin\frac{\tilde\omega t}{2}\right) & -\mathrm{i}\,\mathrm{e}^{-\frac{1}{2}\mathrm{i}t\omega}\frac{2\varepsilon\beta}{\tilde\omega\hbar}\sin\frac{\tilde\omega t}{2} \\ -\mathrm{i}\,\mathrm{e}^{\frac{1}{2}\mathrm{i}t\omega}\frac{2\varepsilon\beta}{\tilde\omega\hbar}\sin\frac{\tilde\omega t}{2} & \mathrm{e}^{\frac{1}{2}\mathrm{i}t\omega}\left(\cos\frac{\tilde\omega t}{2} + \mathrm{i}\frac{\delta}{\tilde\omega}\sin\frac{\tilde\omega t}{2}\right) \end{pmatrix}, \tag{4.8}$$

with $\delta = \varepsilon\omega_0 - \omega$ and $\tilde\omega = (\delta^2 + 4\beta^2\varepsilon^2/\hbar^2)^{1/2}$. Denoting

$$\begin{aligned}\cos\chi &\equiv \cos\frac{\omega t}{2}\cos\frac{\tilde\omega t}{2} - \frac{\delta}{\tilde\omega}\sin\frac{\omega t}{2}\sin\frac{\tilde\omega t}{2} \\ &= \frac{1}{2}\left(1 + \frac{\delta}{\tilde\omega}\right)\cos\frac{(\omega+\tilde\omega)t}{2} + \frac{1}{2}\left(1 - \frac{\delta}{\tilde\omega}\right)\cos\frac{(\omega-\tilde\omega)t}{2},\end{aligned}$$

the eigenvalues of $Y_t(\varepsilon)$ can be expressed as

$$\rho_{1,2}(\varepsilon) = \cos\chi \pm \sqrt{\cos^2\chi - 1}, \tag{4.9}$$

so that $U(t)$ has multiple eigenvalues when $\cos\chi = \pm 1$. This equality is satisfied by $\varepsilon = 0$, which is clearly an extraneous root. The remaining roots of equation (4.1) are obtained from

$$\arccos\left(\cos\frac{\omega t}{2}\cos\frac{\tilde\omega t}{2} - \frac{\delta}{\tilde\omega}\sin\frac{\omega t}{2}\sin\frac{\tilde\omega t}{2}\right) = \pi. \tag{4.10}$$

To simplify the discussion, let us consider the perturbative approximation $\beta \ll \hbar|\delta|/2$. Then $\tilde\omega \approx |\delta|$ and $\cos\chi \approx \cos(\varepsilon\omega_0 t/2)$, so that (4.10) reduces to $\varepsilon\omega_0 t/2 = \pi$. The solution

$$\varepsilon_0 = \frac{2\pi}{\omega_0 t}$$

is a non-extraneous root with $p = 1$ and $q = 2$ ($\rho_{1,2}(\varepsilon_0) = -1$), and thus the radius of convergence of the series (4.3) is precisely $|\varepsilon_0|$. Taking now $\varepsilon = 1$, we get finally the t-domain of convergence of the Magnus expansion $t_c = 2\pi/\omega_0$. Note that for $t = t_c$ and $\omega < \omega_0$ we are outside the convergence disk, and thus the Magnus series diverges, just as noted in [18].

5. Discussion

The Magnus expansion was originally designed by requiring only that $A(t)$ be a linear operator depending on a real variable t in an associative ring and that 'certain unspecified conditions of convergence be satisfied' [21]. The idea was to define, in terms of A, an operator $\Omega(t)$ such that the solution of the initial-value problem

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

for a second operator Y is given as $Y = \exp \Omega$. The proposed expression for Ω was an infinite series satisfying the condition that 'its partial sums become Hermitian after multiplication by i if iA is a Hermitian operator' [21]. The simplest example of an equation of this type is given by a finite system of linear differential equations. In this case, $A(t)$ is the coefficients matrix of the system, and the existence of Ω is assured 'for sufficiently small values of t' [21]. Theorem 2.1 yields an implicit condition in terms of the eigenvalues of the matrix Ω.

Given the importance of the expansion, it has been rediscovered a number of times in different settings along the years. Also a particular attention has been paid to its convergence in the matrix case, and so several bounds on the actual radius of convergence of the form (2.2) have been obtained with different values of r_c. Recently, it has been shown that the optimal value is $r_c = \pi$ for all $n \times n$ real matrices $A(t)$ [27].

In this work, by applying standard techniques of complex analysis and some elementary properties of the unit sphere, we have generalized this result to bounded linear operators in a Hilbert space (theorem 3.4), in the spirit of the original Magnus formulation in the context of quantum mechanics. Obviously, this theorem is also valid for finite-dimensional complex matrices. In our treatment, a complex parameter ε is introduced in the formalism, so that the initial-value problem (3.1) is considered instead. Note that the Magnus expansion is trivially recovered as soon as we fix $\varepsilon = 1$.

Although theorem 3.4 provides the optimal convergence domain, in the sense that π is the largest constant for which the result holds without any further restrictions on the operator $A(t)$, one can easily construct examples showing that $\int_0^T \|A(s)\| \, \mathrm{d}s < \pi$ is *not* necessary for the convergence of the expansion.

With the aim of obtaining a more precise characterization of the convergence, we have considered in section 4 the case of $n \times n$ complex matrices. There, as a straightforward consequence of the theory of analytic matrix functions, and in particular, of the logarithm of an analytic matrix function such as is done in [32], we have established a connection between the convergence of the Magnus series and the existence of multiple eigenvalues of the fundamental matrix $Y(t; \varepsilon)$ for a fixed t, denoted by $Y_t(\varepsilon)$ (theorem 4.2). In essence, 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 $\Omega_t(\varepsilon) = \sum_{k \geqslant 1} \varepsilon^k \Omega_{t,k}$ verifying $\exp \Omega_t(\varepsilon) = Y_t(\varepsilon)$ is precisely $r = |\varepsilon_0|$.

This value r in general will be different for each particular t considered, so that we can write $r = |\varepsilon_0| = F(t)$ for a given function $F(t)$. In particular, for the examples considered in section 4, $F(t) = \frac{2\pi}{3t}$, $F(t) = \frac{\pi}{|\alpha|(t-1)}$ and $F(t) = \frac{2\pi}{\omega_0 t}$, respectively. If we fix $|\varepsilon_0| = 1$, then the convergence t-domain of the Magnus expansion is obtained as the solution of $F(t) = 1$ with the smallest absolute value.

It is interesting at this point to discuss theorem 2.1 on the existence of a differentiable function $\Omega(t)$ and theorem 3.4 in view of the more precise account on the convergence issue provided by theorem 4.2. First, note that under the assumptions of theorem 3.4, all the eigenvalues of $Y(t)$ lie in the region

$$G_\pi = \{z = |z| \, \mathrm{e}^{\mathrm{i}\omega} \in \mathbb{C} : \mathrm{e}^{-\pi} \leqslant |z| \leqslant \mathrm{e}^{\pi}, |\omega| < \pi\},$$

so that automatically all the differences between any two of the eigenvalues of $\Omega(t) = \log Y(t)$ is less than $2\pi\mathrm{i}$ and thus theorem 2.1 holds. Second, if all eigenvalues of $Y_t(\varepsilon)$ are located in G_π, the (possible) multiple eigenvalues take place only at extraneous roots of the parameter ε and thus, according with theorem 4.2, the convergence of the series is assured.

One might think that the practical application of theorem 4.2 to realistic problems is doubtful, since it is necessary to compute in advance the fundamental matrix $Y(t;\varepsilon)$. In this sense, the alternative (but more conservative) estimate provided by theorem 3.4 directly in terms of the operator $A(t)$ is certainly easier to check in practice. In our opinion, however, the characterization of the convergence of the Magnus expansion in terms of the multiple eigenvalues of $Y(t;\varepsilon)$ sheds new light on this issue, has a theoretical interest by itself and, in addition, provides a rigorous justification for the conjecture formulated in [27] on the basis of the exploration of several examples.

Acknowledgments

The author is grateful to Professor W So for providing him the reference [23] and to Professor J A Oteo for useful discussions. This work has been partially supported by Ministerio de Educación y Ciencia (Spain) under project MTM2004-00535 (co-financed by the ERDF of the European Union) and Fundació Bancaixa.

References

[1] Abramowitz M and Stegun I A 1965 *Handbook of Mathematical Functions* (New York: Dover)
[2] Blanes S, Casas F, Oteo J A and Ros J 1998 Magnus and Fer expansions for matrix differential equations: the convergence problem *J. Phys. A: Math. Gen.* **22** 259–68
[3] Blanes S, Casas F, Oteo J A and Ros J 2007 Magnus expansion: mathematical study and physical applications, work in progress
[4] Blanes S, Casas F and Ros J 2000 Improved high order integrators based on the Magnus expansion *BIT* **40** 434–50
[5] Blanes S, Casas F and Ros J 2002 High order optimized geometric integrators for linear differential equations *BIT* **42** 262–84
[6] Chacon R V and Fomenko A T 1991 Recursion formulas for the Lie integral *Adv. Math.* **88** 200–57
[7] Coddington E A and Levinson N 1955 *Theory of Ordinary Differential Equations* (New York: McGraw-Hill)
[8] Dixmier J 1957 L'application exponentielle dans les groupes de Lie résolubles *Bull. Soc. Math. France* **85** 113–21
[9] Dunford N and Schwartz J T 1958 *Linear Operators. Part I: General Theory* (New York: Wiley-Interscience)
[10] Fel'dman E B 1984 On the convergence of the Magnus expansion for spin systems in periodic magnetic fields *Phys. Lett.* A **104** 479–81
[11] Gantmacher F R 1959 *The Theory of Matrices* (New York: Chelsea)
[12] Gronwall T H 1919 Note on the derivatives with respect to a parameter of the solutions of a system of differential equation *Ann. Math.* **20** 292–6
[13] Hairer E, Lubich Ch and Wanner G 2006 *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations* 2nd edn (Berlin: Springer)
[14] Hunter J K and Nachtergaele B 2001 *Applied Analysis* (Singapore: World Scientific)
[15] Iserles A, Munthe-Kaas H Z, Nørsett S P and Zanna A 2000 Lie-group methods *Acta Numer.* **9** 215–365
[16] Iserles A and Nørsett S P 1999 On the solution of linear differential equations in Lie groups *Phil. Trans. R. Soc.* A **357** 983–1019
[17] Karasev M V and Mosolova M V 1977 Infinite products and T products of exponentials *J. Theor. Math. Phys.* **28** 721–30
[18] Klarsfeld S and Oteo J A 1989 Analytic properties of the Magnus operator for two solvable Hamiltonians *Phys. Lett.* A **142** 393–7
[19] Klarsfeld S and Oteo J A 1989 Recursive generation of higher-order terms in the Magnus expansion *Phys. Rev.* A **39** 3270–3273
[20] Kurosh A G 1972 *Higher Algebra* (Paris: Mir)
[21] Magnus W 1954 On the exponential solution of differential equations for a linear operator *Commun. Pure Appl. Math.* **VII** 649–73
[22] Maricq M M 1987 Convergence of Magnus expansion for time-dependent two-level systems *J. Chem. Phys.* **86** 5647–51
[23] Mityagin B S 1990 Unpublished notes

[24] Moan P C 1998 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

[25] Moan P C, Oteo J A and Ros J 1999 On the existence of the exponential solution of linear differential systems *J. Phys. A: Math. Gen.* **32** 5133–9

[26] Moan P C 2002 On backward error analysis and Nekhoroshev stability in the numerical analysis of conservative systems of ODEs *PhD Thesis* University of Cambridge

[27] Moan P C and Niesen J 2006 Convergence of the Magnus series *Technical Report* La Trobe University

[28] Pechukas P and Light J C 1966 On the exponential form of time-displacement operators in quantum mechanics *J. Chem. Phys.* **44** 3897–912

[29] Saito M 1957 Sur certain groupes de Lie résolubles: I, II *Sci. Pap. College Gen. Educ. Univ. Tokyo* **7** 1–11

[30] Salzman W R 1986 Convergence of Magnus and Magnus-like expansions in the Schrödinger representation *J. Chem. Phys.* **85** 4605–13

[31] Wei J 1963 Note on global validity of the Baker–Hausdorff and Magnus theorems *J. Math. Phys.* **4** 1337–41

[32] Yakubovich V A and Starzhinskii V M 1975 *Linear Differential Equations with Periodic Coefficients* (New York: Wiley)