On the existence of the exponential solution of linear differential systems

1999 J. Phys. A: Math. Gen. 32 5133

(http://iopscience.iop.org/0305-4470/32/27/311)

J. Phys. A: Math. Gen. **32** (1999) 5133–5139. Printed in the UK PII: S0305-4470(99)03209-6

On the existence of the exponential solution of linear differential systems

P C Moan†, J A Oteo‡ and J Ros‡§

† Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge CB3 9EW, UK
‡ Departament de Física Teòrica. Universitat de València, Avda. Dr Moliner 50, 46100-Burjassot, València, Spain
§ IFIC, Avda. Dr Moliner 50, 46100-Burjassot, València, Spain

Received 1 April 1999

Abstract. The existence of an exponential representation for the fundamental solutions of a linear differential system is approached from a novel point of view. A sufficient condition is obtained in terms of the norm of the coefficient operator defining the system. The condition turns out to coincide with a previously published one concerning convergence of the Magnus series expansion. Direct analysis of the general evolution equations in the $SU(N)$ Lie group illustrates how the estimate for the domain of existence/convergence becomes larger. Eventually, an application is done for the Baker–Campbell–Hausdorff series.

1. Introduction

The so-called Magnus expansion (ME) is an elegant way to approximately solve the linear operator initial value problem:

$$\frac{dZ}{dt} = A(t)Z \qquad Z(0) = I. \tag{1}$$

For our purposes it suffices to consider A as a complex matrix of dimension $n \times n$ whose matrix elements are integrable functions of t, but the analysis given in this paper applies more generally to bounded operators in a Banach algebra. Here I stands for the identity matrix of dimension n. If we recall that a possible form to write the solution corresponding to a constant operator A reads $Z(t) = \exp(At)$, one may wonder about the advantages of seeking a solution of equation (1) in the exponential form $Z(t) = \exp(\Omega(t))$, $\Omega(0) = 0$. The ME provides one such possibility by taking Ω as the expansion

$$\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t) \tag{2}$$

where the terms are expressed as nested commutators. For the sake of illustration the first two contributions read

$$\Omega_1(t) = \int_0^t A(\tau)\, d\tau \qquad \Omega_2(t) = \frac{1}{2}\int_0^t [A(\tau), \Omega_1(\tau)]\, d\tau. \tag{3}$$

The relevant point for the usefulness of such an exponential representation is that any truncation of (2) leads to an approximate solution of $Z(t)$ which necessarily preserves some

0305-4470/99/275133+07$30.00

intrinsic property of the exact solution. For instance, it is well known that $\det(Z(t)) = \exp(\int_0^t \mathrm{tr}(A(\tau))\,\mathrm{d}\tau$, which is indeed a property shared by a truncated ME. In many problems we are interested in solutions of equation (1) which evolve in t on a Lie group $\mathcal{G}$. As we build Ω with all the terms in the series (2) belonging to the Lie algebra $\mathfrak{g}$ associated to $\mathcal{G}$, any approximation to Z obtained by truncating (2) is still an element of $\mathcal{G}$, provided $A \in \mathfrak{g}$. This feature of ME has been exploited in a number of areas of physics where equation (1) controls the time evolution of a system either in a classical or quantum treatment. A representative list of applications can be found in [1]. For instance, when A is skew-Hermitian, which may correspond to the Schrödinger equation in quantum mechanics, then every exponential approximation to Z is certainly unitary as required by first principles. More recently, ME has been proposed [2] as a qualitatively correct and quantitatively very accurate source of numerical integrators of linear differential systems on Lie groups.

Whereas the use of ME has spread among various branches of science, there are still some fundamental problems open which have not been faced up in the same measure. The purpose of this paper is to shed light on these matters. Actually, we can distinguish two intertwined questions in the Magnus approach for solving the initial value problem (1). First, for what values of t and for what operators A does equation (1) admit a true exponential solution (the existence problem). Second, for what values of t and for what linear operators, A, does the series in equation (2) converge (the convergence problem). More precisely, we obtain conditions on $A(t)$ defining a t-domain where the exponential representation of $Z(t)$ is guaranteed; i.e., we address our study to the first item above. In section 2 we develop a short proof to determine such a condition. Section 3 addresses the second item above. In it, we slightly generalize the recursive proof of convergence of ME in [1]. It turns out that the existence and convergence conditions obtained coincide (up to numerical precision). In view of this striking result, one could be tempted to attribute further significance to the t-domains obtained. For this reason in section 4 we analyse in detail the ME for equation (1) in the particular case of the $su(N)$ Lie algebra. Eventually, in section 5 an application of this result is given to estimate a sufficient condition for convergence of the well known Baker–Campbell–Hausdorff (BCH) series.

2. Existence of Ω

Introducing the form $Z(t) = \exp(\Omega(t))$ into equation (1) one obtains the nonlinear differential equation for Ω (see for instance [3–5])

$$\dot{\Omega} = \frac{\mathrm{ad}_\Omega}{\exp(\mathrm{ad}_\Omega) - 1} A = \sum_{j=0}^{\infty} \frac{B_j}{j!}\,\mathrm{ad}_\Omega^j A \qquad \Omega(0) = 0. \tag{4}$$

Here the dot stands for derivative with respect to t, B_j are Bernoulli numbers [6], and we introduce the adjoint operator: $\mathrm{ad}_\Omega^0 A = A$, $\mathrm{ad}_\Omega(A) = [\Omega, A] \equiv \Omega A - A\Omega$, $\mathrm{ad}_\Omega^j A = \mathrm{ad}_\Omega(\mathrm{ad}_\Omega^{j-1} A)$.

Magnus himself [3] gathers in his original paper the following formal result. The exponential representation exists for a sufficiently small interval of t. For whenever a couple of eigenvalues of Ω, say $\lambda_k(t)$, $\lambda_j(t)$, satisfy the *Magnus condition* $\lambda_k(t) - \lambda_j(t) = 2\pi\mathrm{i}$, then $\dot{\Omega}(t)$ becomes singular in equation (4) because $\lambda_k(t) - \lambda_j(t)$ are precisely the instantaneous eigenvalues of ad_Ω. Alas, that existence theorem has little practical application. But indeed it is the regularity of the exponential map from the Lie algebra to the Lie group that determines the global properties of the Magnus expansion. Dixmier [7] studied the problem globally and gave conditions for the surjectivity of the map. Dixmier's results were completed to some extent by the work of Saito [8]: 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 $\mathrm{ad}_{x_1}\,\mathrm{ad}_{x_2}\ldots\mathrm{ad}_{x_{n-1}}\,x_n = 0$, where x_j are arbitrary elements from the Lie algebra. For solvable real algebras the mapping is surjective if and only if the algebra contains no subalgebra isomorphic to the algebra spanned by the elements x, y, z with the commutator rule $[x, y] = z$, $[x, z] = -y$. As a matter of fact, this guarantees that the eigenvalues of ad_Ω are always real, and hence the Magnus condition on eigenvalues is never satisfied. This paper is intended to extract information about the existence of Ω directly from A.

Equation (4) admits the integral form

$$\Omega(t) = \sum_{j=0}^{\infty} \frac{B_j}{j!} \int_0^t \mathrm{ad}^j_{\Omega(s)}\, A(s)\, \mathrm{d}s \tag{5}$$

provided Ω is sufficiently small, or is e.g., nilpotent. We are interested in finding for which values of t the exponential representation of $Z(t)$ does exist. It is clear in the nilpotent case that the Magnus expansion will give a global representation, hence we will in the following consider only the non-nilpotent case. Now, assume that the Lie algebra generated by $A(t)$ is endowed with a norm, $\|\cdot\|$. With this in mind, the basic commutator will be bounded as

$$\|[\Omega, A]\| \leqslant 2\mu\|\Omega\|\|A\| \tag{6}$$

which implies for the nested commutator

$$\|\,\mathrm{ad}^n_\Omega\, A\| \leqslant (2\mu)^n\|\Omega\|^n\|A\| \tag{7}$$

where $\mu \in [0, 1]$ is introduced as a real factor incorporating any additional knowledge we may have on A useful to tighten the bound. If no further information is added then $\mu = 1$. The norm appearing in the expressions above is as yet unspecified but has to satisfy the *consistency condition* or *submultiplicative property* $\|AB\| \leqslant \|A\|\|B\|$, so that $\|[A, B]\| \leqslant 2\|A\|\|B\|$ be implied.

Using the triangle inequality repeatedly we can obtain from equation (5)

$$\|\Omega(t)\| \leqslant \int_0^t \sum_{j=0}^{\infty} \frac{|B_j|}{j!}(2\mu\|\Omega(s)\|)^j\|A(s)\|\,\mathrm{d}s. \tag{8}$$

The series inside the integral sign can be summed up using

$$g(x) \equiv \sum_{j=0}^{\infty} \frac{|B_j|}{j!}(2x)^j = 2 + x(1 - \cot(x)) \tag{9}$$

which is a positive nondecreasing function in the domain $D \equiv [0, \pi)$.

Thus, if we denote $\|A(s)\| \equiv k(s)$ we arrive at

$$\|\Omega(t)\| \leqslant \int_0^t g(\mu\|\Omega(s)\|)k(s)\,\mathrm{d}s \equiv f(t). \tag{10}$$

From here it is obvious that $\dot{f}(t) = g(\mu\|\Omega(t)\|)k(t)$, and since g is nondecreasing in D, $\dot{f}(t) \leqslant g(\mu f(t))k(t)$. Now the positive character of g in D allows us to write $\dot{f}(t)/g(\mu f(t)) \leqslant k(t)$, and by integration

$$\frac{1}{\mu}\int_0^{\mu f(t)} \frac{\mathrm{d}x}{g(x)} \leqslant \int_0^t k(s)\,\mathrm{d}s \tag{11}$$

provided $\mu f(t) \leqslant \pi$. This is indeed the case if t is such that

$$K(t) \equiv \int_0^t k(s)\mathrm{d}s \leqslant \frac{1}{\mu}\int_0^{\pi} \frac{\mathrm{d}x}{g(x)} \tag{12}$$

due again to the fact that $g > 0$ and consequently $\int_0^y \mathrm{d}z/g(z)$ is a strictly increasing function of y. The value ξ of the integral appearing in the above equation can be determined numerically and the final result reads

$$\int_0^t \|A(s)\| \,\mathrm{d}s \leqslant \xi/\mu = 1.086\,869/\mu. \tag{13}$$

For those values of t satisfying inequality (13) the existence of Ω is then ensured. This statement may depend of the norm chosen. As far as we are dealing with a sufficiency condition for existence of Ω it is enough to find a particular norm satisfying (13).

3. Convergence of the Magnus expansion

Some lower bounds to the radius of convergence of the ME, i.e. equation (2), in terms of A, can be found in the literature [2, 9]. Here we shall be concerned with the result given in [1]. For the sake of completeness we shall sketch and slightly generalize (in order to incorporate the parameter μ above) the induction proof in [1].

Substituting the Magnus series (2) into (4) and gathering terms of the same order the following recursive procedure holds [4]:

$$\begin{aligned}
&\Omega_1 = \int_0^t A(\tau)\,\mathrm{d}\tau \qquad \Omega_n = \sum_{j=1}^{n-1} \frac{B_j}{j!} \int_0^t S_n^{(j)}(\tau)\,\mathrm{d}\tau \qquad n \geqslant 2 \\
&S_n^{(j)} = \sum_{m=1}^{n-j} [\Omega_m, S_{n-m}^{(j-1)}] \qquad 2 \leqslant j \leqslant n-1 \\
&S_n^{(1)} = [\Omega_{n-1}, A] \qquad S_n^{(n-1)} = \mathrm{ad}_{\Omega_1}^{n-1} A.
\end{aligned} \tag{14}$$

Following [1] it is possible to show by induction that (notice the presence of the factor μ)

$$\|S_n^{(j)}(t)\| \leqslant (\mu K(t))^{n-1} k(t) f_n^{(j)} \tag{15}$$

provided the coefficients $f_n^{(j)}$ are obtained from

$$f_n^{(j)} = 2 \sum_{m=1}^{n-j} \sum_{p=0}^{m-1} \frac{|B_p|}{p!m} f_m^{(p)} f_{n-m}^{(j-1)} \tag{16}$$

with $f_1^{(0)} = 1$, $f_n^{(0)} = 0$, for $n > 1$. Thus, the norm of the terms in the Magnus series is bounded as

$$\|\Omega_n(t)\| \leqslant \left[\frac{1}{n} \sum_{p=1}^{n-1} \frac{|B_p|}{p!} f_n^{(p)} \right] (\mu K(t))^n. \tag{17}$$

The series $\sum_{k=1}^{\infty} \|\Omega_k(t)\|$ is then bounded by a power series in $\mu K(t)$ and a numerical application of the D'Alembert criterion of convergence directly leads (up to numerical precision) to equation (13). The fact that both schemes of bounding refer to existence and convergence, respectively, poses the question about a possible especial meaning of (13): does equation (13) provide us with the best possible lower bound for the existence/convergence radius of ME, in absence of any other additional information?

4. Existence of the Magnus operator in $su(N)$

Let us suppose $A \in su(N)$, the set of skew-Hermitian traceless matrices, so that the fundamental matrix Z belongs to the Lie group of unitary matrices, $SU(N)$. This is a familiar case in the study of spin dynamics in quantum mechanics. We shall present the case $N = 2$ in detail first, and then the general $su(N)$ case.

4.1. The su(2) case

The skew-Hermitian traceless matrices A and Ω may be expressed in terms of the Pauli matrices: $\vec{\sigma} = \{\sigma_x, \sigma_y, \sigma_z\}$, which satisfy $[\sigma_x, \sigma_y] = 2\mathrm{i}\sigma_z$, and circular permutations. Thus, we write $\Omega = -\mathrm{i}\vec{\chi}\cdot\vec{\sigma}$ and $A = -\mathrm{i}\vec{a}\cdot\vec{\sigma}$, with $\vec{\chi}(t)$ and $\vec{a}(t)$ real vector valued functions. For two generic matrices $A, B \in su(2)$ we have then $[A, B] = 2\mathrm{i}(\vec{a}\wedge\vec{b})\cdot\vec{\sigma}$, in terms of the vector product. Using the trace norm, $\|A\| = \sqrt{\mathrm{tr}(AA^\dagger)} = a\sqrt{2}$, where $a = \|\vec{a}\|$, we get $\|[A, B]\| = 2\sqrt{2(a^2b^2 - \vec{a}\cdot\vec{b})} = 2\sqrt{2}ab|\sin\theta| \leqslant 2\sqrt{2}ab$, where $0 \leqslant \theta \leqslant \pi$ stands for the angle between $\vec{a}$ and $\vec{b}$. If we compare with the more naive bound $\|[A, B]\| \leqslant 2\|A\|\|B\| = 4ab$ we conclude that for $su(2)$ we can at least take $\mu = 1/\sqrt{2}$, and this leads to $K(t) \leqslant \sqrt{2}\xi = 1.537\,064$ in equation (13), which enlarges the existence/convergence domain. It may still be pertinent to ask about the accuracy of the lower bound extracted in this manner.

Despite the complete solution for Ω is not known, we shall see that some interesting information arises by taking into account that ad_Ω is itself a linear operator. Therefore equation (4) does have a matrix representation, which in the case at hand is of dimension three (the dimension of the Lie algebra).

The relevant nested commutators in equation (4) are now given by

$$\mathrm{ad}_\Omega^n A = -(-1)^n \chi^n (\hat{\chi}\wedge\vec{a})\cdot\vec{\sigma} \qquad (n > 0) \tag{18}$$

where $\hat{\chi} \equiv \vec{\chi}/\chi$. After some straightforward algebra we obtain the nonlinear system of differential equations

$$\frac{\mathrm{d}\vec{\chi}}{\mathrm{d}t} = \vec{a} + \vec{\chi}\wedge\vec{a} + (1 - \chi\cot\chi)[\hat{\chi}\wedge(\hat{\chi}\wedge\vec{a})]. \tag{19}$$

The derivative of $\vec{\chi}$ becomes singular unless $\chi(t) < \pi$. It corresponds to the Magnus condition for the existence of an exponential representation since the eigenvalues of $\Omega(t)$ are $\pm\mathrm{i}\chi(t)$. As it stands this result has no practical application. However, this equation still provides us with a new lower bound for the radius of convergence. Equation (19) and the identity $\dot{\chi} = \hat{\chi}\cdot\mathrm{d}\vec{\chi}/\mathrm{d}t$ yield $\dot{\chi} = \vec{a}\cdot\hat{\chi}$. Consequently,

$$\chi(t) = \int_0^t \vec{a}(\tau)\cdot\hat{\chi}(\tau)\,\mathrm{d}\tau \leqslant \int_0^t a(\tau)\,\mathrm{d}\tau. \tag{20}$$

We then conclude that for $su(2)$ the existence of Ω is ensured if

$$\int_0^t a(\tau)\,\mathrm{d}\tau < \pi. \tag{21}$$

Instead, the bound (13) to $su(2)$ provides us with the existence/convergence domain:

$$\int_0^t a(\tau)\,\mathrm{d}\tau \leqslant \xi\sqrt{2} = 1.537\,064 \tag{22}$$

with $\mu = 1/\sqrt{2}$. These results have to be compared with the crudest estimate obtained with $\mu = 1$, which is

$$\int_0^t a(\tau)\,\mathrm{d}\tau \leqslant \xi = 1.086\,869. \tag{23}$$

4.2. The su(N) case

Suppose the general case $A \in su(N)$. Thus, $\Omega^\dagger = -\Omega$ and $\dot{\Omega}^\dagger = -\dot{\Omega}$. We use the Frobenius norm $\|\Omega\|^2 = \mathrm{tr}(\Omega^\dagger\Omega)$, which is derived from the scalar product $\langle A|B\rangle = \mathrm{tr}(A^\dagger B)$. By differentiating

$$2\|\Omega\|\frac{\mathrm{d}}{\mathrm{d}t}\|\Omega\| = \langle\dot{\Omega}|\Omega\rangle + \langle\Omega|\dot{\Omega}\rangle \tag{24}$$

and using equation (4) we get $\|\Omega\| \frac{d}{dt} \|\Omega\| = \mathrm{tr}(\Omega^\dagger A)$. By Cauchy–Swartz inequality, we have $\|\Omega\| \frac{d}{dt} \|\Omega\| \leqslant \|A\| \|\Omega\|$, and by integrating

$$\|\Omega(t)\| \leqslant \int_0^t \|A(x)\| \, dx. \tag{25}$$

Since Ω is diagonalizable we have, for the Frobenius norm, the bound $\max_k |\lambda_k| \leqslant \|\Omega\|$ in terms of the eigenvalues λ_k $(k = 1 \ldots N)$. Hence if

$$\int_0^t \|A(x)\| \, dx < \pi \tag{26}$$

we are guaranteed that the eigenvalues of Ω satisfy $\max_{k,l} |\lambda_k - \lambda_l| < 2\pi$, and by Magnus condition on eigenvalues Ω does exist.

5. Application to the BCH formula

The so-called BCH formula allows us to obtain (as an infinite series) the logarithm of the product of two noncommuting exponentials: $\log(\exp(X)\exp(Y)) \equiv W = \sum_{k=1}^{\infty} W_k$ (see, for instance, [3]). Since the BCH formula can be easily extracted from ME, equation (13) above gives a condition either for the existence of W or for BCH series convergence in terms of $\|X\|, \|Y\|$.

Suppose $A(t)$ in equation (1) is given by

$$A(t) = \begin{cases} X & t \in [0, 1] \\ Y & t \in (1, 2]. \end{cases} \tag{27}$$

Then the exact solution of (1) at $t = 2$ reads

$$Z(2) = \exp(Y)\exp(X). \tag{28}$$

If, in turn, we look for $Z(2)$ through ME, we have $W = \Omega(2)$. Then the Magnus series, equation (2), provides us with a representation of the BCH series in terms of nested commutators

$$\exp(Y)\exp(X) = \exp\left[\sum_{k=1}^{\infty} \Omega_k(2)\right]. \tag{29}$$

It is unnecessary to say that this representation in terms of nested commutators is by no means unique, due to the Jacobi identity. However, the interesting feature of this expansion is that equation (13) furnishes now a sufficient condition on X, Y so as to ensure existence/convergence, namely

$$\|X\| + \|Y\| \leqslant \xi. \tag{30}$$

We note in passing that ME is sometimes referred to as the continuous analogue to BCH expansion. If we think of equation (27) as a particular type of discontinuity in $A(t)$ then ME is certainly a kind of generalization of BCH formula for continuous $A(t)$.

6. Conclusions

In [1] the existence of the exponential representation for $Z(t)$ is taken for granted and the study addresses the convergence of the expansion (2). Instead, here we do not make use of the expansion (2) at any stage of the proof of (12). The analysis then addresses the existence of a solution $Z(t)$ in an exponential form. We have proved a sufficient condition for the existence of this representation in terms of the norm of the linear operator $A(t)$ defining the differential system. An additional and striking conclusion is that the domain of existence we find coincides with the domain of convergence of the Magnus series found in [1].

In the process of estimating the domain of existence we have found that the treatment of the basic commutator plays an important role. Obviously, the best possibility corresponds to evaluating it in an exact form. Alternatively, we have introduced a parameter μ which conveys the effect of additional information on the algebra. The sequence of equations (21)–(23) illustrates these facts. The lower bound when $\mu = 1$ is by far the poorest one, which is quite natural. As far as we have some new information about the algebra involved we might increase the lower bound and consequently enlarge the domain of existence. In this respect, we have built up a basic way of carrying it out via the introduction of the parameter μ (see equation (7)) This feature is clearly seen from equations (22), (23) in the example in section 4.1 where the existence domain is enlarged by replacing $\mu = 1$ by $\mu = 1/\sqrt{2}$. The situation improves drastically when no bound at all is taken on commutators. This is the case corresponding to equation (21). There the norm of commutators is exactly taken into account since the very beginning, and therefore the bounding involved, is tighter than in previous procedures.

Given the limited number of results concerning the existence of Ω and the convergence of $\sum_{k=1}^{\infty} \Omega_k$, the results presented in this paper can be of utility in the numerous applications of ME, either as a perturbative tool or as a numerical integrator. An illustration of this is given by the example of the convergence analysis of the BCH formula.

Acknowledgments

This work originated from discussions with A Iserles to whom the authors are deeply grateful. JAO was partially supported by DGICyT, Spain, under project No PB97/1227, Spanish MEC and British Council-Valencia. The work of JR was partially supported by DGICyT, Spain (grant No PB87/1139) and Generalitat Valenciana (grant POST98-CB-10-81).

References

[1] 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

[2] 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

[3] Magnus W 1954 On the exponential solution of differential equations for a linear operator *Commun. Pure Appl. Math.* **7** 649–73

[4] Klarsfeld S and Oteo J A 1989 Recursive generation of higher-order terms in the Magnus expansion *Phys. Rev.* A **39** 3270–3

[5] Wilcox R M 1967 Exponential operators and parameter differentiation in quantum physics *J. Math. Phys.* **8** 962–82

[6] Abramowitz M A and Stegun I A 1965 *Handbook of Mathematical Functions* (New York: Dover)

[7] Dixmier J 1957 L'application exponentielle dans les groupes de Lie résolubles *Bull. Soc. Math. France* **85** 113–21

[8] Saito M 1957 Sur certains groupes de Lie résolubles, I; II *Sci. Papers Coll. Gen. Educ. Univ. Tokyo* **7** 1–11

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