On some identities involving exponentials

Fernando Casas
Fernando.Casas@mat.uji.es
www.gicas.uji.es/Fernando/fernando.html

Departament de Matemàtiques
Institut de Matemàtiques i Aplicacions de Castelló (IMAC)
Universitat Jaume I
Castellón

ICMAT, May 2015

Based on work done (along the years) with

- Ander Murua
- Sergio Blanes
- José Ros
- José Ángel Oteo

0. INTRODUCTION

- $A = \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix}$
- $Y' = AY, \; Y(0) = I$
- Solution: $Y(t) = \mathrm{e}^{At}$

$$\mathrm{e}^{tA} = \sum_{k \geq 0} \frac{t^k}{k!} A^k = \begin{pmatrix} 2 - \mathrm{e}^{-t} & -1 + \mathrm{e}^{-t} \\ 2 - 2\mathrm{e}^{-t} & -1 + 2\mathrm{e}^{-t} \end{pmatrix}$$

- Main object: exponential of a matrix
- Basic property of the exponential of a matrix (of dimension $N \geq 2$):

$$\mathrm{e}^{A}\,\mathrm{e}^{B} \neq \mathrm{e}^{A+B} \qquad \text{in general} \quad AB \neq BA$$

- Only if $AB = BA$ it is true that $\mathrm{e}^{A}\,\mathrm{e}^{B} = \mathrm{e}^{A+B}$.

A trivial example

- Consider

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

- A and B do not commute:

$$AB = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad BA = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

- Hence we have

$$\mathrm{e}^A = \begin{pmatrix} \mathrm{e} & \mathrm{e} \\ 0 & \mathrm{e} \end{pmatrix} \qquad \text{and} \qquad \mathrm{e}^B = \begin{pmatrix} \mathrm{e} & 0 \\ \mathrm{e} - 1 & 1 \end{pmatrix}$$

A trivial example

- and

$$e^A e^B = \begin{pmatrix} e & e \\ 0 & e \end{pmatrix} \begin{pmatrix} e & 0 \\ e-1 & 1 \end{pmatrix} = \begin{pmatrix} 2e^2 - e & e \\ e^2 - e & e \end{pmatrix}$$

  whereas

$$e^{(A+B)} = e^{3/2} \begin{pmatrix} c+s & 2s \\ 2s & c-s \end{pmatrix}$$

  with $c = \cosh \frac{\sqrt{5}}{2}$, $s = \frac{1}{\sqrt{5}} \sinh \frac{\sqrt{5}}{2}$.
- Therefore

$$e^A\, e^B \neq e^{A+B}$$

- Important object: the **commutator** $[A, B] = AB - BA$

Problems

1. $\mathrm{e}^{A}\,\mathrm{e}^{B} = \mathrm{e}^{A+B+C}$
2. $\mathrm{e}^{A+B} = e^{A}\,\mathrm{e}^{B}\,\mathrm{e}^{C_1}\,\mathrm{e}^{C_2}\ldots$
3. Given $Y' = A(t)Y$, $Y(0) = I$, with $A(t)$ a $N \times N$ matrix,
   - $N = 1$: $Y(t) = \mathrm{e}^{\int_0^t A(s)ds}$
   - $N > 1$: $Y(t) = \mathrm{e}^{\int_0^t A(s)ds}$ if

$$\left[A(t), \int_0^t A(s)ds\right] = 0 \quad \text{(Coddington \& Levinson)}$$

   - General case: can we write $Y(t) = \mathrm{e}^{\Omega(t)}$ with

$$\Omega(t) = \int_0^t A(s)ds + (\text{something else})?$$

Motivation

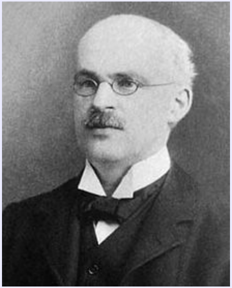
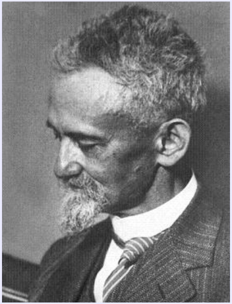
Exponential map:

- Fundamental role played by the exponential transformation in Lie groups and Lie algebras

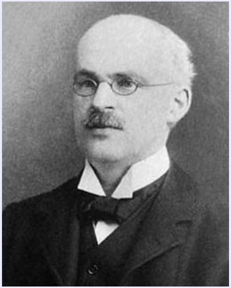
$$\exp : \mathfrak{g} \longmapsto \mathcal{G}$$

- Kashiwara–Vergne conjecture (with important implicaciones in Lie theory, harmonic analysis, etc.), proved as a theorem in 2006
- Lie groups are ubiquitous in physics: symmetries in classical mechanics, Quantum Mechanics, control theory, etc.

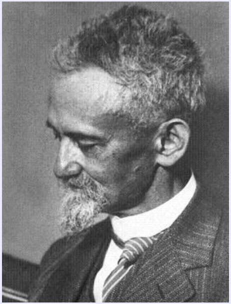
I. MAIN CHARACTERS

H.F. Baker (1866-1956)



J.E. Campbell (1862-1924)



F. Hausdorff (1868-1942)



E.B. Dynkin (1924-2014) W. Magnus (1907-1990) H. Zassenhaus (1912-1991)

Problems

1. $\mathrm{e}^{A}\,\mathrm{e}^{B} = \mathrm{e}^{A+B+C}$ Baker–Campbell–Hausdorff (BCH) Formula
2. $\mathrm{e}^{A+B} = e^{A}\,\mathrm{e}^{B}\,\mathrm{e}^{C_1}\,\mathrm{e}^{C_2}\ldots$ Zassenhaus Formula
3. Given $Y' = A(t)Y$, $Y(0) = I$, $A(t)$ a $N \times N$ matrix,
   - $N = 1$: $Y(t) = \mathrm{e}^{\int_0^t A(s)ds}$
   - $N > 1$: $Y(t) = \mathrm{e}^{\int_0^t A(s)ds}$ if

$$\left[A(t), \int_0^t A(s)ds\right] = 0$$

   - General: can we write $Y(t) = \mathrm{e}^{\Omega(t)}$ with

$$\Omega(t) = \int_0^t A(s)ds + (\text{something else})?$$

Magnus Expansion

These topics have already appeared (several times) at the workshop
We are mainly concerned by

- computational aspects: how to generate *efficiently* the corresponding series
- Convergence of the series

Before starting...

- Lie Product Formula. Let X and Y be $n \times n$ complex matrices. Then

$$\mathrm{e}^{X+Y} = \lim_{m \to \infty} \left(\mathrm{e}^{\frac{X}{m}} \, \mathrm{e}^{\frac{Y}{m}} \right)^m .$$

- A big brother: Trotter product formula. The same result when X and Y are suitable unbounded operators on an infinite-dimensional Hilbert space.

Many applications in the numerical treatment of PDEs

II. BCH FORMULA

$e^X e^Y = e^Z$

- Let X, Y be two non commuting operators. Then

$$e^X e^Y = \sum_{p,q=0}^{\infty} \frac{1}{p!\,q!} X^p Y^q$$

- Substituting this series in the formal series defining the logarithm function

$$\log Z = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (Z-1)^k$$

one gets

$$Z = \log(e^X e^Y) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum \frac{X^{p_1} Y^{q_1} \dots X^{p_k} Y^{q_k}}{p_1!\, q_1! \dots p_k!\, q_k!},$$

- The inner summations extends over all non-negative integers $p_1, q_1, \dots, p_k, q_k$ for which $p_i + q_i > 0$ $(i = 1, 2, \dots, k)$.

- First terms:

$$\begin{aligned} Z &= (X + Y + XY + \frac{1}{2}X^2 + \frac{1}{2}Y^2 + \cdots) \\ &\quad - \frac{1}{2}(XY + YX + X^2 + Y^2 + \cdots) + \cdots \\ &= X + Y + \frac{1}{2}(XY - YX) + \cdots = X + Y + \frac{1}{2}[X, Y] + \cdots \end{aligned}$$

- Baker, Campbell, Hausdorff analyzed whether Z can be written as a series only in terms of (nested) commutators
- The answer is yes, but they weren't able to provide a rigorous proof
- Bourbaki: "*chacun considère que les démonstrations de ses prédécesseurs ne sont pas convaincantes*"
- Finally, Dynkin (1947): explicit formula for Z.
- Sometimes it is called BCH-D (for Dynkin) formula (Bonfiglioli & Fulci)

- Dynkin:

$$Z = \sum_{k=1}^{\infty} \sum_{p_i, q_i} \frac{(-1)^{k-1}}{k} \frac{[X^{p_1} Y^{q_1} \dots X^{p_k} Y^{q_k}]}{\left(\sum_{i=1}^{k}(p_i + q_i)\right) p_1!\, q_1! \dots p_k!\, q_k!} \quad (1)$$

- Inner summation over all non-negative integers $p_1, q_1, \dots, p_k$, q_k for which $p_1 + q_1 > 0, \dots, p_k + q_k > 0$
- $[X^{p_1} Y^{q_1} \dots X^{p_k} Y^{q_k}]$ denotes the right nested commutator based on the *word* $X^{p_1} Y^{q_1} \dots X^{p_k} Y^{q_k}$:

$$[XY^2X^2Y] \equiv [X, [Y, [Y, [X, [X, Y]]]]]$$

- Gathering terms together

$$Z = \log(\mathrm{e}^X \mathrm{e}^Y) = X + Y + \sum_{m=2}^{\infty} Z_m, \quad (2)$$

- $Z_m(X, Y)$: homogeneous Lie polynomial in X, Y of degree m, i.e., a $\mathbb{Q}$-linear combination of commutators of the form $[V_1, [V_2, \dots, [V_{m-1}, V_m] \dots]]$ with $V_i \in \{X, Y\}$ for $1 \le i \le m$.

First terms

$$
\begin{aligned}
Z_2 &= \frac{1}{2}[X, Y] \\
Z_3 &= \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] \\
Z_4 &= -\frac{1}{24}[Y, [X, [X, Y]]] \\
Z_5 &= \frac{1}{720}[X, [X, [X, [X, Y]]]] - \frac{1}{180}[Y, [X, [X, [X, Y]]]] \\
&\quad + \frac{1}{180}[Y, [Y, [X, [X, Y]]]] + \frac{1}{720}[Y, [Y, [Y, [X, Y]]]] \\
&\quad - \frac{1}{120}[[X, Y], [X, [X, Y]]] - \frac{1}{360}[[X, Y], [Y, [X, Y]]]
\end{aligned}
$$

Applications

Fundamental role in different fields:

- Mathematics: theory of linear differential equations, Lie groups, numerical analysis of differential equations
- Theoretical Physics: perturbation theory, quantum mechanics, statistical mechanics, quantum computing
- Control theory: design and analysis of nonlinear control mechanisms, nonlinear filters, stabilization of rigid bodies,...

Applications

Quantum Mechanics

- $i\hbar\dot{U} = HU(t)$, $U(t_0) = I$, so that $\psi(t) = U(t)\psi_0$
- $H(t) = K + V = -\frac{\hbar^2}{2m}p^2 + V$
- Solution: $U(t) = \mathrm{e}^{-iHt/\hbar}$.
- Very often, computing $\mathrm{e}^{-iKt/\hbar}$, $\mathrm{e}^{-iVt/\hbar}$ is easier

Applications

Quantum Monte Carlo methods

- Partition function

$$Z = \mathrm{Tr}(\mathrm{e}^{-\beta H}) = \sum_{\alpha} \langle \alpha | \mathrm{e}^{-\beta H} | \alpha \rangle,$$

  for the orthogonal complete set of states $|\alpha\rangle$. Here $\beta = 1/T$ and $H = K + V$

- All practical implementations intended for Monte Carlo estimations of Z rely on approximating

$$\mathrm{e}^{-\beta(K+V)} = \left(\mathrm{e}^{-\varepsilon(K+V)}\right)^M,$$

  with $\varepsilon = \beta/M$ and M is the number of convolution terms (*beads*).

- Product of exponentials

$$\mathrm{e}^{-\varepsilon(K+V)} \simeq \prod_{i=1}^{m} \mathrm{e}^{-a_i \varepsilon K} \mathrm{e}^{-b_i \varepsilon V}$$

Applications

- Lie groups theory: Lie algebra $\leftrightarrow$ Lie group. Multiplication law in the group is determined uniquely by the Lie algebra structure, at least in a neighborhood of the identity
- Helpful also to prove the existence of a local Lie group with a given Lie algebra
- The particular structure of the series is not very important in this setting...
- ...But in other fields it *is* relevant to analyze the combinatorial aspects and its efficient computation

Applications

- Lie groups theory: Lie algebra $\leftrightarrow$ Lie group. Multiplication law in the group is determined uniquely by the Lie algebra structure, at least in a neighborhood of the identity
- Helpful also to prove the existence of a local Lie group with a given Lie algebra
- The particular structure of the series is not very important in this setting...
- ...But in other fields it *is* relevant to analyze the combinatorial aspects and its efficient computation

Example

- Splitting and composition methods

$$\dot{u} = F(u) = A(u) + B(u), \qquad u(0) = u_0$$

- Flows of $A(u)$ and $B(u)$, e^{tA}, e^{tB} can be obtained explicitly
- Approximation (for $t = h$, the time step)

$$\Psi_h \equiv \exp(ha_1 A) \exp(hb_1 B) \cdots \exp(ha_k A) \exp(hb_k B)$$

- Order conditions to be satisfied by a_i, b_i so that

$$\Psi_h \equiv \exp(ha_1 A) \exp(hb_1 B) \cdots \exp(ha_k A) \exp(hb_k B)$$

verifies $\Psi_h(u_0) = u(h) + \mathcal{O}(h^{r+1})$ when $h \to 0$.

- They are obtained by applying BCH in sequence:

$$\Psi_h = \exp(p_{1,1} A + p_{1,2} B + p_{2,1} [A, B] + \cdots)$$

with $p_{i,j}$ polynomials in a_i, b_i.

- $p_{1,1} = p_{1,2} = 1$, $p_{2,1} = 0$, etc.

How to obtain the BCH formula

- Different procedures in the literature:
  - Goldberg form + Dynkin (Specht-Wever) theorem (explicit)

$$Z = X + Y + \sum_{n=2}^{\infty} \frac{1}{n} \sum_{w,|w|=n} g_w \, [w], \tag{3}$$

with $w = w_1 w_2 \ldots w_n$, each w_i is X or Y, $[w] = [w_1, [w_2, \ldots [w_{n-1}, w_n] \ldots]]$, the coefficient g_w is a rational number and n is the word length.

  - Varadarajan (recursive)

$$\begin{aligned} Z_1 &= X + Y \qquad (4) \\ (n+1) Z_{n+1} &= \frac{1}{2}[X - Y, Z_n] + \sum_{p=1}^{[n/2]} \frac{B_{2p}}{(2p)!} \\ & \quad \sum [Z_{k_1}, [\cdots [Z_{k_{2p}}, X + Y] \cdots]], \quad n \geq 1 \end{aligned}$$

Second sum: over all positive integers such that $k_1 + \cdots + k_{2p} = n$

Software packages

- *Diffman* (Trondheim-Bergen). Matlab toolbox for computations in a free Lie algebra.
- The computation of the BCH formula is carried out in *Diffman* by integrating numerically

$$Z' = d\exp_Z^{-1}(X) \equiv \sum_{k=1}^{\infty} \frac{B_k}{k!} \mathrm{ad}_Z^k X, \quad Z(0) = Y$$

  from $t = 0$ to $t = 1$ using a single step of a Runge–Kutta method. ($\mathrm{e}^{Z(t)} = \mathrm{e}^{tX} \mathrm{e}^{Y}$)
- Koseleff (1993): explicit expression in the Lyndon basis up to $n = 10$ *by using only manipulations of Lie polynomials*, without resorting to the associative algebra

Software packages

- Reinsch (2000): Simple derivation with matrices of rational numbers. Mathematica program. The expression is *not* written in terms of commutators
- A recent (much simpler) modification in 2015
- *Lie Tools Package (LTP)* (Torres-Torriti & Michalska, 2003).
    - Package in Maple for carrying out Lie algebraic symbolic computations.
    - Special function for the computation of BCH formula in the Dynkin form in terms of Lie monomials in the Hall basis.
    - Reported results in 2003: up to order 10 in 25 hours with maximum memory usage of 17.5 Mbytes on a Pentium III, 550 MHz, 256 Mbytes RAM, Maple 7, Linux

Bottleneck in the computation

- The iterated commutators are not all linearly independent, due to the Jacobi identity

$$[X_1,[X_2,X_3]] + [X_2,[X_3,X_1]] + [X_3,[X_1,X_2]] = 0$$

  (and other identities involving nested commutators of higher degree originated by it)
- All the previous expressions for the BCH series are *not* formulated directly in terms of a basis of the free Lie algebra $\mathcal{L}(X,Y)$
- Problematic when designing numerical integrators for ODEs (*one condition per element in the basis*)
- Very difficult to study specific properties of the series: distribution of coefficients, combinatorial properties, etc.

- It is possible to rewrite the formulas in terms of a basis of $\mathcal{L}(X, Y)$, but this process is very time consuming and requires lots of memory resources
- The complexity grows exponentially with m: the number of terms involved in the series grows as the dimension c_m of the homogeneous subspace $\mathcal{L}(X, Y)_m$
- $c_m = \mathcal{O}(2^m/m)$ (Witt's formula)

Goal

- To express the BCH series as

$$Z = \log(\exp(X)\exp(Y)) = \sum_{i \geq 1} z_i\, E_i, \tag{5}$$

where $z_i \in \mathbb{Q}$ $(i \geq 1)$ and $\{E_i \ : \ i = 1, 2, 3, \ldots\}$ is a basis of $\mathcal{L}(X, Y)$ whose elements are of the form

$$E_1 = X, \quad E_2 = Y, \quad \text{and} \quad E_i = [E_{i'}, E_{i''}] \quad i \geq 3, \tag{6}$$

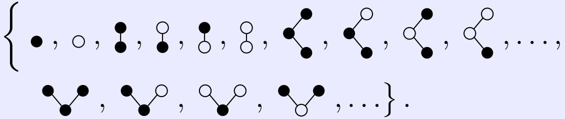
for appropriate values of the integers $i', i'' < i$ $(i = 3, 4, \ldots)$.

- In particular: classical Hall basis, Lyndon basis
- Design an efficient algorithm
- Analyze the series (coefficients, convergence)

Summary of the algorithm

- Starting point: vector space $\mathfrak{g}$ of maps $\alpha : \mathcal{T} \to \mathbb{R}$
- $\mathcal{T}$: set of rooted trees with black and white vertices

$$\mathcal{T} =$$



- $\mathcal{T}$ is typically referred to as the set of labeled rooted trees with two labels, ‘black’ and ‘white’.
- elements of $\mathcal{T}$: bicoloured rooted trees.

- $\mathfrak{g}$ is endowed with a Lie algebra structure by defining the Lie bracket $[\alpha, \beta] \in \mathfrak{g}$, of two arbitrary maps $\alpha, \beta \in \mathfrak{g}$ as
- For each $u \in \mathcal{T}$,

$$[\alpha, \beta](u) = \sum_{j=1}^{|u|-1} \left(\alpha(u_{(j)})\beta(u^{(j)}) - \alpha(u^{(j)})\beta(u_{(j)})\right), \tag{7}$$

- $|u|$ denotes the number vertices of u
- each of the pairs of trees $(u_{(j)}, u^{(j)}) \in \mathcal{T} \times \mathcal{T}$, $j = 1, \ldots, |u| - 1$, is obtained from u by removing one of the $|u| - 1$ edges of the rooted tree u, the root of $u_{(j)}$ being the original root of u.

- For instance

$$
\begin{aligned}
[\alpha,\beta]\left(\genfrac{}{}{0pt}{}{\bullet}{\circ}\right) &= \alpha(\circ)\beta(\bullet)-\alpha(\bullet)\beta(\circ), \quad [\alpha,\beta]\left(\genfrac{}{}{0pt}{}{\circ}{\circ}\right)=0\\
[\alpha,\beta]\left(\overset{\bullet\;\;\bullet}{\circ}\right) &= 2\big(\alpha\left(\genfrac{}{}{0pt}{}{\bullet}{\circ}\right)\beta(\bullet)-\alpha(\bullet)\beta\left(\genfrac{}{}{0pt}{}{\bullet}{\circ}\right)\big)\\
[\alpha,\beta]\left(\overset{\bullet\;\;\circ}{\circ}\right) &= \alpha\left(\genfrac{}{}{0pt}{}{\circ}{\circ}\right)\beta(\bullet)+\alpha\left(\genfrac{}{}{0pt}{}{\bullet}{\circ}\right)\beta(\circ)\\
&\quad -\alpha(\bullet)\beta\left(\genfrac{}{}{0pt}{}{\circ}{\circ}\right)-\alpha(\circ)\beta\left(\genfrac{}{}{0pt}{}{\bullet}{\circ}\right)
\end{aligned}
$$

- The Lie subalgebra of $\mathfrak{g}$ generated by the maps $X, Y \in \mathfrak{g}$ defined as

$$
X(u)=\begin{cases} 1 & \text{if} \quad u=\bullet \\ 0 & \text{if} \quad u\in\mathcal{T}\backslash\{\bullet\}\end{cases}, \qquad Y(u)=\begin{cases} 1 & \text{if} \quad u=\circ \\ 0 & \text{if} \quad u\in\mathcal{T}\backslash\{\circ\}\end{cases}
$$

  is a free Lie algebra over the set $\{X, Y\}$
- $\mathcal{L}(X, Y)$: Lie subalgebra of $\mathfrak{g}$ generated by the maps X and Y.

- For each particular Hall–Viennot basis $\{E_i \ : \ i = 1, 2, 3, \ldots\}$, and X and Y as above, *one can associate a bicoloured rooted tree u_i to each element E_i* such that, for any map $\alpha \in \mathcal{L}(X, Y)$,

$$\alpha \quad = \quad \sum_{i \geq 1} \frac{\alpha(u_i)}{\sigma(u_i)} E_i, \tag{9}$$

- For each i, $\sigma(u_i)$ is certain positive integer associated to the bicoloured rooted tree u_i (the number of symmetries of u_i)
- if $\alpha \in \mathcal{L}(X, Y)$, then its projection α_n to the homogeneous subspace $\mathcal{L}(X, Y)_n$ is given by

$$\alpha_n(u) = \begin{cases} \alpha(u) & \text{if} \qquad |u| = n \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

for each $u \in \mathcal{T}$.

- Lie series:

$$\alpha = \alpha_1 + \alpha_2 + \alpha_3 + \cdots, \quad \text{where} \quad \alpha_n \in \mathcal{L}(X, Y)_n.$$

- A map $\alpha \in \mathfrak{g}$ is then a Lie series if and only if (9) holds
- The corresponding BCH series *is* a Lie series

$$\begin{aligned} Z &= \sum_{i \geq 1} z_i E_i = \sum_{i \geq 1} \frac{Z(u_i)}{\sigma(u_i)} E_i \\ &= Z(\bullet)X + Z(\circ)Y + Z(\overset{\bullet}{\circ})[Y, X] \\ &\quad + \frac{Z(\overset{\bullet\ \bullet}{\circ})}{2}[[Y, X], X] + Z(\overset{\bullet\ \circ}{\circ})[[Y, X], Y] + \cdots, \end{aligned}$$

- The coefficients $Z(u_i)$ can be determined by recursive procedures for BCH (Varadarajan)

In summary:

- Construct algorithmically a sequence of labeled rooted trees in a one-to-one correspondence with a Hall basis
- In addition, they must verify

$$\alpha = \sum_{i \geq 1} \frac{\alpha(u_i)}{\sigma(u_i)} E_i,$$

- In this way, one can build Lie series
- In particular, the BCH series
- A very efficient algorithm written in Mathematica allows us to get the BCH series up to a prescribed value of m in the Hall and Lyndon basis

Some results

- Comparison:
  - with the best previous algorithm: 17.5 MBytes up to $m = 10$.
  - ours: 5.4 MBytes
- In less than 15 min. of CPU (2008) and 1.5 GBytes we get up to $m = 20$
- 109697 non vanishing terms out of 111013 elements E_i of grade $|i| \leq 20$ in the Hall basis
- Last element:

$$\begin{aligned} E_{111013} &= [[[[[Y,X],Y],[Y,X]],[[[Y,X],X],[Y,X]]], \\ &\quad [[[[Y,X],Y],[Y,X]],[[[[Y,X],Y],Y],Y]]], \end{aligned}$$

with coefficient

$$z_{111013} = -\frac{19234697}{140792940288}.$$

- An observation: In the basis of P. Hall there are 1316 zero coefficients out of 111013 up to degree m $= 20$, whereas in the Lyndon basis the number of vanishing terms rises to 34 253 (more than 30% of the total number of coefficients!!)
- More remarkably, one notices that the distribution of these vanishing coefficients in the Lyndon basis follows a very specific pattern
- It is possible to explain this pattern
- In a sense, the Lyndon basis seems the natural choice to get systematically the BCH series with the minimum number of terms
- Variations: symmetric BCH formula

$$\mathrm{e}^{\frac{1}{2}X}\mathrm{e}^{Y}\mathrm{e}^{\frac{1}{2}X} = \mathrm{e}^{W}$$

Convergence

Theorem

(Mityagin) The Baker–Campbell–Hausdorff series converges absolutely when $\|X\| + \|Y\| < \pi$.

- This result can be generalized to any set $X_1, X_2, \ldots, X_k$ of bounded operators in a Hilbert space $\mathcal{H}$:

$$\mathrm{e}^{X_1}\mathrm{e}^{X_2}\cdots\mathrm{e}^{X_k} = \mathrm{e}^{Z}$$

  converges if

$$\|X_1\| + \|X_2\| + \cdots + \|X_k\| < \pi$$

- Optimal bound

An example

Let

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

and let $X = \alpha X_1$, $Y = \beta X_2$, with $\alpha, \beta \in \mathbb{C}$. Then

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

analytic function for $|\alpha| < \pi$ with first singularities at $\alpha = \pm i\pi$. Then BCH cannot converge if $|\alpha| \geq \pi$, independently of $\beta \neq 0$.

- From the above theorem: convergence if $|\alpha| + |\beta| < \pi$
- In the limit $|\beta| \to 0$ this result is optimal

Second example

$$X = \begin{pmatrix} 0 & 0 \\ \alpha & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}$$

with $\alpha > 0$. Then

$$\mathrm{e}^{\varepsilon X}\mathrm{e}^{\varepsilon Y} = \begin{pmatrix} 1 & \alpha\varepsilon \\ \alpha\varepsilon & 1+\alpha^2\varepsilon^2 \end{pmatrix} \tag{11}$$

- convergence of the BCH series in this case whenever $2\alpha|\varepsilon| < \pi$, or $|\varepsilon| < \frac{\pi}{2\alpha}$
- conservative estimate since convergence can be shown for $|\varepsilon| < \frac{2}{\alpha}$

Numerical check of convergence for $\alpha = 2$

- $Z^{[N]}(\varepsilon) = \sum_{n=1}^{N} Z_n(\varepsilon)$
- Compute $E_r(\varepsilon) = \|\mathrm{e}^X \mathrm{e}^Y \mathrm{e}^{-Z^{[N]}(\varepsilon)} - I\|$
- Convergence if $\varepsilon < 1$
- $\varepsilon = 1/4$; with $N = 10$, $E_r(\varepsilon) \approx 10^{-7}$. With $N = 15$, $E_r(\varepsilon) \approx 10^{-10}$
- $\varepsilon = 0.9$; to get $E_r(\varepsilon) \approx 10^{-8}$ we need $N = 150$; with $N = 200$ then $E_r(\varepsilon) \approx 10^{-10}$

Other results on convergence

- The Baker–Campbell–Hausdorff formula expressed as a series of homogeneous Lie polynomials in $X, Y \in \mathfrak{g}$ (a Banach Lie algebra), converges absolutely in the domain $D_1 \cup D_2$ of $\mathfrak{g} \times \mathfrak{g}$, where

$$\begin{aligned} D_1 &= \left\{ (X,Y) \, : \, \mu \|X\| < \int_{\mu\|Y\|}^{2\pi} \frac{1}{g(x)} dx \right\} \\ D_2 &= \left\{ (X,Y) \, : \, \mu \|Y\| < \int_{\mu\|X\|}^{2\pi} \frac{1}{g(x)} dx \right\} \end{aligned}$$

  and $g(x) = 2 + \frac{x}{2}(1 - \cot \frac{x}{2})$. (Michel 1974, F.C. & S. Blanes 2004)

- Biagi & Bonfiglioli 2014: generalization to arbitrary infinite-dimensional Banach-Lie algebras (in particular, without using the exponential map)

III. ZASSENHAUS FORMULA

- In the paper dealing with ME expansion, Magnus (1954) cites an unpublished reference by Zassenhaus, reporting that there exists a formula which may be called the dual of the (Baker–Campbell–)Hausdorff formula. More specifically,

Theorem

(Zassenhaus Formula). Let $\mathcal{L}(X, Y)$ be the free Lie algebra generated by X and Y. Then, e^{X+Y} can be uniquely decomposed as

$$\mathrm{e}^{X+Y} = \mathrm{e}^{X}\,\mathrm{e}^{Y}\prod_{n=2}^{\infty}\mathrm{e}^{C_n(X,Y)} = \mathrm{e}^{X}\,\mathrm{e}^{Y}\,\mathrm{e}^{C_2(X,Y)}\cdots\mathrm{e}^{C_n(X,Y)}\cdots,$$

where $C_n(X, Y) \in \mathcal{L}(X, Y)$ is a homogeneous Lie polynomial in X and Y of degree n.

Zassenhaus formula

- The existence of this formula is an immediate consequence of the BCH theorem.
- By comparing with the BCH formula it is possible to obtain the first terms as

$$C_2(X,Y) = -\frac{1}{2}[X,Y], \qquad C_3(X,Y) = \frac{1}{3}[Y,[X,Y]] + \frac{1}{6}[X,[X,Y]].$$

- Less familiar than the BCH formula but still important in several fields: statistical mechanics, many-body theories, quantum optics, path integrals, q-analysis in quantum groups, particle accelerators physics, etc.
- Numerical analysis: Iserles *et al.*, 2014.
- Again, two important aspects: *efficient computation* and *convergence* of the formula.

Some (brief) history

- Several systematic computations of the terms C_n for $n > 3$ have been carried out in the literature: Wilcox (1967), Volkin (1968), Suzuki (1976), Baues (1980). All of them give results for C_n as a linear combination of nested commutators.
- Scholz and Weyrauch (2006): recursive procedure based on upper triangular matrices.
- Weyrauch and Scholz (2009): C_n up to $n = 15$ in less than 2 minutes (with another procedure)
- Now
$$C_n = \sum_{w,|w|=n} g_w\, w, \tag{12}$$
where g_w is a rational coefficient and the sum is taken over all words w with length $|w| = n$ in the symbols X and Y, i.e., $w = a_1 a_2 \cdots a_n$, each a_i being X or Y.
- Applying Dynkin–Specht–Wever theorem it is possible to express them in terms of commutators, but in a way that there are redundancies

Our contribution

- To present a new recurrence that allows one to express the Zassenhaus terms C_n up to a prescribed degree directly in terms of independent commutators involving n operators X and Y.
- We are able to express directly C_n with the minimum number of commutators required at each degree n.
- We obtain sharper bounds for the terms of the Zassenhaus formula which show that the product converges in a larger domain than previous results.

A new recurrence

- We introduce a parameter λ,

$$\mathrm{e}^{\lambda(X+Y)} = \mathrm{e}^{\lambda X}\,\mathrm{e}^{\lambda Y}\,\mathrm{e}^{\lambda^2 C_2}\,\mathrm{e}^{\lambda^3 C_3}\,\mathrm{e}^{\lambda^4 C_4}\cdots \tag{13}$$

  so that the original Zassenhaus formula is recovered when $\lambda = 1$.

- Consider the compositions

$$R_1(\lambda) = \mathrm{e}^{-\lambda Y}\,\mathrm{e}^{-\lambda X}\,\mathrm{e}^{\lambda(X+Y)} \tag{14}$$

  and for each $n \geq 2$,

$$R_n(\lambda) = \mathrm{e}^{-\lambda^n C_n}\cdots\mathrm{e}^{-\lambda^2 C_2}\,\mathrm{e}^{-\lambda Y}\,\mathrm{e}^{-\lambda X}\,\mathrm{e}^{\lambda(X+Y)} = \mathrm{e}^{-\lambda^n C_n}\,R_{n-1}(\lambda).$$

  Then,

$$R_n(\lambda) = \mathrm{e}^{\lambda^{n+1} C_{n+1}}\,\mathrm{e}^{\lambda^{n+2} C_{n+2}}\cdots$$

- Finally,

$$F_n(\lambda) \equiv \left(\frac{d}{d\lambda} R_n(\lambda)\right) R_n(\lambda)^{-1}, \qquad n \geq 1. \qquad (15)$$

- We have for $n \geq 1$

$$\begin{aligned} F_{n+1}(\lambda) &= \mathrm{e}^{-\lambda^{n+1} \mathrm{ad}_{C_{n+1}}} G_{n+1}(\lambda), & (16) \\ C_{n+1} &= \frac{1}{(n+1)!} F_n^{(n)}(0), & (17) \\ G_{n+1}(\lambda) &= F_n(\lambda) - \frac{\lambda^n}{n!} F_n^{(n)}(0). & (18) \end{aligned}$$

Expressions (16)–(18) allow one to compute recursively the Zassenhaus terms C_n starting from $F_1(\lambda)$. The sequence is

$$F_n(\lambda) \longrightarrow C_{n+1} \longrightarrow G_{n+1}(\lambda) \longrightarrow F_{n+1}(\lambda) \longrightarrow \cdots$$

- For $n = 1$,

$$F_1(\lambda) = \mathrm{e}^{-\lambda \mathrm{ad}_Y}(\mathrm{e}^{-\lambda \mathrm{ad}_X} - I)Y,$$

that is,

$$F_1(\lambda) = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty} \frac{(-\lambda)^{i+j}}{i!j!} \mathrm{ad}_Y^i \mathrm{ad}_X^j Y \tag{19}$$

or equivalently

$$F_1(\lambda) = \sum_{k=1}^{\infty} f_{1,k}\, \lambda^k, \quad \text{with} \quad f_{1,k} = \sum_{j=1}^{k} \frac{(-1)^k}{j!(k-j)!} \mathrm{ad}_Y^{k-j} \mathrm{ad}_X^j Y. \tag{20}$$

- In general ($n \geq 2$),

$$F_n(\lambda) = \sum_{k=n}^{\infty} f_{n,k}\, \lambda^k, \quad \text{with} \quad f_{n,k} = \sum_{j=0}^{[k/n]-1} \frac{(-1)^j}{j!} \mathrm{ad}_{C_n}^j f_{n-1,k-nj}, \tag{21}$$

- It turns out that

$$F_n(\lambda) = \sum_{k=n+1}^{2n+2} k\, C_k\, \lambda^{k-1} + \lambda^{2n+2} H_n(\lambda)$$

where $H_n(\lambda)$ involves commutators of C_j, $j \geq n+1$

- Notice that the terms $C_{n+1}, \ldots, C_{2n+2}$ of the Zassenhaus formula can be then directly obtained from $F_n(\lambda)$.
- In particular,

$$C_{n+1} = \frac{1}{n+1} f_{1,n} = \frac{1}{n+1} \sum_{i=0}^{n-1} \frac{(-1)^n}{i!(n-j)!} \mathrm{ad}_Y^i \mathrm{ad}_X^{n-j} Y, \qquad (22)$$

for $n = 1, 2, 3$, and

$$C_{n+1} = \frac{1}{n+1}\, f_{[n/2],n} \qquad n \geq 5, \qquad (23)$$

Algorithm

$$
\begin{aligned}
&\text{Define } f_{1,k} = \sum_{j=1}^{k} \frac{(-1)^k}{j!(k-j)!} \mathrm{ad}_Y^{k-j} \mathrm{ad}_X^j Y \\
&C_2 = (1/2)\, f_{1,1} \\
&\text{Define } f_{n,k} \quad n \geq 2,\ k \geq n \ \text{ by :} \\
&f_{n,k} = \sum_{j=0}^{[k/n]-1} \frac{(-1)^j}{j!} \mathrm{ad}_{C_n}^j f_{n-1,k-nj} \\
&C_n = (1/n) f_{[(n-1)/2],n-1} \quad n \geq 3.
\end{aligned}
\tag{24}
$$

- Important property: it provides expressions for C_n that, by construction, involve only independent commutators. In other words, they cannot be simplified further by using the Jacobi identity and the antisymmetry property of the commutator.
- This can be easily proved by repeated application of the Lazard elimination principle.

Computational aspects

- The algorithm can be easily implemented in a symbolic algebra package. We need to define an object inheriting only the linearity property of the commutator, the adjoint operator and the functions $f_{n,k}$ and C_n.
- We have expressions of C_n up to $n = 20$ with a reasonable computational time and memory requirements (35 MB).

n	CPU time (seconds)		Memory (MB)	
	W-S	*New*	*W-S*	*New*
14	29.18	0.14	122.90	0.88
16	203.85	0.59	764.32	4.09
18		3.01		11.12
20		19.18		35.27

- C_{16} has 54146 terms when expressed as combinations of words, but only 3711 terms with the new formulation

Algorithm

```
Clear[Cmt, ad, ff, cc];
$RecursionLimit= 1024;
Cmt[a_, a_]:= 0;
Cmt[a___, 0, b___]:= 0;
Cmt[a___, c_ + d_, b___] := Cmt[a, c, b] + Cmt[a, d, b];
Cmt[a___, n_ c_Cmt, b___]:= n Cmt[a, c, b];
Cmt[a___, n_ X, b___]:= n Cmt[a, X, b];
Cmt[a___, n_ Y, b___]:= n Cmt[a, Y, b];
Cmt /: Format[Cmt[a_, b_]]:= SequenceForm["[", a, ",", b, "]"];

ad[a_, 0, b_]:= b;
ad[a_, j_Integer, b_]:= Cmt[a, ad[a, j-1, b]];
ff[1, k_]:= ff[1, k] =
   Sum[((-1)^k/(j! (k-j)!)) ad[Y, k-j, ad[X, j, Y]], {j, 1, k}];
cc[2] = (1/2) ff[1, 1];
ff[p_, k_]:= ff[p, k] =
   Sum[((-1)^j/j!) ad[cc[p], j, ff[p-1, k - p j]], {j, 0,
       IntegerPart[k/p] - 1}];
cc[p_Integer]:= cc[p] =
   Expand[(1/p) ff[IntegerPart[(p-1)/2], p-1]];
```

Convergence

- Suppose now that X and Y are defined in a Banach algebra $\mathcal{A}$
- Then it makes sense to analyze the convergence of the Zassenhaus formula.
- Only two previous results establishing sufficient conditions for convergence of the form $\|X\| + \|Y\| < r$ with a given $r > 0$.
- Suzuki (1976): $r_s = \log 2 - \frac{1}{2} \approx 0.1931$
- Bayen (1979): r_b given by the unique positive solution of the equation

$$z^2 \left(1 + 2\int_0^z \frac{\mathrm{e}^{2w} - 1}{w} dw\right) = 4(2\log 2 - 1).$$

  Numerically, $r_b = 0.59670569\ldots$.
- Thus, for $\|X\| + \|Y\| < r_b$ one has

$$\lim_{n\to\infty} \mathrm{e}^X \, \mathrm{e}^Y \, \mathrm{e}^{C_2} \cdots \mathrm{e}^{C_n} = \mathrm{e}^{X+Y}. \qquad (25)$$

Our treatment

- Next we use recursion (16)–(18) to show that it converges for $(x, y) \equiv (||X||, ||Y||) \in \mathbb{R}^2$ in a domain that is larger than $\{(x, y) \in \mathbb{R}^2 \ : \ 0 \leq x + y < r_b\}$.
- There is convergence if $\lim_{n\to\infty} ||R_n(1)|| = 1$.
- But $R_n(\lambda)$ is also the solution of

$$\frac{d}{d\lambda} R_n(\lambda) = F_n(\lambda) R_n(\lambda), \qquad R_n(0) = I. \tag{26}$$

- If $\int_0^1 \|F_n(\lambda)\| d\lambda < \infty$, then there exists a unique solution $R_n(\lambda)$ of (26) for $0 \leq \lambda \leq 1$, and $\|R_n(1)\| \leq \exp(\int_0^1 \|F_n(\lambda)\| d\lambda)$
- In consequence, convergence is guaranteed whenever $(x, y) = (\|X\|, \|Y\|) \in \mathbb{R}^2$ is such that

$$\lim_{n\to\infty} \int_0^1 \|F_n(\lambda)\| d\lambda = 0.$$

- We have that $\|C_{n+1}\| \leq \delta_{n+1}$, where $\delta_2 = x\,y$ and for $n \geq 2$,

$$\delta_{n+1} = \frac{1}{n+1} \sum_{(i_0,i_1,\ldots,i_n)\in\mathcal{I}_n} \frac{2^{i_0+\cdots+i_n}}{i_0!i_1!\cdots i_n!} \delta_n^{i_n} \cdots \delta_2^{i_2} y^{i_1} x^{i_0} y.$$

- Similarly, $\|F_n(\lambda)\| \leq f_n(\lambda)$ and

$$\int_0^1 f_n(\lambda) d\lambda \leq \sum_{k=n}^{\infty} \delta_k,$$

- Then, $\lim_{n\to\infty} ||R_n(1)|| = 1$ if the series $\sum_{k=2}^{\infty} \delta_k$ converges.
- Let's analyze each term in this series...

- We get from our recurrence

$$\begin{aligned} \|f_{1,k}\| &\leq d_{1,k} \equiv 2^k y \sum_{j=1}^{k} \frac{1}{j!(k-j)!} x^j y^{k-j} = \frac{2^k}{k!} y\big((x+y)^k - y^k\big) \\ \|f_{n,k}\| &\leq d_{n,k} = \sum_{j=0}^{[k/n]-1} \frac{2^j}{j!} \delta_n^j d_{n-1,k-nj} \end{aligned} \tag{27}$$

- Therefore

$$\|C_n\| \leq \delta_n = \frac{1}{n} d_{[(n-1)/2],n-1}, \qquad n \geq 3.$$

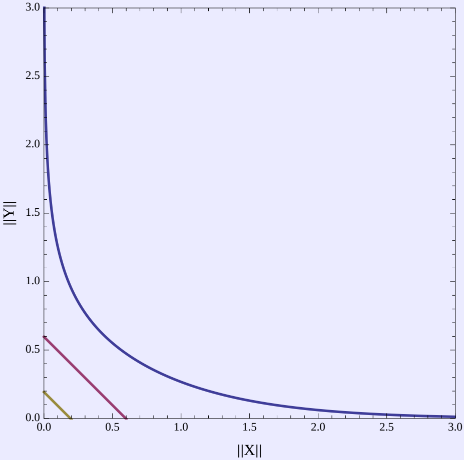
- A sufficient condition for convergence is obtained by imposing

$$\lim_{n\to\infty} \frac{\delta_{n+1}}{\delta_n} < 1. \tag{28}$$

- Recall that both $d_{n,k}$ and δ_n depend on $(x, y) = (\|X\|, \|Y\|)$, so condition (28) implies a constraint on the convergence domain $(x, y) \in \mathbb{R}^2$

Convergence domain

Computing numerically for each point the coefficients $d_{n,k}$ and δ_n up to $n = 1000$ we get



Example

$$X = \alpha \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad Y = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

- Compute $R_1 = \mathrm{e}^{-Y}\mathrm{e}^{-X}\mathrm{e}^{X+Y}$
- Compute $R_2(m) = \mathrm{e}^{C_2}\mathrm{e}^{C_3}\cdots\mathrm{e}^{C_m}$
- Finally $E_m = \|R_1 - R_2(m)\|$
- Particular case: $\alpha = 0.2$. Then we are outside the guaranteed domain of convergence
- $m = 10$, $E_{10} \approx 1.3345 \cdot 10^{-4}$
- $m = 15$, $E_{15} \approx 1.9180 \cdot 10^{-6}$
- $m = 20$, $E_{20} \approx 4.7958 \cdot 10^{-9}$

Generalization

- Sometimes one has to deal with

$$\exp(\lambda A_1 + \lambda^2 A_2 + \cdots + \lambda^n A_n + \cdots)$$

  with A_k non-commuting operators
- In that case it is still possible to generalize the expansion and to get

$$\mathrm{e}^{\lambda A_1 + \lambda^2 A_2 + \cdots} = \mathrm{e}^{\lambda C_1} \mathrm{e}^{\lambda^2 C_2} \ldots \mathrm{e}^{\lambda^n C_n} \ldots$$

- Recursive procedure to obtain C_k

IV. MAGNUS EXPANSION

General linear differential equation

Goal

Given the matrix $A(t)$ $N \times N$, solve the initial value problem

$$Y'(t) = A(t)Y(t), \qquad Y(t_0) = Y_0. \tag{29}$$

- If $N = 1$, the solution reads

$$Y(t) = \exp\left(\int_{t_0}^{t} A(s)ds\right) Y_0. \tag{30}$$

- This is also valid when $N > 1$ if $[A(t), \int_0^t A(s)ds] = 0$. Particular case: $A(t_1)A(t_2) = A(t_2)A(t_1)$ for all t_1 y t_2. In particular, when A is constant.
- In general, (30) is *not* the solution

- Typical procedure (Neumann, Dyson):

$$Y(t) = \int_{t_0}^{t} A(s)ds + \int_{t_0}^{t} A(s_1) \int_{t_0}^{s_1} A(s_2)ds_2 ds_1 + \cdots$$

- Magnus (1954): construct $Y(t)$ as a genuine *exponential representation*
- Motivation: problems arising in Quantum Mechanics (in this way, unitary is preserved, and is essential in QM)

Magnus expansion

- W. Magnus proposal: to express the solution as the exponential of a certain matrix function $\Omega(t, t_0)$,

$$Y(t) = \exp \Omega(t, t_0)\, Y_0 \tag{31}$$

- Ω is built as a series expansion

$$\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t). \tag{32}$$

- For simplicity, $\Omega(t) \equiv \Omega(t, t_0)$ y $t_0 = 0$.

- First terms:

$$
\begin{aligned}
\Omega_1(t) &= \int_0^t A(t_1)\, \mathrm{d}t_1, \\
\Omega_2(t) &= \frac{1}{2}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \ \ [A(t_1), A(t_2)] \\
\Omega_3(t) &= \frac{1}{6}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3 \ ([A(t_1),[A(t_2),A(t_3)]] + \\
&\quad [A(t_3),[A(t_2),A(t_1)]])
\end{aligned}
\tag{33}
$$

$[A, B] \equiv AB - BA$

- $\Omega_1(t)$ is exactly the exponent in the scalar case
- If we insist in keeping an exponential representation for $Y(t)$, then the exponent must be corrected
- The rest of the series (32) accounts for this correction

How the series is obtained?

- Insert $Y(t) = \exp \Omega(t)$ in $Y' = A(t)Y$, $Y(0) = I$
- Differential equation satisfied by Ω:

$$\frac{d\Omega}{dt} = \sum_{n=0}^{\infty} \frac{B_n}{n!} \operatorname{ad}_{\Omega}^{n} A, \qquad (34)$$

  where $\quad \operatorname{ad}_{\Omega}^{0} A = A, \qquad \operatorname{ad}_{\Omega}^{k+1} A = [\Omega, \operatorname{ad}_{\Omega}^{k} A],$
  and B_j are the Bernoulli number.
- At first sight, a very bad idea!: we replace a *linear* differential equation by another which is *highly nonlinear!*
- ... But this is defined for Ω

- We apply Picard's iteration:

$$\Omega^{[0]} = O, \qquad \Omega^{[1]} = \int_0^t A(t_1) dt_1,$$

$$\Omega^{[n]} = \int_0^t \left(A(t_1) dt_1 - \frac{1}{2}[\Omega^{[n-1]}, A] + \frac{1}{12}[\Omega^{[n-1]}, [\Omega^{[n-1]}, A]] + \cdots \right)$$

  so that $\lim_{n\to\infty} \Omega^{[n]}(t) = \Omega(t)$ in a neighborhood of $t = 0$
- Another recursive procedure to obtain the series, based on a generator

- When the recursion is worked out explicitly,

$$\Omega_n(t) = \sum_{j=1}^{n-1} \frac{B_j}{j!} \sum_{\substack{k_1+\cdots+k_j=n-1 \\ k_1\ge 1,\ldots,k_j\ge 1}} \int_0^t \mathrm{ad}_{\Omega_{k_1}(s)}\, \mathrm{ad}_{\Omega_{k_2}(s)} \cdots \mathrm{ad}_{\Omega_{k_j}(s)} A(s)\, ds$$

- Ω_n is a linear combination of n-multiple integrals of $n-1$-nested commutators containing n operators A evaluated at different times
- The expression is increasingly complicated when n grows

Some properties

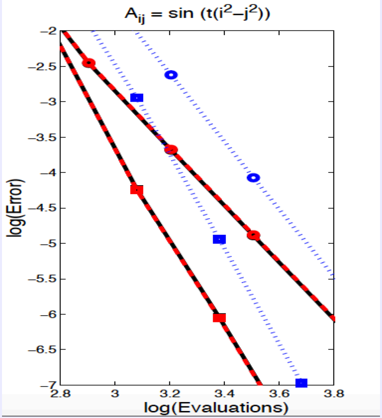
- If $A(t)$ belongs to some Lie algebra $\mathfrak{g}$, then $\Omega(t)$ (and truncation of the Magnus series) also belongs to $\mathfrak{g}$ and therefore $\exp(\Omega) \in \mathcal{G}$, where $\mathcal{G}$ is the Lie group with Lie algebra $\mathfrak{g}$.
  1. Symplectic group (in Hamiltonian mechanics)
  2. Unitary group (for the Schrödinger equation)
- The resulting approximations preserve important qualitative properties of the exact solution (e.g., unitarity, etc.)

- Analytic approximations
- Starting point for the construction of new families of numerical integrators for $Y' = A(t)Y$
- *Very efficient* high order numerical methods
- *Lie group integrators*, special class of *geometric numerical integration methods*

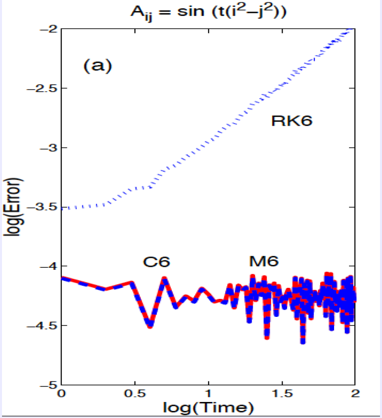
Example: $Y' = A(t)Y, \quad Y(0) = I$

$$A_{ij} = \sin(t(j^2 - i^2)), \quad 1 \leq i < j \leq 10$$

Efficiency diagram



Error as a function of time



Convergence

- Is this result only formal? What about convergence?
- Specifically, given a certain operator $A(t)$, when it is possible to get $\Omega(t)$ in (31) as the sum of the series $\Omega(t) = \sum_{n=1}^{\infty} \Omega_n(t)$?
- It turns out that the Magnus series converges for $t \in [0, T)$ such that
$$\int_0^T \|A(s)\| ds < \pi$$
where $\|\cdot\|$ is the 2-norm
- This is a generic result, in the sense that it is possible to find particular matrices $A(t)$ so that the series diverges for all $t > T$.
- ... But is is only a *sufficient* condition: there exist matrices $A(t)$ so that the expansion converges for $t > T$.
- Analysis of the eigenvalues

Remarks

- The result is valid for *complex* matrices $A(t)$
- In fact, for any given bounded operator $A(t)$ in a Hilbert space $\mathcal{H}$ if Y is a normal operator (in particular, if $\mathrm{i}Y$ is unitary).
- This results can be used in turn to prove the convergence of the Baker–Campbell–Hausdorff formula

BCH and the Magnus expansion

- Consider the initial value problem

$$U' = A(t)U, \qquad U(0) = I, \tag{35}$$

with

$$A(t) = \begin{cases} Y & 0 \le t \le 1 \\ X & 1 < t \le 2 \end{cases}$$

The exact solution of (35) at $t = 2$ is $U(2) = \mathrm{e}^X \, \mathrm{e}^Y$.

- But we can apply Magnus: $U(2) = \mathrm{e}^{\Omega(2)}$.
- In this way it is possible to get BCH as a particular case of the Magnus expansion. (Sometimes it is called the *continuous BCH formula BCH*).

Work in progress

- 'Symmetric' Zassenhaus formula: useful for obtaining new numerical methods for certain classes of PDEs (Bader *et al.* 2014)

  $$\mathrm{e}^{X+Y} = \cdots \mathrm{e}^{C_3}\,\mathrm{e}^{C_2}\,\mathrm{e}^{\frac{1}{2}Y}\,\mathrm{e}^{X}\,\mathrm{e}^{\frac{1}{2}Y}\,\mathrm{e}^{C_2}\,\mathrm{e}^{C_3}\cdots$$

  in an efficient way
- 'Continuous analogue' of the Zassenhaus formula (M. Nadinic's Thesis 2015): Given $U' = A(t)Y$, $U(0) = I$,

  $$U(t) = \mathrm{e}^{W_1(t)}\,\mathrm{e}^{W_2(t)}\,\cdots\mathrm{e}^{W_r(t)}\cdots$$

  as efficiently as possible

Some references

- S. Blanes, F.C., J.A. Oteo, J. Ros, The Magnus expansion and some of its applications, *Physics Reports* **470** (2009), 151-238.
- S. Blanes and F.C., On the convergence and optimization of the Baker–Campbell–Hausdorff formula, *Lin. Alg. Appl.* **378** (2004), 135-158.
- A. Bonfiglioli and R. Fulci, *Topics in Noncommutative Algebra*, Lecture Notes in Mathematics 2034. Springer, 2012.
- F.C. and A. Murua, An efficient algorithm for computing the Baker–Campbell–Hausdorff series and some of its applications. *J. Math. Phys.* **50** (2009), 033513.
- F.C., A. Murua and M. Nadinic, Efficient computation of the Zassenhaus formula, *Comput. Phys. Comm.* **183** (2012), 2386-2391.
- W. Magnus, On the exponential solution of differential equations for a linear operator. *Commun. Pure Appl. Math.* **VII** (1954), 649-673.