The Baker-Campbell-Hausdorff formula and the convergence of the Magnus expansion



1989 J. Phys. A: Math. Gen. 22 4565

(http://iopscience.iop.org/0305-4470/22/21/018)

J. Phys. A: Math. Gen. **22** (1989) 4565–4572. Printed in the UK

# The Baker–Campbell–Hausdorff formula and the convergence of the Magnus expansion

S Klarsfeld and J A Oteo

Division de Physique Théorique†, Institut de Physique Nucléaire, 91406 Orsay Cedex, France

Received 5 June 1989

**Abstract.** We show that for a wide class of dynamical systems (described by Hamiltonians of the form usually considered in time-dependent perturbation theory) the divergence of the Magnus expansion in the Schrödinger picture for large time intervals is due to pole singularities inherent to the Baker–Campbell–Hausdorff formula.

## 1. Introduction

It has been reported recently in the literature [1–4] that for several simple dynamical models the Magnus expansion of the unitary time-displacement operator $U = U(t, 0)$ in the Schrödinger picture does not converge for time intervals larger than the natural periods of the systems. The purpose of the present paper is to trace down the origin of the divergence, and to show that the a similar result indeed holds under much more general conditions.

For completeness we recall that the Magnus expansion [5] formally provides an exponential solution of the time-dependent Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}U = HU \qquad U = I \text{ at } t = 0 \tag{1.1}$$

where $H = H(t)$ is the Hamiltonian of the system. While the existence of an exponential solution $U = \exp(\Omega)$, with $\Omega = \Omega(t, 0)$, is ensured for some sufficiently short time interval, the general conditions for the convergence of the Magnus expansion are still under debate. Of special interest for applications are Hamiltonians of the form

$$H = H_0 + V(t) \tag{1.2}$$

where $H_0$ and $V(t)$ are two non-commuting operators, the first of which is time-independent and has a known spectrum. Dynamical systems of this type are frequently dealt with in time-dependent perturbation theory. Salzman [2, 3] has investigated the convergence problem for special choices of $H_0$ and $V(t)$. He succeeded in calculating directly many terms in the Magnus expansion of $\Omega$ in the Schrödinger picture to first order in $V$, from which he was able to infer the form of the general term and to determine the convergence radius. Recursive methods for generating higher-order

† Unité de Recherche des Universités Paris 11 et Paris 6, associée au CNRS.

0305-4470/89/214565+08$02.50  

terms can be of great help in such calculations [6, 7]. On the other hand, in a few cases the Magnus operator can be obtained exactly in closed form by using Lie algebra methods [5, 8], and it appears that the lack of convergence is related to the presence of poles in the $t$ plane [9].

In order to understand the full significance of Salzman's results for dynamical systems of this kind it is, however, both simpler and more instructive to start with the interaction picture ($\Omega_{\mathrm{I}}$), and then to build up $\Omega_{\mathrm{S}}$ by a unitary transformation. This may be done by using a special form of the so-called Baker–Campbell–Hausdorff (BCH) formula, which is derived in the appendix. In § 2 the method is applied to two model systems previously discussed by Salzman [2, 3] and yields closed-form expressions for the Magnus operator $\Omega_{\mathrm{S}}$ to first order in $V$. At the same time these examples pave the road for consideration of more general systems. This is done in § 3, where we show that the presence of poles in the Magnus operator $\Omega_{\mathrm{S}}$ occurs for any Hamiltonian of the form (1.2). Finally § 4 contains a few concluding remarks.

## 2. Two simple models

In order to illustrate the virtues of our method we shall first re-examine two simple physical systems belonging to the class (1.2): (i) the forced harmonic oscillator (HO), and (ii) the driven two-level system. The Hamiltonians for these systems are respectively

$$H = \hbar\omega a^{\dagger} a + \beta f(t)(a^{\dagger} + a) \tag{2.1}$$

where $a^{\dagger}$, $a$ are the raising and lowering operators, and

$$H = \tfrac{1}{2}\hbar\omega\sigma_z + \beta f(t)\sigma_x \tag{2.2}$$

where $\sigma_x$, $\sigma_z$ are Pauli matrices. Here $f(t)$ contains the time dependence, $\beta$ is a coupling constant and $\hbar\omega$ defines the internal energy scale of the system.

Let us denote by $U_{\mathrm{S}}$ and $U_{\mathrm{I}}$ the time-displacement operators in the Schrödinger and in the interaction pictures respectively. It is readily verified that

$$U_{\mathrm{S}} = \exp(\tilde{H}_0 t)\, U_{\mathrm{I}} \qquad \tilde{H}_0 = H_0/\mathrm{i}\hbar. \tag{2.3}$$

Introducing the associated Magnus operator in each picture, one therefore has

$$\exp(\Omega_{\mathrm{S}}) = \exp(\tilde{H}_0 t)\, \exp(\Omega_{\mathrm{I}}). \tag{2.4}$$

The above equation may be solved by means of the BCH formula [5, 8, 10, 11]. As shown in the appendix (cf (A2) and (A12)), this gives

$$\Omega_{\mathrm{S}} = \tilde{H}_0 t + \sum_{n=0}^{\infty} (-1)^n \frac{B_n t^n}{n!} \{\tilde{H}_0^n, \Omega_{\mathrm{I}}\} + \mathrm{O}(\Omega_{\mathrm{I}}^2) \tag{2.5}$$

where $B_n$ are Bernoulli numbers and the symbol $\{\ldots\}$ represents a multiple commutator defined in (A3).

On the other hand, it is obvious that to lowest order in $\beta$ one simply has $\Omega_{\mathrm{I}} \simeq \Omega_{1\mathrm{I}}$, where

$$\Omega_{1\mathrm{I}} = \int_0^t \mathrm{d}t'\, \tilde{H}_{\mathrm{I}}(t') \qquad \tilde{H}_{\mathrm{I}} = H_{\mathrm{I}}/\mathrm{i}\hbar \tag{2.6}$$

is the first term in the Magnus expansion, and

$$H_{\mathrm{I}} = \exp(-\tilde{H}_0 t)\, V(t)\, \exp(\tilde{H}_0 t) \tag{2.7}$$

is the Hamiltonian in the interaction picture. Using (A14) one finds after a bit of algebra

$$\Omega_{1\mathrm{I}} = \frac{\beta}{\mathrm{i}\hbar}[P_+(t)a^\dagger + P_-(t)a] \tag{2.8}$$

for the forced HO, and

$$\Omega_{1\mathrm{I}} = \frac{\beta}{\mathrm{i}\hbar}[P_x(t)\sigma_x + P_y(t)\sigma_y] \tag{2.9}$$

for the two-level system. Here

$$P_\pm = P_x \mp \mathrm{i}P_y = \int_0^t \mathrm{d}t'\, f(t')\exp(\pm\mathrm{i}\omega t'). \tag{2.10}$$

Notice that the form of the function $f(t)$ has not yet been specified.

We must now substitute (2.8) and (2.9) into (2.5). Taking into account the commutation relations satisfied by $a$, $a^\dagger$ and by the Pauli matrices it is easy to show that one has

$$\{\tilde{H}_0^{2k}, \Omega_{1\mathrm{I}}\} = (\mathrm{i}\omega)^{2k}\Omega_{1\mathrm{I}} \qquad k = 0, 1, 2, \ldots \tag{2.11}$$

for both systems. Hence (2.5) can be rewritten as

$$\Omega_{\mathrm{S}} = \tilde{H}_0 t + \Omega_{1\mathrm{I}} \sum_{k=0}^{\infty} \frac{B_{2k}}{(2k)!}(\mathrm{i}\omega t)^{2k} - B_1 t[\tilde{H}_0, \Omega_{1\mathrm{I}}] + \mathrm{O}(\beta^2). \tag{2.12}$$

The above expression allows us to calculate $\Omega_{\mathrm{S}}$ only on a restricted time interval because the series diverges for $\omega t > 2\pi$. This follows simply from the fact that the generator function for the Bernoulli numbers, namely

$$\frac{z}{\mathrm{e}^z - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n \tag{2.13}$$

has pole singularities at $z = \pm 2\pi N\mathrm{i}$, $N \neq 0$. Since $B_{2k+1} = 0$ for $k > 0$, the series in (2.12) differs from that in (2.13) by just one term, so that it may be readily summed up in closed form giving

$$\Omega_{\mathrm{S}} = \tilde{H}_0 t + \tfrac{1}{2}\omega t \cot(\tfrac{1}{2}\omega t)\Omega_{1\mathrm{I}} + \tfrac{1}{2}t[\tilde{H}_0, \Omega_{1\mathrm{I}}] + \mathrm{O}(\beta^2). \tag{2.14}$$

More compact forms are further obtained by specialising the last formula for each of the systems under scrutiny. Thus for system (i) one finds

$$\Omega_S \simeq \tilde{H}_0 t + \frac{\beta}{i\hbar}\,\frac{\frac{1}{2}\omega t}{\sin(\frac{1}{2}\omega t)}[\exp(\tfrac{1}{2}i\omega t)P_- a + \exp(-\tfrac{1}{2}i\omega t)P_+ a^\dagger] \tag{2.15}$$

and for system (ii) equation (2.14) becomes

$$\Omega_S \simeq \tilde{H}_0 t + \frac{\beta}{i\hbar}\,\frac{\frac{1}{2}\omega t}{\sin(\frac{1}{2}\omega t)}[\exp(\tfrac{1}{2}i\omega t)P_- \sigma_- + \exp(-\tfrac{1}{2}i\omega t)P_+ \sigma_+] \tag{2.16}$$

where $\sigma_\pm = (\sigma_x \pm i\sigma_y)/2$.

In order to recover Salzman's results it is necessary to determine the functions $P_\pm(t)$ explicitly before expanding the right-hand sides of (2.15) and (2.16) in power series of $t$. For instance, insertion of $f(t) = t$ in (2.10) gives

$$P_\pm = [(1 \mp i\omega t)\exp(\pm i\omega t) - 1]/\omega^2 \tag{2.17}$$

and equations (2.15) and (2.16) become respectively

$$\Omega_S \simeq \tilde{H}_0 t + \frac{\beta t^2}{2i\hbar}[a + a^\dagger + i\chi(\tfrac{1}{2}\omega t)(a - a^\dagger)] \tag{2.18}$$

$$\Omega_S \simeq \tilde{H}_0 t + \frac{\beta t^2}{2i\hbar}[\sigma_x + \chi(\tfrac{1}{2}\omega t)\sigma_y] \tag{2.19}$$

where $\chi(z) = \cot z - 1/z$. The Taylor expansion of $\chi(z)$, which is related to (2.13) as seen above, then immediately yields the expressions derived by Salzman (cf equations (9*a*), (9*b*), (10*a*), (10*b*) in [3]). We have checked in the same way his results for $f(t) = t^2$.

## 3. The general case

The two dynamical systems discussed in § 2 were rather special in that $H_0$ had a unique characteristic frequency (equally spaced levels in the HO case). It is, however, easy to extend the result to more complex dynamical systems. To show how this works, we consider now a general Hamiltonian of type (1.2). In practice the interaction term usually takes the form $V(t) = \sum f_m(t)V_m$, where $V_m$ are also time-independent operators like $H_0$, but this will not be assumed here. All we need are the eigenvectors of $H_0$ with their corresponding eigenvalues $E_1, E_2, \ldots$. Let $X$ be any other (possibly time-dependent) operator. In the basis defined by the eigenvectors of $H_0$ one has

$$[H_0, X]_{kl} = (E_k - E_l)X_{kl} \tag{3.1}$$

and more generally (cf (A3))

$$\{H_0^n, X\}_{kl} = (E_k - E_l)^n X_{kl} \qquad n = 0, 1, 2, \ldots. \tag{3.2}$$

Notice that the diagonal matrix elements of all these commutators vanish. From (3.2) we get in particular

$$\{\tilde{H}_0^n, \Omega_I\}_{kl} = (-i\omega_{kl})^n(\Omega_I)_{kl} \qquad \omega_{kl} = (E_k - E_l)/\hbar \tag{3.3}$$

which when substituted in (2.5) leads to

$$(\Omega_S)_{kl} = -(iE_k t/\hbar)\delta_{kl} + (\Omega_I)_{kl} \sum_{n=0}^{\infty} \frac{B_n t^n}{n!}(i\omega_{kl})^n + \ldots. \tag{3.4}$$

For diagonal matrix elements the sum on the right-hand side of (3.4) is equal to 1, so that $(\Omega_S)_{kk} \simeq -iE_k t/\hbar + (\Omega_I)_{kk}$. In the off-diagonal case we use (2.13) to rewrite (3.4) as

$$(\Omega_S)_{kl} \simeq \frac{i\omega_{kl}t}{\exp(i\omega_{kl}t) - 1}(\Omega_I)_{kl} \qquad k \neq l. \tag{3.5}$$

To first order with respect to $V$, one has simply to replace $\Omega_I$ by $\Omega_{1I}$ in (3.5). By using (2.6) and (2.7) the matrix elements of the latter are readily found to be

$$(\Omega_{1I})_{kl} = (i\hbar)^{-1} \int_0^t dt'\, V_{kl}(t') \exp(i\omega_{kl}t'). \tag{3.6}$$

Thus, unless $(\Omega_{1I})_{kl} = 0$, the corresponding matrix element of the Magnus operator in the Schrödinger picture displays poles for $\omega_{kl}t = \pm 2\pi N$, $N \neq 0$ (in special circumstances some of the poles, or all of them, might be cancelled by the vanishing of the integral, e.g. when $V(t) =$ constant). Generally speaking the convergence radius of the Magnus expansion will therefore be determined by the largest transition frequency for which $V_{kl}$ does not vanish.

The results obtained in § 2 readily follow as particular cases of this general statement. For the two-level system this is completely trivial. For the HO one should simply recall that in the basis in which the energy is diagonal the operators $a$ and $a^\dagger$ have non-zero matrix elements only between adjacent levels.

## 4. Conclusions

In the present paper we have investigated the properties of the Magnus operator as a function of time for a wide class of Hamiltonians. By using the BCH formula we were able to derive closed-form expressions for the contributions to order $V$ to the matrix elements of this operator in the Schrödinger picture. The off-diagonal elements exhibit poles in the $t$ plane, which explains why the Magnus expansion of $\Omega_S$ does not converge for large $t$. The radius of convergence of the expansion restricted to terms of first order in $V$ is explicitly determined in the general case.

A remarkable feature of our method is that the sum which is responsible for the divergence factorises in a manner manifestly independent of the specific form and strength of the perturbation. It should be mentioned, however, that similar convergence difficulties occur for Hamiltonians which do not contain a time-independent 'unperturbed' part [12]. It is obvious that the method used here does not apply in such cases.

Finally it is important to notice that in some problems exponential product representations of $U_S$ might offer a useful alternative to the Magnus expansion [8, 13, 14]. A striking example is (2.4), which was our starting point.

**Acknowledgments**

One of the authors (JAO) thanks the Division de Physique Théorique de l'I.P.N. for its hospitality, and the Ministerio de Educacion y Ciencia (Spain) for financial support under the programme MEC/MRES. We are indebted to Dr M Ginocchio for calling our attention on Goldberg's paper.

**Appendix.**

Given two non-commuting operators $X$, $Y$, the famous BCH theorem states that the operator $Z$ defined by $\exp X \exp Y = \exp Z$ can be expressed as a sum of $X$, $Y$ and their multiple commutators. To fourth order one finds

$$Z = X + Y + \tfrac{1}{2}[X,Y] + \tfrac{1}{12}[X,[X,Y]] + \tfrac{1}{12}[Y,[Y,X]] - \tfrac{1}{24}[X,[Y,[X,Y]]] + \ldots . \qquad \text{(A1)}$$

The fifth-order terms were given by Magnus [5] (see also [11]), while Richtmyer and Greenspan [15] extended the expansion up to order ten by using a computer.

The general BCH expansion becomes rapidly unwieldy, but it is not too difficult to obtain explicitly all the terms which are linear in one of the operators, say $Y$. The shortest way is to use Hausdorff's method [5, 11] which yields $Z = X + Z_1 + Z_2 + \ldots$, where $Z_n$ contains all the terms of degree $n$ with respect to $Y$. For the linear part, $Z_1$, one finds

$$Z_1 = Y\hat{X}(\mathrm{e}^{\hat{X}} - 1)^{-1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} Y\hat{X}^n = \sum_{n=0}^{\infty} (-1)^n \frac{B_n}{n!} \hat{X}^n Y \qquad \text{(A2)}$$

where $B_n$ are Bernoulli numbers (e.g. $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_3 = B_5 = \ldots = 0$) and $\hat{X}$ is the linear (adjoint) operator defined by $\hat{X}Y = -Y\hat{X} = [X,Y]$. Hence $\hat{X}^n Y$ and $Y\hat{X}^n$ are multiple commutators :

$$\begin{aligned} \hat{X}^n Y &= \{X^n, Y\} = \underbrace{[X,[\ldots[X}_{n},Y]\ldots]] \\ Y\hat{X}^n &= \{Y, X^n\} = [[\ldots[Y,\underbrace{X]\ldots],X]}_{n} = (-1)^n \hat{X}^n Y. \end{aligned} \qquad \text{(A3)}$$

The curly bracket symbol introduced above has the advantage of emphasising the commutator structure of the terms. Note also that the standard mathematical notation for the adjoint operator is $\mathrm{ad}\,X$ rather than $\hat{X}$.

An alternative derivation (which has some pedagogical value on its own) is based on two important results due to Goldberg [16] who investigated the formal power series expansion of $Z$. Denoting the coefficient of the product $X^{s_1} Y^{s_2} \ldots X^{s_m} (Y^{s_m})$ for $m$ odd (even) by $C_X(s_1, \ldots, s_m)$ one has according to his theorem 1,

$$C_X(s_1, \ldots, s_m) = \int_0^1 \mathrm{d}t\, t^{m'} (t-1)^{m''} G_{s_1}(t) \ldots G_{s_m}(t) \qquad \text{(A4)}$$

where $m' = [m/2]$, $m'' = [(m-1)/2]$, and the polynomials $G_k(t)$ are defined recursively by

$$kG_k(t) = (\mathrm{d}/\mathrm{d}t)[t(t-1)G_{k-1}(t)] \qquad G_1(t) = 1. \tag{A5}$$

Here we assume $s_1, \ldots, s_m \neq 0$. Similarly, the coefficient of the product $Y^{s_1}X^{s_2}\ldots Y^{s_m}(X^{s_m})$ is written as $C_Y(s_1, \ldots, s_m)$ and is given by $C_Y = (-1)^{n-1}C_X$, where $n = \sum_{i=1}^{m} s_i$. The second result we need is Goldberg's theorem 3 which provides for the coefficients of the product $X^pY^q$ the remarkable formula:

$$C_X(p,q) = \frac{(-1)^p}{p!q!}\sum_{k=1}^{q}\binom{q}{k}B_{p+q-k}. \tag{A6}$$

Let us specialise now for terms of the form $X^{n-k}YX^k$ where $n > 0$ . When $k = 0$ and $k = n$ we can use (A6) to obtain

$$C_X(n,1) = (-1)^nB_n/n! \qquad C_Y(1,n) = B_n/n!. \tag{A7}$$

On the other hand, on account of the recurrence relation (A5) for $G_k(t)$, repeated partial integration applied to (A4) leads to (here we assume $n \geq 2$)

$$C_X(n-k,1,k) = (-1)^k\binom{n}{k}C_X(n,1) \qquad 0 < k < n. \tag{A8}$$

Thus, collecting all these terms with $n$ fixed in the power series of $Z$ one finds

$$(-1)^n\frac{B_n}{n!}\sum_{k=0}^{n}(-1)^k\binom{n}{k}X^{n-k}YX^k. \tag{A9}$$

The sum in the preceding formula is easily recognised as the expanded form of the multiple commutator $\hat{X}^nY$ defined in (A3). The proof is given by induction. For $n = 0$ and $n = 1$ the result is trivial. Assume now that the sum in (A9) equals $\hat{X}^nY$ for some $n > 1$. From (A3) it then follows

$$\begin{aligned}\hat{X}^{n+1}Y &= [X,\hat{X}^nY] \\ &= X^{n+1}Y + \sum_{k=1}^{n}(-1)^k\left[\binom{n}{k}+\binom{n}{k-1}\right]X^{n+1-k}YX^k + (-1)^{n+1}YX^{n+1}.\end{aligned} \tag{A10}$$

Taking into account the relation

$$\binom{n}{k}+\binom{n}{k-1}=\binom{n+1}{k} \tag{A11}$$

completes the proof. To summarise, we see again that to first order in $Y$ the general BCH formula can be written as

$$Z = X + \sum_{n=0}^{\infty}(-1)^n\frac{B_n}{n!}\{X^n, Y\} + \mathrm{O}(Y^2) \tag{A12}$$

or in the more compact form

$$Z = X - [\mathrm{e}^{-\hat{X}} - 1]^{-1}\hat{X}Y + \mathrm{O}(Y^2). \tag{A13}$$

For the reader's convenience we recall here another important formula relating to exponential operators. This is [8]

$$\mathrm{e}^{X}\, Y\, \mathrm{e}^{-X} = \sum_{n=0}^{\infty} \frac{1}{n!}\hat{X}^n Y = \mathrm{e}^{\hat{X}}\, Y. \tag{A14}$$

The result is expressed again in terms of multiple commutators defined via the adjoint operator.

## References


[1] Fel'dman E B 1984 *Phys. Lett.* **104A** 479
[2] Salzman W R 1986 *Chem. Phys. Lett.* **124** 531
[3] Salzman W R 1987 *Phys. Rev.* A **36** 5074
[4] Matti Maricq M 1987 *J. Chem. Phys.* **86** 5647
[5] Magnus W 1954 *Commun. Pure Appl. Math.* **7** 649
[6] Burum D P 1981 *Phys. Rev.* B **24** 3684
[7] Klarsfeld S and Oteo J A 1989 *Phys. Rev.* A **39** 3270
[8] Wilcox R M 1967 *J. Math. Phys.* **8** 962
[9] Fernández F M 1988 *J. Chem. Phys.* **88** 490
[10] Steinberg S 1986 *Lie Methods in Optics* (*Lecture Notes in Physics* **250**) ed J Sánchez Mondragón and K B Wolf (Berlin: Springer)
[11] Weiss G H and Maradudin A A 1962 *J. Math. Phys.* **3** 771
[12] Fernández F M, Echave J and Castro E A 1987 *Chem. Phys.* **117** 101
[13] Wei J and Norman E 1963 *J. Math. Phys.* **4** 575
[14] Klarsfeld S and Oteo J A 1989 *J. Phys. A: Math. Gen.* **22** 2687
[15] Richtmyer R D and Greenspan S 1965 *Commun. Pure Appl. Math.* **18** 107
[16] Goldberg K 1956 *Duke. Math. J.* **23** 13