



Article

A Unifying Framework for Perturbative Exponential Factorizations

Ana Arnal ^{1,*}, Fernando Casas ¹, Cristina Chiralt ¹ and José Angel Oteo ²

¹ Departament de Matemàtiques and IMAC, Universitat Jaume I, 12071 Castellón, Spain; Fernando.Casas@uji.es (F.C.); chiralt@uji.es (C.C.)

² Departament de Física Teòrica, Universitat de València and Institute for Integrative Systems Biology (I2SysBio), 46100 Burjassot, Spain; oteo@uv.es

* Correspondence: ana.arnal@uji.es

Abstract: We propose a framework where Fer and Wilcox expansions for the solution of differential equations are derived from two particular choices for the initial transformation that seeds the product expansion. In this scheme, intermediate expansions can also be envisaged. Recurrence formulas are developed. A new lower bound for the convergence of the Wilcox expansion is provided, as well as some applications of the results. In particular, two examples are worked out up to a high order of approximation to illustrate the behavior of the Wilcox expansion.

Keywords: sequences of linear transformations; Wilcox expansion; Fer expansion; Zassenhaus formula; Bellman problem



Citation: Arnal, A.; Casas, F.; Chiralt, C.; Oteo, J.A. A Unifying Framework for Perturbative Exponential Factorizations.

Mathematics **2021**, *9*, 637.
<https://doi.org/10.3390/math9060637>

Academic Editor: Higinio Ramos

Received: 25 January 2021
Accepted: 13 March 2021
Published: 17 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Linear differential equations of the form

$$\dot{x} \equiv \frac{dx}{dt} = A(t)x, \quad x(0) = x_0 \in \mathbb{C}^d, \quad (1)$$

with $A(t)$ a $d \times d$ matrix, whose entries are integrable functions of t , are ever-present in many branches of science, the fundamental evolution equation of Quantum Mechanics, the Schrödinger equation, being a particular case. In consequence, solving Equation (1) is of the greatest importance. In spite of their apparent simplicity, however, they are seldom solvable in terms of elementary functions, and so different procedures have been proposed over the years to render approximate solutions. These are specially useful in the analytical treatment of perturbative problems, such as those arising in the time evolution of quantum systems [1], control theory, or problems where time-ordered products are involved [2]. Among them, exponential perturbative expansions have received a great deal of attention, due to some remarkable properties they possess. In particular, if Equation (1) is defined in a Lie group, the approximations they furnish also evolve in the same Lie group. As a consequence, important qualitative properties of the exact solution are also preserved by the approximations. Thus, if Equation (1) represents the time-dependent Schrödinger equation, then the approximate evolution operator is still unitary, and as a consequence, the total sum of (approximate) transition probabilities is the unity, no matter where the expansion is truncated. There are, in fact, many physical problems (in non-linear mechanics, optical spectroscopy, magnetic resonance, etc.) involving periodic fast-oscillating external fields that are also modeled by Equation (1), with $A(t)$ periodic. In that case, especially tailored expansions incorporating the well-known Floquet theorem [3], such as the average Hamiltonian theory [4] and the Floquet–Magnus expansion [5,6], have also been proposed.

When dealing with the general problem (1), one of the most widely used exponential approximations corresponds to the Magnus expansion [7]

$$x(t) = e^{\Omega(t)} x_0, \quad (2)$$

where Ω is an infinite series

$$\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t) \tag{3}$$

whose terms are linear combinations of time-ordered integrals of nested commutators of A evaluated at different times (see [8] for a review, including applications to several physical and mathematical problems). What is more interesting for our purposes here is that this expansion can be related with a coordinate transformation $x \mapsto X$ rendering the original system (1) into the trivial equation

$$\frac{dX}{dt} = 0, \tag{4}$$

with the *static* solution $X(t) = X(0) = x_0$, and that the transformation is given precisely by $x(t) = \exp(\Omega(t))X(t)$ [9].

In contrast to the Magnus expansion, the Floquet–Magnus expansion obtains the solution with two exponential transformations when $A(t)$ is periodic, whereas other exponential perturbative expansions are based on infinite product factorizations of $x(t)$,

$$x(t) = e^{\Omega_1(t)} e^{\Omega_2(t)} \dots e^{\Omega_n(t)} \dots x(0), \tag{5}$$

such as those proposed by Fer and Wilcox. In fact, as pointed out in [10], both expansions have a curious history, which is worth describing. It was Fer who proposed the expansion that bears his name in [11], although he never applied it to solve any specific problem. Bellman reviewed this paper in the *Mathematical Reviews* (MR0104009), and even proposed the expansion as an exercise in [12]. Nevertheless, Wilcox identified it in [1] with an alternative factorization, Equation (5), which was indeed a different and new type of expansion. From then on, their historical trajectories move apart. Thus, Fer expansion was rediscovered by Iserles [13] as a tool for the numerical integration of linear differential equations and later on used in Quantum Mechanics [14] and solid-state nuclear magnetic resonance [15], but also as a Lie-group integrator [16–18]. On the other hand, Wilcox expansion has been rediscovered several times in the literature, in particular in [19] in the context of nonlinear control systems, and in [20] as a general tool for approximating the time evolution operator in Quantum Mechanics.

The first goal in this work is to recast both infinite product expansions within a unifying framework. This is done by considering, instead of just *one* exponential transformation, as in the case of the Magnus expansion, a *sequence* of such transformations, $e^{\Omega_1(t)}, \dots, e^{\Omega_k(t)}, \dots$, chosen to satisfy certain requirements. To be more specific, suppose one replaces $A(t)$ in Equation (1) by $\lambda A(t)$, where $\lambda > 0$ is a parameter. Then, if the transformations are chosen so that each $\Omega_k(t)$ is proportional to λ^k , we recover the Wilcox expansion, whereas we end up with the Fer expansion when each $\Omega_k(t)$ is an infinite series in λ whose first term is proportional to $\lambda^{2^{k-1}}$.

We also show that further alternative descriptions yield new factorizations. This additional degree of freedom can be indeed used to deal better with the features of the matrix $A(t)$, as in Floquet–Magnus, when $A(t)$ is periodic.

One might then consider this sequence of linear transformations as a generalization of the concept of picture in Quantum Mechanics when Equation (1) refers to the Schrödinger equation.

Our second goal consists of obtaining, on the basis of this framework, new results concerning Wilcox expansion. Thus, we develop a recursive procedure to obtain every order of approximation in terms of nested commutators, as well as a convergence radius bound. We also establish a formal connection of the Wilcox expansion with the Zassenhaus formula [7,21].

Eventually, the important problem of expanding the exponential $\exp(A + \varepsilon B)$ for $\varepsilon > 0$ and A, B , two generic non-commuting operators will be addressed, and two applications of the obtained results.

2. A Sequence of Transformations: The General Case

Given the initial value problem (1), let us consider a linear change in variables $x \mapsto X_1$ of the form

$$x(t) = e^{\Omega_1(t)} X_1(t), \quad \Omega_1(0) = 0 \tag{6}$$

transforming the original system into

$$\frac{dX_1}{dt} = B_1(t) X_1. \tag{7}$$

For the time being, the generator $\Omega_1(t)$ of the transformation is not specified. Then, $B_1(t)$ can be expressed in terms of $A(t)$ and $\Omega_1(t)$ as follows. First, by inserting Equation (6) into Equation (1), and taking Equation (7) into account, one can obtain

$$\frac{d}{dt} \exp(\Omega_1) = A(t) \exp(\Omega_1) - \exp(\Omega_1) B_1(t), \tag{8}$$

whence

$$B_1(t) = e^{-\Omega_1} A(t) e^{\Omega_1} - e^{-\Omega_1} \frac{d}{dt} e^{\Omega_1}. \tag{9}$$

The derivative of the matrix exponential can be written as [8]

$$\frac{d}{dt} \exp(\Omega_1(t)) = d \exp_{\Omega_1(t)}(\dot{\Omega}_1(t)) \exp(\Omega_1), \tag{10}$$

where the symbol $d \exp_{\Omega}(C)$ stands for the (everywhere convergent) power series

$$d \exp_{\Omega}(C) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \text{ad}_{\Omega}^k(C) \equiv \frac{\exp(\text{ad}_{\Omega}) - I}{\text{ad}_{\Omega}}(C). \tag{11}$$

Here $\text{ad}_{\Omega}^0 C = C$, $\text{ad}_{\Omega}^k C = [\Omega, \text{ad}_{\Omega}^{k-1} C]$, and $[\Omega, C]$ denotes the usual commutator. Therefore

$$B_1(t) = e^{-\Omega_1} \left(A(t) - d \exp_{\Omega_1(t)}(\dot{\Omega}_1(t)) \right) e^{\Omega_1} \equiv e^{-\text{ad}_{\Omega_1}}(B_0 - G_1) \tag{12}$$

where

$$B_0(t) \equiv A(t), \quad G_1(t) \equiv d \exp_{\Omega_1(t)}(\dot{\Omega}_1(t)) \tag{13}$$

and

$$e^{-\text{ad}_{\Omega_1}} F = \sum_{k \geq 0} \frac{(-1)^k}{k!} \text{ad}_{\Omega_1}^k F = e^{-\Omega_1} F e^{\Omega_1}. \tag{14}$$

Of course, nothing prevents us from repeating the whole procedure above and introducing a second transformation to Equation (7) of the form

$$X_1(t) = e^{\Omega_2(t)} X_2(t), \quad \Omega_2(0) = 0, \tag{15}$$

so that the new variables verify

$$\frac{dX_2}{dt} = B_2(t) X_2. \tag{16}$$

In general, for the n -th such linear transformation

$$X_{n-1}(t) = e^{\Omega_n(t)} X_n(t), \quad \Omega_n(0) = 0 \tag{17}$$

with

$$\frac{dX_n}{dt} = B_n(t) X_n, \tag{18}$$

one has

$$B_n(t) = e^{-\text{ad}_{\Omega_n}}(B_{n-1} - G_n), \quad \text{with} \quad G_n(t) = d \exp_{\Omega_n(t)}(\dot{\Omega}_n(t)) \tag{19}$$

so that the solution of Equation (1) is expressed as

$$x(t) = e^{\Omega_1(t)} e^{\Omega_2(t)} \dots e^{\Omega_n(t)} X_n(t). \tag{20}$$

Alternatively, we can write $B_n(t)$ in Equation (19) as follows. Since it is also true that [8]

$$\begin{aligned} d \exp_{\Omega_n(t)}(\dot{\Omega}_n(t)) &= \left(\frac{d}{dt} e^{\Omega_n} \right) e^{-\Omega_n} = \left(\int_0^1 e^{x\Omega_n} \dot{\Omega}_n e^{(1-x)\Omega_n} dx \right) e^{-\Omega_n} \\ &= \int_0^1 e^{x\Omega_n} \dot{\Omega}_n e^{-x\Omega_n} dx, \end{aligned} \tag{21}$$

we have

$$\begin{aligned} B_n &= e^{-\Omega_n} B_{n-1} e^{\Omega_n} - \int_0^1 e^{-u\Omega_n} \dot{\Omega}_n e^{u\Omega_n} du \\ &= \sum_{k \geq 0} \frac{(-1)^k}{(k+1)!} \left((k+1) \text{ad}_{\Omega_n}^k B_{n-1} - \text{ad}_{\Omega_n}^k \dot{\Omega}_n \right), \quad n \geq 1. \end{aligned} \tag{22}$$

The important point is, of course, how to choose B_n , or, alternatively, Ω_n , i.e., the specific requirements each transformation has to satisfy in order to be useful to approximately solve Equation (1). There are obviously many possibilities, and in the following we analyze two of them, leading to two different and well-known exponential perturbation factorizations mentioned in the Introduction, namely the Wilcox [1] and Fer [11] expansions.

3. Wilcox Expansion

3.1. Recurrences

Let us introduce the (dummy) parameter λ in Equation (1) and replace A with λA . This is helpful when collecting coefficients, and at the end we can always take $\lambda = 1$.

Since the solution of Equation (1) when A is constant, or more generally when $A(t_1)A(t_2) = A(t_2)A(t_1)$ for all $t_1 \neq t_2$, is $x(t) = \exp(\int_0^t A(u)du)$, it makes sense to take the generator for the first transformation as

$$\Omega_1(t) = \int_0^t B_0(u) du = \lambda \int_0^t A(u) du \equiv \lambda W_1(t). \tag{23}$$

Then, according to Equations (12) and (13), we have

$$C_1 \equiv B_0 - G_1 = \sum_{k \geq 1} \lambda^k (b_{0,k} - g_{1,k}) \tag{24}$$

where

$$\begin{aligned} b_{0,1} &= A(t), & b_{0,l} &= 0, \quad l > 1 \\ g_{1,1} &= \dot{W}_1, & g_{1,l} &= \frac{1}{l!} \text{ad}_{W_1}^{l-1} \dot{W}_1, \quad l > 1. \end{aligned} \tag{25}$$

With the choice $\dot{W}_1 = A(t)$, it turns out that B_1 is a power series in λ starting with λ^2 ,

$$B_1(t) = e^{-\text{ad}_{\Omega_1}}(B_0 - G_1) = e^{-\lambda \text{ad}_{W_1}} C_1 = \sum_{l \geq 2} \lambda^l b_{1,l}, \tag{26}$$

where

$$b_{1,l} = \sum_{k=0}^{l-2} \frac{(-1)^k}{k!} \text{ad}_{W_1}^k (b_{0,l-k} - g_{1,l-k}). \tag{27}$$

We can analogously choose the second transformation proportional to λ^2 , i.e., as $\Omega_2(t) \equiv \lambda^2 W_2(t)$ for a given W_2 to be determined. Then, a straightforward calculation shows that

$$C_2 \equiv B_1 - G_2 = \sum_{l \geq 2} \lambda^l (b_{1,l} - g_{2,l}) \tag{28}$$

with

$$g_{2,2} = \dot{W}_2, \quad g_{2,2l} = \frac{1}{l!} \text{ad}_{W_2}^{l-1} \dot{W}_2, \quad g_{2,r} = 0, \quad r \neq 2l. \tag{29}$$

The generator W_2 is then obtained by imposing that $b_{1,2} - g_{2,2} = 0$ in Equation (28), i.e.,

$$g_{2,2} = \dot{W}_2 = b_{1,2} = b_{0,2} - g_{1,2} = -\frac{1}{2} \text{ad}_{W_1} \dot{W}_1. \tag{30}$$

In this way, $B_2(t)$ is a power series in λ , starting with λ^3 ,

$$B_2(t) = e^{-\lambda^2 \text{ad}_{W_2}} C_2 = \sum_{l \geq 3} \lambda^l b_{2,l} \tag{31}$$

with

$$b_{2,l} = \sum_{k=0}^{[(l-1)/2]-1} \frac{(-1)^k}{k!} \text{ad}_{W_2}^k (b_{1,l-2k} - g_{2,l-2k}), \tag{32}$$

where $[\cdot]$ stands for the integer part of the argument. In general, the n -th transformation $\Omega_n(t) \equiv \lambda^n W_n(t)$ is determined in such a way that the power series of $B_n(t)$ starts with λ^{n+1} . This can be done as follows: from $B_{n-1} = \sum_{l \geq n} \lambda^l b_{n-1,l}$, we compute

$$C_n \equiv B_{n-1} - G_n = \sum_{r=n}^{\infty} \lambda^r c_{n,r} = \sum_{r=n}^{\infty} \lambda^r (b_{n-1,r} - g_{n,r}) \tag{33}$$

with

$$g_{n,r} = \begin{cases} \frac{1}{l!} \text{ad}_{W_n}^{l-1} \dot{W}_n, & r = nl \\ 0, & r \neq nl \end{cases} \quad l = 1, 2, \dots \tag{34}$$

Then, W_n is obtained by taking $b_{n-1,n} - g_{n,n} = 0$, i.e.,

$$\dot{W}_n = b_{n-1,n} \tag{35}$$

and, finally, B_n is determined as

$$B_n = e^{-\lambda^n \text{ad}_{W_n}} C_n = \sum_{l=n+1}^{\infty} \lambda^l b_{n,l} \tag{36}$$

with

$$b_{n,l} = \sum_{k=0}^{[\frac{l-1}{n}]-1} \frac{(-1)^k}{k!} \text{ad}_{W_n}^k c_{n,l-nk}. \tag{37}$$

Notice that, in view of Equations (34) and (37), Equation (35) simplifies to

$$\dot{W}_n = b_{n-2,n} \quad \text{for} \quad n \geq 3. \tag{38}$$

The solution of Equation (1) is expressed, after n such transformations, as

$$x(t) = e^{\lambda W_1(t)} e^{\lambda^2 W_2(t)} \dots e^{\lambda^n W_n(t)} X_n(t). \tag{39}$$

An approximation to the exact solution containing all the dependence up to $\mathcal{O}(\lambda^n)$ is obtained by taking $X_n = x_0$ in Equation (39).

In spite of this, the truncated factorization $e^{\lambda W_1(t)} e^{\lambda^2 W_2(t)} \dots e^{\lambda^n W_n(t)}$ still shares relevant qualitative properties with the evolution operator, such as orthogonality, unitarity, etc.

Recursion (33)–(38) allows one to construct *any* generator W_k of the expansion in terms of W_1, \dots, W_{k-1} . For illustration, we next collect the first terms

$$\begin{aligned} \dot{W}_1 &= A(t), & \dot{W}_2 &= -\frac{1}{2} \text{ad}_{W_1} \dot{W}_1, & \dot{W}_3 &= \frac{1}{3} \text{ad}_{W_1}^2 \dot{W}_1 \\ \dot{W}_4 &= -\frac{1}{8} \text{ad}_{W_1}^3 \dot{W}_1 - \frac{1}{2} \text{ad}_{W_2} \dot{W}_2, & \dot{W}_5 &= \frac{1}{30} \text{ad}_{W_1}^4 \dot{W}_1 - \text{ad}_{W_2} \dot{W}_3 \end{aligned} \tag{40}$$

This is the way the Wilcox expansion is built up.

3.2. Explicit Expressions for $W_n(t)$

Although the recursive procedure (33)–(38) turns out to be very computationally efficient to construct the exponents $W_n(t)$ for a given $A(t)$ in practice, it is clear that much insight about the expansion can be gained if an explicit expression for *any* W_n can be constructed, thus generalizing the treatment originally done by Wilcox up to $n = 4$ [1].

Such an expression could be obtained, in principle, by working out the recurrence (33)–(38), but a more direct approach consists of comparing the Dyson perturbation series of $U(t)$ [22] in the associated initial value problem

$$\dot{U} = \lambda A(t)U, \quad U(0) = I, \tag{41}$$

i.e.,

$$U(t) = I + \sum_{n=1}^{\infty} \lambda^n P_n(t), \quad \text{with} \quad P_k(t) = \int_0^t dt_1 \dots \int_0^{t_{k-1}} dt_k A(t_1) \dots A(t_k) \tag{42}$$

with the expansion in λ of the factorization

$$U(t) = e^{\lambda W_1(t)} e^{\lambda^2 W_2(t)} \dots e^{\lambda^n W_n(t)} \dots \tag{43}$$

Thus, for the first terms, one has

$$\begin{aligned} P_1 &= W_1, & P_2 &= W_2 + \frac{1}{2} W_1^2, & P_3 &= W_3 + W_1 W_2 + \frac{1}{3!} W_1^3, \\ P_4 &= W_4 + W_1 W_3 + \frac{1}{2} W_1^2 W_2 + \frac{1}{2} W_2^2 + \frac{1}{4!} W_1^4. \end{aligned} \tag{44}$$

In general, we can write

$$P_n = \sum_{\substack{p(n) \\ i_1 \leq i_2 \leq \dots \leq i_k}} \frac{1}{r_k!} W_{i_1} W_{i_2} \dots W_{i_k}, \quad i_1 \leq i_2 \leq \dots \leq i_k, \tag{45}$$

where the sum is extended over the total number of partitions $p(n)$ of the integer n . We recall that a partition $p(n)$ of the integer n is an n -tuple (i_1, i_2, \dots, i_k) , such that $i_1 + i_2 + \dots + i_k = n$, with ordering $i_1 \leq i_2 \leq \dots \leq i_k$. Thus, the seven partitions of $n = 5$ with the chosen ordering are (5), (1, 4), (2, 3), (1, 1, 3), (1, 2, 2), (1, 1, 1, 2) and (1, 1, 1, 1, 1). In Equation (45), r_k is the number of repeated indices in the partition considered.

By working out Equation (45), one can invert the relations and express W_n in terms of P_1, \dots, P_{n-1} for any $n \geq 1$. Thus, one obtains

$$\begin{aligned} W_1 &= P_1, & W_2 &= P_2 - \frac{1}{2} P_1^2, & W_3 &= P_3 - P_1 P_2 + \frac{1}{3} P_1^3, \\ W_4 &= P_4 - P_1 P_3 + \frac{3}{4} P_1^2 P_2 + \frac{1}{4} P_2 P_1^2 - \frac{1}{2} P_2^2 - \frac{1}{4} P_1^4. \end{aligned} \tag{46}$$

Notice that $W_n, n \geq 2$, is expressed in terms of products of iterated integrals $P_{i_1} \cdots P_{i_j}$. Interestingly, it is possible to express these products as proper time-ordered integrals by using a procedure developed in [23]. If we denote

$$A(i_1 i_2 \dots i_n) \equiv \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n A(t_{i_1}) A(t_{i_2}) \cdots A(t_{i_n}), \tag{47}$$

so that $P_n(t) = A(12 \dots n)$, then

$$\begin{aligned} W_2 &= A(12) - \frac{1}{2} A(1) \cdot A(1) \\ W_3 &= A(123) - A(1) \cdot A(12) + \frac{1}{3} A(1) \cdot A(1) \cdot A(1), \end{aligned} \tag{48}$$

etc. Taking into account Fubini’s theorem,

$$\int_0^\alpha dy \int_y^\alpha f(x, y) dx = \int_0^\alpha dx \int_0^x f(x, y) dy, \tag{49}$$

it is clear that $A(1) \cdot A(1) = A(12) + A(21)$, and thus

$$W_2 = A(12) - \frac{1}{2} (A(12) + A(21)) = \frac{1}{2} (A(12) - A(21)). \tag{50}$$

We can proceed analogously with the following products

$$\begin{aligned} A(1) \cdot A(12) &= A(123) + A(213) + A(312) \\ A(1) \cdot A(1) \cdot A(1) &= A(123) + A(132) + A(213) + A(231) + A(312) + A(321), \end{aligned} \tag{51}$$

so that

$$W_3 = \frac{1}{3} A(123) + \frac{1}{3} A(132) - \frac{2}{3} A(213) + \frac{1}{3} A(231) - \frac{2}{3} A(312) + \frac{1}{3} A(321). \tag{52}$$

Carrying out this argument to any order, we can expand all the products of integrals appearing in W_n . As a result, each product is replaced by the sum of all possible permutations of time ordering consistent with the time ordering in the factors of this product [24].

At this point, it is illustrative to consider some examples in detail. Thus, the product $A(1) \cdot A(12)$ gives the sum of all permutations of three elements, such that the second index is less than the third one. With respect to $A(1) \cdot A(1) \cdot A(1)$, since there is no special ordering, then all possible permutations have to be taken into account. Finally, for the product $P_1 P_3$ appearing in W_4 one has

$$P_1 P_3 = A(1) \cdot A(123) = A(4123) + A(3124) + A(2134) + A(1234). \tag{53}$$

Proceeding in a similar way, one can show that any product of iterated integrals can be expressed as a sum of iterated integrals. This property is, in fact, related to a much deeper characterization of the group of permutations [23]. If $\mathfrak{S}Symm$ denotes the graded \mathbb{Q} -vector space with the fundamental basis given by the disjoint union of the symmetric groups S_n for all $n \geq 0$, then it is possible to define a product $*$ of permutations and a coproduct δ in $\mathfrak{S}Symm$, so that there is a one-to-one correspondence between iterated integrals and permutations

$$A(\sigma) \cdot A(\tau) = A(\sigma * \tau). \tag{54}$$

The product $*$ was introduced in [25], and, together with the coproduct δ , endows $\mathfrak{S}Symm$ with a structure of Hopf algebra [26], the so-called Malvenuto–Reutenauer Hopf algebra of permutations [27].

In sum, the *general* structure of the Wilcox expansion terms therefore reads as

$$W_n = \sum_{k=1}^{n!} \omega_k^{(n)} A(\sigma_k), \tag{55}$$

where the summation extends over *all* the $n!$ permutations $\sigma_k \in S_n$ of $\{1, 2, \dots, n\}$. The weights $\omega_k^{(n)}$ are given by rational numbers that can be determined algorithmically for any n , although the general expression for them is not obvious, in contrast with the Magnus expansion, for which such a closed formula exists [28]. This can be then considered as an open problem.

Moreover, if one is interested in obtaining a compact expression for W_n in terms of independent nested commutators of $A(t)$, as is done in [1] up to $n = 4$; one can use the class of bases proposed by Dragt & Forest in [24] for the Lie algebra generated by the operators $A(t_1), A(t_2), \dots, A(t_n)$. The same procedure as carried out in [23] for the Magnus expansion can be applied here, so that one obtains the general formula

$$W_n(t) = \sum_{\tau_k} c_{\tau_k}^{(n)} \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n [A(t_{\tau(2)}), [A(t_{\tau(3)}) \cdots [A(t_{\tau(n)}), A(t_1)] \cdots]]. \tag{56}$$

Here, the sum extends over the $(n - 1)!$ permutations τ_k of the elements $\{2, 3, \dots, n\}$ and $c_{\tau_k}^{(n)}$ is a rational number that depends on the particular permutation. For a given permutation, say τ_k , its value coincides with the prefactor in Equation (55) of the particular term $A(\sigma_k)$, corresponding to the permutation $\sigma_k \in S_n$, such that

$$\{\sigma_k(1), \sigma_k(2), \dots, \sigma_k(n - 1), \sigma_k(n)\} = \{\tau_k(2), \tau_k(3), \dots, \tau_k(n), 1\}. \tag{57}$$

Thus, if we denote

$$A[i_1 i_2 \dots i_n] \equiv \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n [A(t_{i_1}), [A(t_{i_2}), \cdots [A(t_{i_{n-1}}), A(t_{i_n})] \cdots]]. \tag{58}$$

we obtain for the first terms

$$\begin{aligned} W_2 &= -\frac{1}{2}A[21], \\ W_3 &= \frac{1}{3}A[231] + \frac{1}{3}A[321], \\ W_4 &= -\frac{1}{4}A[3241] - \frac{1}{4}A[4231] - \frac{1}{4}A[4321], \\ W_5 &= -\frac{2}{15}A[23451] - \frac{2}{15}A[23541] - \frac{2}{15}A[24351] - \frac{2}{15}A[24531] \\ &\quad - \frac{2}{15}A[25341] - \frac{2}{15}A[25431] + \frac{1}{5}A[32451] + \frac{1}{5}A[32541] - \frac{2}{15}A[34251] \\ &\quad - \frac{2}{15}A[34521] - \frac{2}{15}A[35241] - \frac{2}{15}A[35421] + \frac{1}{5}A[42351] + \frac{1}{5}A[42531] \\ &\quad + \frac{1}{5}A[43251] + \frac{1}{5}A[43521] - \frac{2}{15}A[45231] - \frac{2}{15}A[45321] + \frac{1}{5}A[52341] \\ &\quad + \frac{1}{5}A[52431] + \frac{1}{5}A[53241] + \frac{1}{5}A[53421] + \frac{1}{5}A[54231] + \frac{1}{5}A[54321]. \end{aligned} \tag{59}$$

In sum, the general structure of the Wilcox expansion terms *via* commutators uses the same weights as in Equation (55) and reads

$$W_n = \sum_{k=1}^{(n-1)!} \omega_k^{(n)} A[\sigma_k], \tag{60}$$

where the primed sum requires the rightmost element in the permutation $\sigma(k)$ to be invariant (as in Equation (56)). This element may be chosen at will and, whatever that value, the permutations are build up with the remaining $n - 1$ elements. Different, but equivalent, expressions for W_n in terms of commutators are obtained depending on the value fixed at the rightmost position. We stress once again that, although only the first terms have been collected here for simplicity, the whole procedure is algorithmic in nature and has been implemented in a computer algebra system furnishing to evaluate $W_n(t)$ explicitly for any n [29]. Note that $W_n(t)$ involves a linear combination of $(n - 1)!$ iterated integrals.

3.3. Convergence of Wilcox Expansion

Recursion (33)–(38) is also very useful to provide estimates for the radius of convergence of the Wilcox expansion when Equation (1) is defined in a Banach algebra \mathcal{A} , i.e., an algebra that is also a complete normed linear space with a sub-multiplicative norm,

$$\|XY\| \leq \|X\| \|Y\|. \tag{61}$$

If this is the case, then $\|\text{ad}_X Y\| \leq 2 \|X\| \|Y\|$ and, in general, $\|\text{ad}_X^n Y\| \leq 2^n \|X\|^n \|Y\|$.

As shown in [30,31], if the series

$$M(\lambda; t) = \sum_{j=1}^{\infty} \lambda^j \|W_j(t)\| \tag{62}$$

has a certain radius of convergence r_c for a given t , then, for $\lambda < r < r_c$, the sequence of functions

$$\Psi_n \equiv e^{\lambda W_1(t)} e^{\lambda^2 W_2(t)} \dots e^{\lambda^n W_n(t)} \tag{63}$$

converges uniformly on any compact subset of the ball $B(0, r_c)$. Thus, studying the convergence of the Wilcox expansion reduces analysis of the series $M(\lambda; t)$ and, in particular, its radius of convergence r_c .

Let $k(t)$ be a function such that $\|A(t)\| \leq k(t)$ and denote $K(t) = \int_0^t k(s) ds$. Then, clearly

$$\|b_{0,1}\| \leq k(t); \quad \|b_{0,l}\| = 0, \quad l = 2, 3, \dots \tag{64}$$

and

$$\|W_1\| \leq K(t), \quad \|W_2\| \leq \frac{1}{2} K^2(t). \tag{65}$$

In general, the following bounds can be established by induction

$$\begin{aligned} \|g_{n,r}(t)\| &\leq \beta_{n,r} K^{r-1}(t) k(t), \\ \|b_{n,l}(t)\| &\leq \alpha_{n,l} K^{l-1}(t) k(t), & n = 1, 2, \dots; \quad l > n \\ \|W_n(t)\| &\leq c_n K^n(t), \end{aligned} \tag{66}$$

where

$$\begin{aligned}
 \alpha_{0,1} &= 1; & \alpha_{0,l} &= 0, \quad l > 1 \\
 \alpha_{n,l} &= \sum_{j=0}^{\lfloor \frac{l-1}{n} \rfloor - 1} \frac{1}{j!} 2^j c_n^j (\alpha_{n-1,l-nj} + \beta_{n,l-nj}) & n \geq 1, \quad l > n \\
 \beta_{n,r} &= \begin{cases} \frac{1}{l!} 2^{l-1} n c_n^l, & \lfloor \frac{r}{n} \rfloor = \frac{r}{n} = l \\ 0, & \lfloor \frac{r}{n} \rfloor \neq \frac{r}{n} \end{cases} & (67) \\
 c_1 &= 1, \quad c_2 = \frac{1}{2}, \quad c_n = \frac{1}{n} \alpha_{n-2,n}, & n > 2.
 \end{aligned}$$

It is clear that if the series $\sum_{j \geq 1} c_j K^j(t)$ converges, so does $M(\lambda = 1; t)$. Therefore, a sufficient condition for convergence of the Wilcox expansion is obtained by imposing

$$\lim_{n \rightarrow \infty} \frac{c_{n+1} K^{n+1}(t)}{c_n K^n(t)} = K(t) \lim_{n \rightarrow \infty} D_n < 1, \tag{68}$$

where

$$D_n \equiv \frac{n}{n+1} \frac{\alpha_{n-1,n+1}}{\alpha_{n-2,n}}. \tag{69}$$

We have computed this quantity up to $n = 2000$ and then extrapolated to the limit $(1/n) \rightarrow 0$. Then $D_n \rightarrow D_\infty = 1.51868$, as seen in Figure 1, and thus the convergence of the Wilcox expansion is ensured at least for values of time t such that

$$\int_0^t \|A(s)\| ds \leq K(t) < \xi_W = \frac{1}{D_\infty} \approx 0.65846 \tag{70}$$

This type of extrapolation has also been used to estimate the convergence radius of the Magnus expansion [32]. Although the estimate is not completely analytic, the same type of computation has provided accurate results in other settings. In particular, for the Magnus expansion, such an estimate fully agrees with a purely theoretically deduced bound [8,10,32].

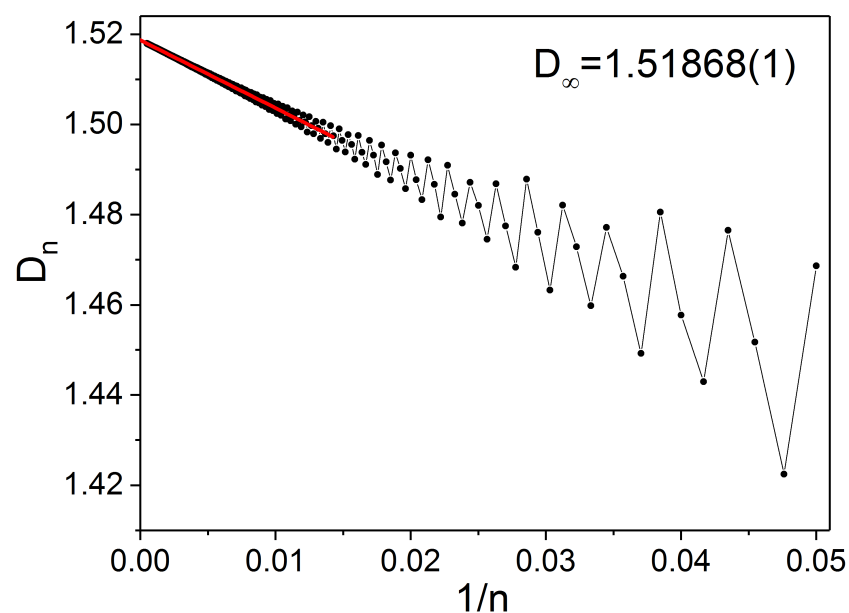


Figure 1. D_n as a function of $1/n$, and linear extrapolation (red line).

4. Fer-Like Expansions

4.1. Standard Fer Expansion

In forming the Wilcox expansion, the first transformation is chosen in such a way that $\dot{\Omega}_1 = B_0(t)$, whereas $\dot{\Omega}_n \neq B_{n-1}$ for $n \geq 2$. It makes sense, then, to analyze what happens if we impose this condition at each step of the procedure

$$\dot{\Omega}_n(t) = B_{n-1}(t) \quad \text{or, equivalently,} \quad \Omega_n(t) = \int_0^t B_{n-1}(u)du \tag{71}$$

for all $n \geq 1$. In that way, expression (22) for B_n clearly simplifies to

$$B_n(t) = \sum_{k \geq 1} \frac{(-1)^k k}{(k+1)!} \text{ad}_{\Omega_n}^k B_{n-1}, \quad n \geq 1. \tag{72}$$

In doing, so we recover the precise Fer expansion, see [10,11]. Again, after n transformations, we get

$$x(t) = e^{\Omega_1(t)} e^{\Omega_2(t)} \dots e^{\Omega_n(t)} X_n(t) \tag{73}$$

so that, if we impose $X_n(t) = x_0$, we are left with another approximation to the exact solution. Notice that this approximation clearly differs from the previous Wilcox expansion for $n \geq 2$, as can be seen by analyzing the dependence on λ of each transformation. Whereas $\Omega_n = \lambda^n W_n(t)$ for the Wilcox expansion, now Ω_n contains terms of order $\lambda^{2^{n-1}}$ and higher. This can be easily shown by induction: Ω_1 is proportional to λ , so that B_1 , according to Equation (72) contains terms of order λ^2 (coming from $[\Omega_1, B_0]$) and higher. In general, B_{n-2} and Ω_{n-1} contain terms of order $\lambda^{2^{n-2}}$ and higher, so the first term in the series (72) for B_{n-1} , i.e., the commutator $[\Omega_{n-1}, B_{n-2}]$, produces a term of order $(\lambda^2)^{2^{n-2}} = \lambda^{2^{n-1}}$ in Ω_n .

Alternatively, expressing Equation (72) as

$$B_n(t) = \int_0^1 dx \int_0^x du e^{-(1-u)\Omega_n} [B_{n-1}, \Omega_n] e^{(1-u)\Omega_n} \tag{74}$$

and taking norms, it is then possible to show that the Fer expansion converges for values of t such that [10]

$$\int_0^t \|A(s)\| ds < 0.8604065. \tag{75}$$

4.2. Intermediate Fer-Like Expansions

Notice that the λ -power series of Ω_n in the Fer expansion contains infinite terms starting with $\lambda^{2^{n-1}}$, but the corresponding truncated factorization obtained from Equation (73) by taking $X_n(t) = x_0$ is correct only up to terms of order $\lambda^{2^{n-1}}$. One might then consider yet another sequence of transformations so that each Ω_k contains only the relevant terms leading to a correct approximation up to this order. Of course, both factorizations would be different, but nevertheless they would produce the correct power series up to order $\lambda^{2^{n-1}}$. The corresponding factorization can be properly called a modified Fer expansion.

Our starting point is, once again, Equation (22). Clearly, the first transformation is the same as in Fer (and Wilcox), i.e.,

$$\dot{\Omega}_1 = B_0 = \lambda A(t), \tag{76}$$

and thus

$$B_1 = \sum_{k=0}^{\infty} \frac{(-1)^k k}{(k+1)!} \text{ad}_{\Omega_1}^k B_0 = \mathcal{O}(\lambda^2), \tag{77}$$

where the rightmost term points out the lowest λ contribution in the sum.

Next, to reproduce the same dependence on λ as the Fer expansion, we need to enforce that $B_2 = \mathcal{O}(\lambda^4)$, and the question is how to choose Ω_2 guaranteeing this feature. An analysis of Equation (22) with $n = 2$ reveals that this is achieved by taking $\dot{\Omega}_2$ as the sum of terms in B_1 in Equation (77) contributing to λ^2 and λ^3 , i.e.,

$$\dot{\Omega}_2 = -\frac{1}{2}\text{ad}_{\Omega_1} B_0 + \frac{1}{3}\text{ad}_{\Omega_1}^2 B_0, \tag{78}$$

since the next term appearing in the expression of B_2 involves the computation of $\text{ad}_{\Omega_1}^3 B_0 = \mathcal{O}(\lambda^4)$. Thus

$$B_2 = \sum_{k=1}^{\infty} \frac{(-1)^k}{(k+1)!} \left((k+1)\text{ad}_{\Omega_2}^k B_1 - \text{ad}_{\Omega_2}^k \dot{\Omega}_2 \right) = \mathcal{O}(\lambda^4). \tag{79}$$

Likewise, Ω_3 is to be designed so that

$$B_3 = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} \left((k+1)\text{ad}_{\Omega_3}^k B_2 - \text{ad}_{\Omega_3}^k \dot{\Omega}_3 \right) = \mathcal{O}(\lambda^8), \tag{80}$$

and this is guaranteed by taking Ω_3 as the sum all the terms in B_2 contributing to powers from λ^4 up to λ^7 . From Equation (79) it is clear that

$$B_2 = \underbrace{-\frac{1}{2}(2\text{ad}_{\Omega_2} B_1 - \text{ad}_{\Omega_2} \dot{\Omega}_2)}_{\mathcal{O}(\lambda^4)} + \underbrace{\frac{1}{3!}(3\text{ad}_{\Omega_2}^2 B_1 - \text{ad}_{\Omega_2}^2 \dot{\Omega}_2)}_{\mathcal{O}(\lambda^6)} + \mathcal{O}(\lambda^8), \tag{81}$$

where only the relevant terms in the expansion in B_1 have to be taken into account. In this way, we can take

$$\dot{\Omega}_3 = -\text{ad}_{\Omega_2} B_1^{[4]} + \frac{1}{2}\text{ad}_{\Omega_2} \dot{\Omega}_2 + \frac{1}{2}\text{ad}_{\Omega_2}^2 B_1^{[2]} - \frac{1}{6}\text{ad}_{\Omega_2}^2 \dot{\Omega}_2 \tag{82}$$

with

$$B_1^{[j]} \equiv \sum_{k=1}^j \frac{(-1)^k}{(k+1)!} k \text{ad}_{\Omega_1}^k B_0. \tag{83}$$

Notice, that since the second term in Ω_2 in Equation (78) is $\mathcal{O}(\lambda^3)$, the expression (82) does contain some contributions in λ^8 and λ^9 that in principle could be removed. We prefer, however, to maintain them in order to have a more compact expression.

For this modified Fer expansion, Ω_n is generally chosen, so that $\dot{\Omega}_n$ is precisely the sum of all terms of B_{n-1} containing terms of powers from $\lambda^{2^{n-1}}$ up to λ^{2^n-1} and then appropriately truncating the series of B_{n-2}, \dots, B_1 .

Other possibilities for choosing B_k at the successive stages clearly exist, and according to the particular election, different intermediate Fer-like expansions result. In practice, one of those combinations of commutators could be more easily computed for a specific problem.

5. Applications

5.1. Wilcox Expansion as the Continuous Analogue of Zassenhaus Formula

The Zassenhaus formula may be considered as the dual of the Baker–Campbell–Hausdorff (BCH) formula [33] in the sense that it relates the exponential of the sum of two non-commuting operators X and Y with an infinite product of exponentials of these operators and their nested commutators. More specifically,

$$e^{X+Y} = e^X e^Y \prod_{n=2}^{\infty} e^{C_n(X,Y)} = e^X e^Y e^{C_2(X,Y)} e^{C_3(X,Y)} \dots e^{C_k(X,Y)} \dots, \tag{84}$$

where $C_k(X, Y)$ is a homogeneous Lie polynomial in X and Y of degree k [1,7,21,34,35]. A very efficient procedure to generate all the terms in Equation (84) is presented in [21] and allows to construct C_n up to a prescribed value of n directly, in terms of the minimum number of independent commutators involving n operators X and Y .

In view of the formal similarity between Equations (39) and (84), Wilcox expansion also has been described as the “continuous analogue of the Zassenhaus formula” [10], just as the Magnus expansion is sometimes called the continuous version of the BCH formula. To substantiate this claim, we next reproduce the Zassenhaus Formula (84) by applying the procedure of Section 3 to a particular initial value problem, namely, the abstract equation

$$\dot{U} = \lambda(X + Y) U, \quad U(0) = I, \tag{85}$$

where X and Y are two non-commuting constant operators.

The formal solution is, of course, $U(t) = e^{t\lambda(X+Y)}$, but we can also solve Equation (85) by first integrating $\dot{U}_0 = \lambda XU_0$ and factorizing $U(t)$ as $U(t) = U_0 U_I = e^{t\lambda X} U_I$, where U_I obeys the equation

$$\dot{U}_I = \lambda e^{-t\lambda X} Y e^{t\lambda X} U_I \equiv A_\lambda(t) U_I, \tag{86}$$

and finally apply this to Equation (86), the sequence of transformations leading to the Wilcox expansion. Notice, however, that now the coefficient matrix $A_\lambda(t)$ is an infinite series in λ

$$A_\lambda(t) = \lambda e^{-t\lambda \text{ad}_X} Y = \sum_{j \geq 0} \frac{(-1)^j}{j!} t^j \lambda^{j+1} \text{ad}_X^j Y, \tag{87}$$

so that, when applying the recursion (33)–(38), $\dot{\Omega}_1$ is no longer $B_0 \equiv A_\lambda(t)$, but the term in $A_\lambda(t)$ which is proportional to λ . In other words,

$$\dot{W}_1 = Y, \quad \text{and thus} \quad W_1(t) = tY. \tag{88}$$

After some computation, one arrives at

$$b_{0,l} = \frac{(-1)^{l-1}}{(l-1)!} t^{l-1} \text{ad}_X^{l-1} Y. \tag{89}$$

Since $\dot{W}_1 = b_{0,1} = Y$, then clearly $g_{1,l} = 0$ for all $l > 1$ and

$$b_{1,l} = \sum_{k=0}^{l-2} \frac{(-1)^k}{k!} \text{ad}_{W_1}^k b_{0,l-k}. \tag{90}$$

By imposing $b_{1,2} = g_{2,2} = \dot{W}_2$, we get

$$\dot{W}_2 = -t \text{ad}_X Y, \quad \text{so that} \quad W_2(t) = -\frac{1}{2} t^2 \text{ad}_X Y. \tag{91}$$

In general, $g_{n,n} = \dot{W}_n$, $g_{n,r} = 0$ when $r \neq n$ and the recurrence (33)–(38) reads now

$$\begin{aligned} \dot{W}_1 &= b_{0,1} \\ b_{n,l} &= \sum_{k=0}^{[l/n]-1} \frac{(-1)^k}{k!} \text{ad}_{W_n}^k b_{n-1,l-nk}, \quad n = 1, 2, \dots \\ \dot{W}_n &= b_{n-2,n} \quad n \geq 2 \end{aligned} \tag{92}$$

together with Equation (89). Working out this recursion we obtain, for the first terms

$$\begin{aligned}
 W_3(t) &= \frac{1}{6}t^3 \text{ad}_X^2 Y + \frac{1}{3}t^3 \text{ad}_Y \text{ad}_X Y \\
 W_4(t) &= -\frac{1}{24}t^4 \text{ad}_X^3 Y - \frac{1}{8}t^4 \text{ad}_Y \text{ad}_X^2 Y - \frac{1}{8}t^4 \text{ad}_Y^2 \text{ad}_X Y \\
 W_5(t) &= \frac{1}{120}t^5 \text{ad}_X^4 Y + \frac{1}{30}t^5 \text{ad}_Y \text{ad}_X^3 Y + \frac{1}{20}t^5 \text{ad}_Y^2 \text{ad}_X^2 Y \\
 &\quad + \frac{1}{30}t^5 \text{ad}_Y^3 \text{ad}_X Y + \frac{1}{20}t^5 \text{ad}_{[X,Y]} \text{ad}_X^2 Y + \frac{1}{10}t^5 \text{ad}_{[X,Y]} \text{ad}_Y \text{ad}_X Y
 \end{aligned}
 \tag{93}$$

One can see that this procedure agrees with the algorithm presented in [21] for every term $W_n, n \geq 1$, in

$$U(t) = e^{t\lambda(X+Y)} = e^{t\lambda X} e^{\lambda W_1(t)} e^{\lambda^2 W_2(t)} e^{\lambda^3 W_3(t)} \dots
 \tag{94}$$

The Zassenhaus formula is recovered by taking $t = 1$, i.e., $C_n(X, Y) = W_n(t = 1)$.

5.2. Expanding the Exponential $\exp(A + \varepsilon B)$

Bellman, in his classic book [36], states that “one of the great challenges of modern physics is that of obtaining useful approximate relations for $e^{(A+\varepsilon B)t}$ in the case where $AB \neq BA$ ”. One such approximation was proposed and left undisclosed in ([12], p. 175). Assuming that $e^{A+\varepsilon B}$ can be written in the form

$$e^{A+\varepsilon B} = e^A e^{\varepsilon C_1} e^{\varepsilon^2 C_2} e^{\varepsilon^3 C_3} \dots,
 \tag{95}$$

Bellman proposed to determine the first three terms C_1, C_2, C_3 , and pointed out that, contrary to other expansions, the product expansion (95) is unitary if A and B are skew-Hermitian.

It turns out that the Wilcox expansion can be used to provide explicit expressions for C_n for any two indeterminates A and B , as we will see in the sequel.

Before proceeding, it is important to remark that this problem differs from the Zassenhaus formula, in the sense that the expansion parameter affects only one of the operators in the exponential. The solution goes as follows. We write

$$U(t) \equiv e^{t(A+\varepsilon B)} = e^{tA} V,
 \tag{96}$$

and solve the differential equation satisfied by V

$$\frac{dV}{dt} = \varepsilon e^{-tA} B e^{tA} V \equiv \varepsilon \tilde{B}(t) V, \quad V(0) = I
 \tag{97}$$

with the Wilcox expansion, so that

$$V(t) = e^{\varepsilon W_1(t)} e^{\varepsilon^2 W_2(t)} e^{\varepsilon^3 W_3(t)} \dots
 \tag{98}$$

The operators C_i in Equation (95) are then obtained by taking $t = 1$.

The successive terms $W_j(t)$ in Equation (98) can be determined either by the recursion (33)–(38) or the explicit expression (56). For the first term, we get

$$W_1(t) = \int_0^t \tilde{B}(s) dt = \int_0^t e^{-s \text{ad}_A} B ds = \sum_{k=0}^{\infty} \frac{(-1)^k t^{k+1}}{(k+1)!} \text{ad}_A^k B
 \tag{99}$$

and so the expression for C_1 is given by

$$C_1 = \frac{1 - e^{\text{ad}_A}}{\text{ad}_A} B = B - \frac{1}{2}[A, B] + \frac{1}{3!}[A, [A, B]] - \frac{1}{4!}[A, [A, [A, B]]] + \dots \tag{100}$$

Although it is possible in principle to construct explicit expressions for W_2, W_3 , etc., it is perhaps more convenient to apply the recursion (33)–(38) for each particular application.

5.3. Illustrative Examples

We next particularize the *Bellman problem* (95) to matrices where closed expressions for C_1, C_2, \dots can be obtained. The idea is to illustrate the behaviour of the product expansion by computing explicitly high-order terms with matrices in the $SU(2)$ and the $SO(3)$ Lie algebras.

5.3.1. Matrices X and Y in $SU(2)$

In the first example, we chose $A = ia\sigma_z$ and $B = i\sigma_x$, where $i = \sqrt{-1}$, a is a real parameter and

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{101}$$

are Pauli matrices. This instance is borrowed from Quantum Mechanics, where $\exp[i(a\sigma_z + \varepsilon\sigma_x)]$ is a matrix that transforms the $\frac{1}{2}$ -spin wave function in a Hilbert space.

Using the scalar product notation to write down a linear combination of Pauli matrices: $\vec{v} \cdot \vec{\sigma} = v_x\sigma_x + v_y\sigma_y + v_z\sigma_z$, the matrix exponential reads

$$\exp(i\vec{v} \cdot \vec{\sigma}) = \cos v I + i \frac{\sin v}{v} \vec{v} \cdot \vec{\sigma}. \tag{102}$$

In the sequel, we work out the expansion

$$e^{i(a\sigma_z + \varepsilon\sigma_x)} = e^{ia\sigma_z} e^{i\varepsilon W_1} e^{i\varepsilon^2 W_2} e^{i\varepsilon^3 W_3} \dots \tag{103}$$

up to order eleven in ε and analyze the increasing accuracy of the product expansion as far as more terms are considered. The lhs in Equation (103) may be thought as a transformation involving σ_x and σ_z . In turn, the rhs is a pure σ_z transformation, i.e., $\exp(ia\sigma_z)$, followed by an infinite succession of transformations, $\exp(i\varepsilon^k W_k)$, whose effect should decrease with k . The truncated product expansion is expected to be accurate as far as $\varepsilon \ll a$.

In Table 1 we write down the first five contributions for a generic t (expressions for $k > 5$ are too involved to be collected here). Wilcox–Bellman’s Formula (103) corresponds then to $t = 1$. All the terms have been obtained with the recurrences of Section 3, starting from

$$\dot{W}_1(t) = e^{-iat\sigma_z} \sigma_x e^{iat\sigma_z} = C\sigma_x + S\sigma_y, \tag{104}$$

where $C \equiv \cos(2at)$ and $S \equiv \sin(2at)$.

The formulas in Table 1 show that ε/a may be considered as an effective expansion parameter. In Figure 2, we illustrate, for $a = 1$, the accuracy of the Wilcox–Bellman product expansion in the example at hand as a function of ε . We plot the squared modulus of the non-diagonal matrix element, say $|U_{1,2}|^2$, of Equation (103) for every analytic approximation up to order eleven in ε , as well as the exact result. Even orders do not contribute in this test, because W_{2k} is always proportional to σ_z , and therefore $\exp(i\varepsilon^{2k} W_{2k})$ is a diagonal matrix.

Table 1. First five orders in Bellman problem for generic t . The operators C_k in Equation (95) are obtained by taking $t = 1$, i.e., $C_k = i W_k(1)$. We have defined $S \equiv \sin(2at)$ and $C \equiv \cos(2at)$.

k	$W_k(t)$
1	$\frac{1}{2a}[S\sigma_x + (1 - C)\sigma_y]$
2	$\frac{1}{4a^2}(2at - S)\sigma_z$
3	$\frac{1}{12a^3}\{[6at + (C - 4)S]\sigma_x - (1 - C)^2\sigma_y\}$
4	$-\frac{1}{16a^4}[6at + (C - 4)S]\sigma_z$
5	$\frac{1}{240a^5}\{[56S - (4C + 7)SC - 10at(7 + 4C - 2C^2)]\sigma_x + [4C^3 - 7C^2 - 28C + 31 + 2at(2SC - 4S + 3at)]\sigma_y\}$

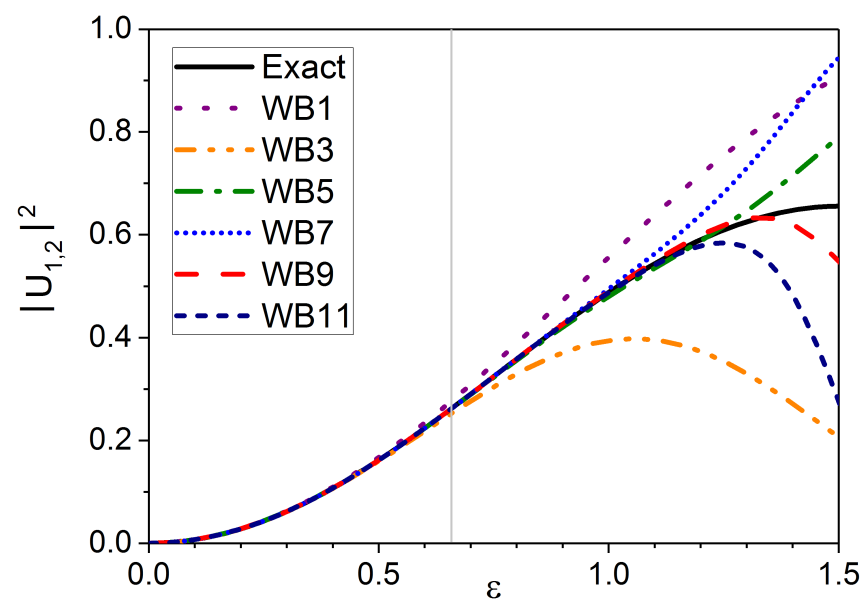


Figure 2. Accuracy of Wilcox–Bellman product expansion up to order eleven as a function of the ratio ε/a , with $a = 1$. The quantity plotted is the squared modulus of the non-diagonal element of the matrix. The vertical grey line stands for the convergence lower bound $\varepsilon = 0.658$.

As regards convergence of the product expansion, the lower bound of Equation (70) leads to

$$\int_0^1 \|e^{-iat\sigma_z} \varepsilon \sigma_x e^{iat\sigma_z}\| dt = \varepsilon < 0.658. \tag{105}$$

In turn, the behaviour of the curves in Figure 2 points out that convergence of the product expansion extends well beyond that lower bound for this particular example.

Eventually, Figure 3 shows the logarithm of the absolute error in the approximations given by curves in Figure 2.

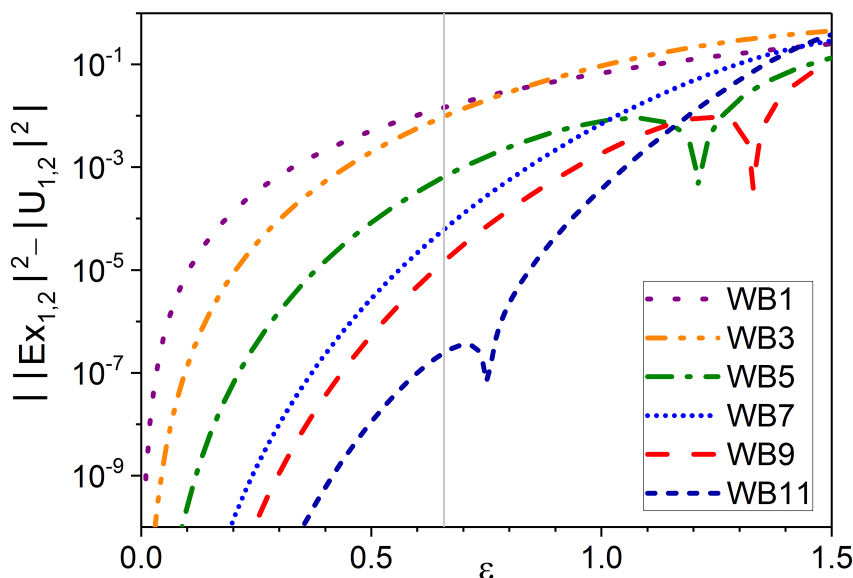


Figure 3. Absolute error of the approximations given by curves in Figure 2 with $a = 1$. The vertical grey line is located at the value of the convergence lower bound $\epsilon = 0.658$.

5.3.2. Matrices X and Y in SO(3)

The second example refers to the matrix that describes a rotation in three dimensions defined by the vector $\vec{a} = \alpha \hat{a}$. Here, α stands for the rotation angle around the axis given by the unitary vector \hat{a} . A generic 3D rotation matrix can be written as $\exp(\vec{a} \cdot \vec{\rho})$, where the components of $\vec{\rho}$ are the three fundamental rotation matrices

$$\rho_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \rho_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \rho_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (106)$$

We study the particular case $\vec{a} \cdot \vec{\rho} = \alpha (\cos \theta \rho_z + \sin \theta \rho_x)$, and compare the rotation of angle α around the unit vector $(\sin \theta, 0, \cos \theta)$

$$e^{\alpha (\cos \theta \rho_z + \sin \theta \rho_x)} \quad (107)$$

with the sequence of transformations

$$e^{\alpha \cos \theta \rho_z} e^{\alpha \sin \theta W_1} e^{(\alpha \sin \theta)^2 W_2} e^{(\alpha \sin \theta)^3 W_3} \dots \quad (108)$$

In other words, the question we address is how the pure z-axis rotation in Equation (108), $\exp(\alpha \cos \theta \rho_z)$, has to be corrected by an infinite composition of further rotations to reproduce the one defined by $\vec{a} \cdot \vec{\rho}$. When $\alpha \sin \theta$ is small enough, the approach is expected to converge, since the expansion convergence lower bound reads, in this case, $|\alpha \sin \theta| < 0.658$.

Here, the accuracy of the product expansion will depend on both the rotation angle α and the relative orientation of the rotation axis, determined by the angle θ . This is illustrated in Figures 4 and 5 for the first five orders of approximation, with $\alpha = \pi/4, \pi/2, 3\pi/4$ and π .

In order to test the expansion, we have computed the matrix trace of the successive approximants and compared them with the exact result

$$\text{tr}(\exp(\alpha (\cos \theta \rho_z + \sin \theta \rho_x))) = 1 + 2 \cos \alpha. \quad (109)$$

The first two approximants are simple enough to be written down

$$\begin{aligned} \text{tr}\left(e^{\alpha \cos \theta \rho_z} e^{\alpha \sin \theta W_1}\right) &= \left(1 + \cos(\alpha \cos \theta)\right) \cos\left(2 \tan \theta \sin \frac{\alpha \cos \theta}{2}\right) \\ \text{tr}\left(e^{\alpha \cos \theta \rho_z} e^{\alpha \sin \theta W_1} e^{(\alpha \sin \theta)^2 W_2}\right) &= \cos\left(\frac{1}{2} \tan^2 \theta (\sin(\alpha \cos \theta) - \alpha \cos \theta)\right) \cdot \cos\left(\alpha \cos \theta - \frac{1}{2} \tan \theta \sin \frac{\alpha \cos \theta}{2}\right) \end{aligned} \quad (110)$$

Interestingly, the third approximation order is worse than the second one in all four cases. In the case of $\alpha = \pi$, the fourth order is better than the fifth one.

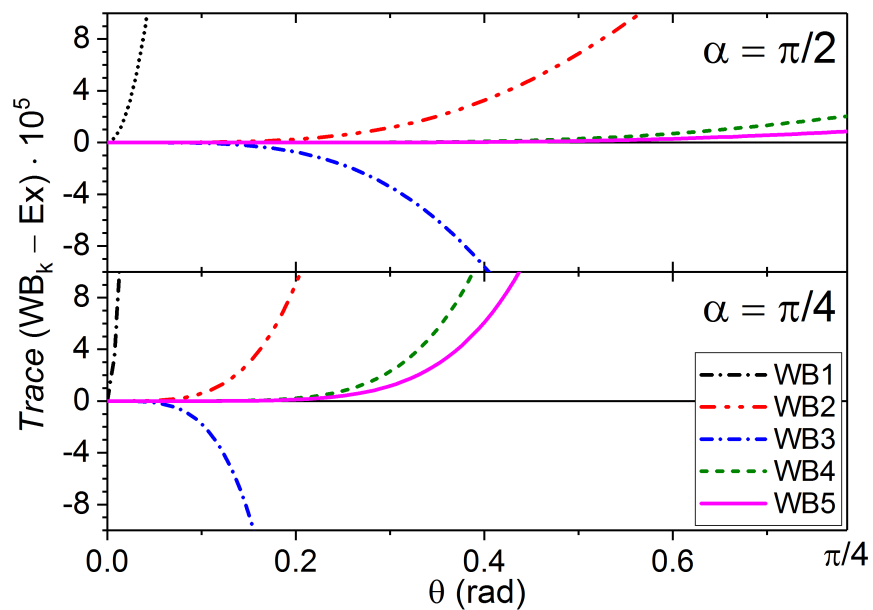


Figure 4. Error in the approximations to the matrix trace as a function of the rotation angle θ , for two values $\alpha = \pi/2$ and $\pi/4$.

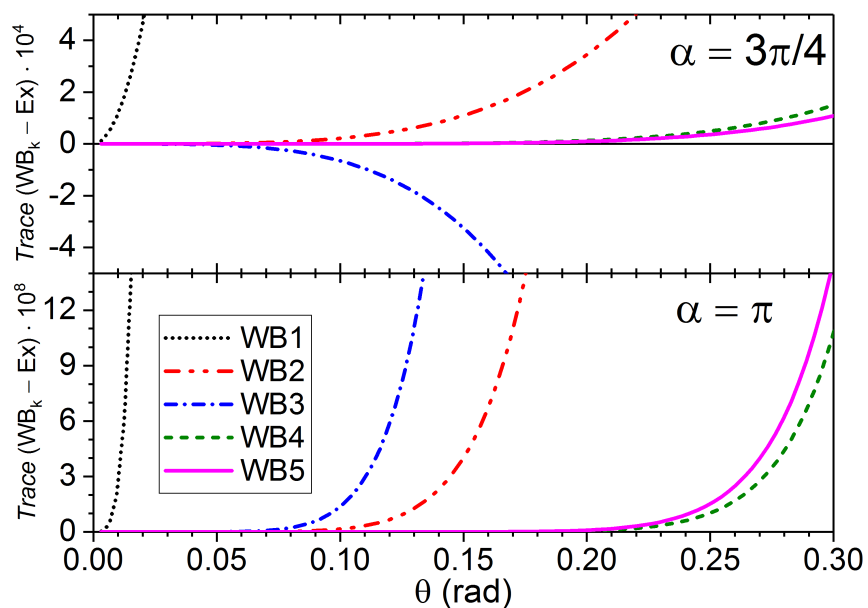


Figure 5. Error in the approximations to the matrix trace as a function of the rotation angle θ , for two values $\alpha = 3\pi/4$ and π .

6. Conclusions

When a linear system of differential equations, defined by the coefficient matrix $\lambda A(t)$, is transformed under $\exp(\lambda \int_0^t A(t') dt')$, the coefficient matrix in the new representation becomes an infinite power series in λ , say $\tilde{A}(t) = \sum_{k \geq 1} \lambda^k a_k(t)$. That is the first step of all matrix exponential methods to approximate the time-evolution operator. In the framework that we have introduced, it is the first move in a sequence of exponential transformations that change the linear system from one representation to another with the goal that dynamics will become less and less relevant. Choosing the transformation $\exp(\lambda \int_0^t \tilde{A}(x) dx)$ as the second move and iterating this procedure afterwards yields the Fer expansion. Instead, choosing the transformation $\exp(\lambda \int_0^t a_1(x) dx)$, i.e., the leading term of the new coefficient matrix, opens up Wilcox expansion. The framework allows for intermediate expansions, taking $\exp(\sum_{k=1}^n \lambda^k \int_0^t a_k(x) dx)$ as an initiator, as well as jumping between schemes, in accordance with the particular requirements of the problem at hand.

We have seen that the theory of linear transformations (or changes of picture in the language of Quantum Mechanics) provides a unified framework to deal with many different exponential perturbative expansions. Whereas only one linear transformation reducing the dynamics to the trivial Equation (4) or to a system with a constant matrix renders the Magnus [9] and the Floquet–Magnus [37] expansion, respectively, a sequence of such transformations with different choices of the new matrices lead to Wilcox and Fer factorizations. From this perspective, other factorizations are possible depending on the particular problem at hand: one only has to appropriately select the successive transformations.

In the case of Wilcox expansion, we have provided an efficient recursive procedure to compute this. In addition, we have developed a method to build up an explicit expression for any W_n in terms of commutators. This is possible by using similar tools, as in the case of the Magnus expansion, namely by relating products of iterated integrals with the structure of the Hopf algebra of permutations, and by using special bases of nested commutators. A sufficient condition for the expansion convergence has also been obtained.

We have presented some application examples of the results about Wilcox expansion. Firstly, we have shown how to obtain Zassenhaus formula from Wilcox expansion which, in turn, may be interpreted as its continuous analogue. Secondly, we point out that Wilcox expansion solves the problem of expanding the exponential $\exp(A + \epsilon B)$ when A and B are non-commuting operators. We refer to this as Wilcox–Bellman expansion. Two practical cases, in this respect, have been analyzed up to high order. Interestingly, in one of them the convergence seems to not be uniform. For convenience, the interested reader can find, in [29], a *Mathematica* code generating general explicit expressions and recurrences for the Wilcox expansion.

While a full assessment of the Wilcox expansion in comparison with Fer expansion is not the main purpose of this work, we can still mention some of their most distinctive features. Both types of expansion construct the solution of Equation (1) as an infinite exponential factorization, but in Wilcox, the exponent of each factor is proportional to successive powers of the expansion parameter λ , whereas, in Fer, each exponent contains an infinite sum of powers of λ . This means that, when truncated after a given number of transformations, say n , the Wilcox expansion differs from the exact solution in the power λ^{n+1} . In other words, each term $W_k(t)$ in the Wilcox expansion collects the effect of the perturbation at order k . On the other hand, the Fer expansion, when truncated after n transformations, provides a much more accurate approximation. This is true, of course, if the infinite sums involved in each transformation are exactly computed, an almost impossible task unless the time dependence of $A(t)$ is simple enough. By contrast, we have explicit expressions for each exponent $W_k(t)$ in the Wilcox expansion for a generic $A(t)$ and, by using the same techniques as in the Magnus expansion, we can construct appropriate approximations of the iterated integrals if necessary. As the examples collected here and in some other studies show [20], Wilcox expansion can provide accurate results after only a few such transformations.

Author Contributions: All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: The first three authors have been funded by Ministerio de Ciencia e Innovación (Spain) through projects MTM2016-77660-P and PID2019-104927GB-C21 (AEI/FEDER, UE), and by Universitat Jaume I (grants UJI-B2019-17 and GACUJI/2020/05). The work of J.A.O. has been partially supported by the Spanish MINECO (grant numbers AYA2016-81065-C2-2 and PID2019-109592GB-I00).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data that support the findings of this study are available from the corresponding author upon request. Codes generating explicit expressions and recurrences for the Wilcox expansion are openly available in <http://www.gicas.uji.es/Research/Wilcox.html> (accessed on 16 March 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wilcox, R. Exponential operators and parameter differentiation in quantum physics. *J. Math. Phys.* **1967**, *8*, 962–982. [[CrossRef](#)]
2. Oteo, J.; Ros, J. From time-ordered products to Magnus expansion. *J. Math. Phys.* **2000**, *41*, 3268–3277. [[CrossRef](#)]
3. Hale, J. *Ordinary Differential Equations*; Krieger Publishing Company: Malabar, FL, USA, 1980.
4. Maricq, M.M. Application of average Hamiltonian theory to the NMR of solids. *Phys. Rev. B* **1982**, *25*, 6662–6632. [[CrossRef](#)]
5. Casas, F.; Oteo, J.A.; Ros, J. Floquet theory: Exponential perturbative treatment. *J. Phys. A Math. Gen.* **2001**, *34*, 3379–3388. [[CrossRef](#)]
6. Mananga, E.; Charpentier, T. On the Floquet–Magnus expansion: Applications in solid-state nuclear magnetic resonance and physics. *Phys. Rep.* **2016**, *609*, 1–50. [[CrossRef](#)]
7. Magnus, W. On the Exponential Solution of Differential Equations for a Linear Operator. *Comm. Pure Appl. Math.* **1954**, *VII*, 649–673. [[CrossRef](#)]
8. Blanes, S.; Casas, F.; Oteo, J.; Ros, J. The Magnus expansion and some of its applications. *Phys. Rep.* **2009**, *470*, 151–238. [[CrossRef](#)]
9. Casas, F.; Chartier, P.; Murua, A. Continuous changes of variables and the Magnus expansion. *J. Phys. Commun.* **2019**, *3*, 095014. [[CrossRef](#)]
10. Blanes, S.; Casas, F.; Oteo, J.; Ros, J. Magnus and Fer expansions for matrix differential equations: The convergence problem. *J. Phys. A Math. Gen.* **1998**, *22*, 259–268. [[CrossRef](#)]
11. Fer, F. Résolution de l'équation matricielle $\dot{U} = pU$ par produit infini d'exponentielles matricielles. *Bull. Classe Sci. Acad. R. Bel.* **1958**, *44*, 818–829.
12. Bellman, R. *Introduction to Matrix Analysis*, 2nd ed.; McGraw-Hill: New York, NY, USA, 1970.
13. Iserles, A. Solving Linear Ordinary Differential Equations by Exponentials of Iterated Commutators. *Numer. Math.* **1984**, *45*, 183–199. [[CrossRef](#)]
14. Klarsfeld, S.; Oteo, J. Exponential infinite-product representations of the time displacement operator. *J. Phys. A Math. Gen.* **1989**, *22*, 2687–2694. [[CrossRef](#)]
15. Mananga, E. On the Fer expansion: Applications in solid-state nuclear magnetic resonance and physics. *Phys. Rep.* **2016**, *608*, 1–41. [[CrossRef](#)]
16. Casas, F. Fer's Factorization as a Symplectic Integrator. *Numer. Math.* **1996**, *74*, 283–303. [[CrossRef](#)]
17. Zanna, A. Collocation and Relaxed Collocation for the Fer and the Magnus Expansions. *SIAM J. Numer. Anal.* **1999**, *36*, 1145–1182. [[CrossRef](#)]
18. Iserles, A.; Munthe-Kaas, H.; Nørsett, S.; Zanna, A. Lie-Group Methods. *Acta Numer.* **2000**, *9*, 215–365. [[CrossRef](#)]
19. Huillet, T.; Monin, A.; Salut, G. Lie algebraic canonical representation in nonlinear control systems. *Math. Syst. Theory* **1987**, *20*, 193–213. [[CrossRef](#)]
20. Zagury, N.; Aragão, A.; Casanova, J.; Solano, E. Unitary expansion of the time evolution operator. *Phys. Rev. A* **2010**, *82*, 042110. [[CrossRef](#)]
21. Casas, F.; Murua, A.; Nadinic, M. Efficient computation of the Zassenhaus formula. *Comput. Phys. Commun.* **2012**, *183*, 2386–2391. [[CrossRef](#)]
22. Galindo, A.; Pascual, P. *Quantum Mechanics*; Springer: Berlin/Heidelberg, Germany, 1990.
23. Arnal, A.; Casas, F.; Chiralt, C. A general formula for the Magnus expansion in terms of iterated integrals of right-nested commutators. *J. Phys. Commun.* **2018**, *2*, 035024. [[CrossRef](#)]
24. Dragt, A.; Forest, E. Computation of nonlinear behavior of Hamiltonian systems using Lie algebraic methods. *J. Math. Phys.* **1983**, *24*, 2734–2744. [[CrossRef](#)]
25. Agrachev, A.; Gamkrelidze, R. The shuffle product and symmetric groups. In *Differential Equations, Dynamical Systems, and Control Science*; Elworthy, K., Everitt, W., Lee, E., Eds.; Marcel Dekker: New York, NY, USA, 1994; pp. 365–382.

26. Hazewinkel, M.; Gubareni, N.; Kirichenko, V. *Algebras, Rings and Modules. Lie Algebras and Hopf Algebras*; AMS: Providence, RI, USA, 2010.
27. Malvenuto, C.; Reutenauer, C. Duality between quasi-symmetric functions and the Solomon descent algebra. *J. Algebra* **1995**, *177*, 967–982. [[CrossRef](#)]
28. Strichartz, R.S. The Campbell–Baker–Hausdorff–Dynkin Formula and Solutions of Differential Equations. *J. Funct. Anal.* **1987**, *72*, 320–345. [[CrossRef](#)]
29. Geometric Integration Research Group. 2021. Available online: <http://www.gicas.uji.es/Research/Wilcox.html> (accessed on 18 February 2021).
30. Bayen, F. On the convergence of the Zassenhaus formula. *Lett. Math. Phys.* **1979**, *3*, 161–167. [[CrossRef](#)]
31. Arnal, A.; Casas, F.; Chiralt, C. On the structure and convergence of the symmetric Zassenhaus formula. *Comput. Phys. Commun.* **2017**, *217*, 58–65. [[CrossRef](#)]
32. Moan, P.; Oteo, J. Convergence of the exponential Lie series. *J. Math. Phys.* **2001**, *42*, 501–508. [[CrossRef](#)]
33. Casas, F.; Murua, A. An efficient algorithm for computing the Baker–Campbell–Hausdorff series and some of its applications. *J. Math. Phys.* **2009**, *50*, 033513. [[CrossRef](#)]
34. Suzuki, M. On the convergence of exponential operators—The Zassenhaus formula, BCH formula and systematic approximants. *Commun. Math. Phys.* **1977**, *57*, 193–200. [[CrossRef](#)]
35. Weyrauch, M.; Scholz, D. Computing the Baker–Campbell–Hausdorff series and the Zassenhaus product. *Comput. Phys. Commun.* **2009**, *180*, 1558–1565. [[CrossRef](#)]
36. Bellman, R. *Perturbation Techniques in Mathematics, Engineering & Physics*; Dover Publications: Mineola, NY, USA, 1972.
37. Arnal, A.; Casas, F.; Chiralt, C. Exponential perturbative expansions and coordinate transformations. *Math. Comput. Appl.* **2020**, *25*, 50. [[CrossRef](#)]