

PAPER • OPEN ACCESS

Continuous changes of variables and the Magnus expansion

To cite this article: Fernando Casas *et al* 2019 *J. Phys. Commun.* **3** 095014

View the [article online](#) for updates and enhancements.



PAPER

Continuous changes of variables and the Magnus expansion

OPEN ACCESS

RECEIVED
24 July 2019REVISED
4 September 2019ACCEPTED FOR PUBLICATION
9 September 2019PUBLISHED
23 September 2019Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Fernando Casas^{1,4} , Philippe Chartier² and Ander Murua³¹ Universitat Jaume I, IMAC, Departament de Matemàtiques, 12071 Castellón, Spain² INRIA, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes, France³ Konputazio Zientziak eta A.A. Saila, Infomatika Fakultatea, UPV/EHU, 20018 Donostia-San Sebastián, Spain⁴ Author to whom any correspondence should be addressed.E-mail: fernando.casas@uji.es, philippe.chartier@inria.fr and ander.murua@ehu.es

Keywords: magnus expansion, continuous changes of variables, averaging

Abstract

In this paper, we are concerned with a formulation of Magnus and Floquet-Magnus expansions for general nonlinear differential equations. To this aim, we introduce suitable continuous variable transformations generated by operators. As an application of the simple formulas so-obtained, we explicitly compute the first terms of the Floquet-Magnus expansion for the Van der Pol oscillator and the nonlinear Schrödinger equation on the torus.

1. Introduction

The Magnus expansion constitutes nowadays a standard tool for obtaining both analytic and numerical approximations to the solutions of non-autonomous linear differential equations. In its simplest formulation, the Magnus expansion [1] aims to construct the solution of the linear differential equation

$$\dot{Y}(t) = A(t)Y(t), \quad Y(0) = I, \quad (1)$$

where $A(t)$ is a $n \times n$ matrix, as

$$Y(t) = \exp \Omega(t), \quad (2)$$

where Ω is an infinite series

$$\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t), \quad \text{with} \quad \Omega_k(0) = 0, \quad (3)$$

whose terms are increasingly complex expressions involving iterated integrals of nested commutators of the matrix A evaluated at different times.

Since the 1960s the Magnus expansion (often with different names) has been used in many different fields, ranging from nuclear, atomic and molecular physics to nuclear magnetic resonance and quantum electrodynamics, mainly in connection with perturbation theory. More recently, it has also been the starting point to construct numerical integration methods in the realm of geometric numerical integration (see [2] for a review), when preserving the main qualitative features of the exact solution, such as its invariant quantities or the geometric structure is at issue [3, 4]. The convergence of the expansion is also an important feature and several general results are available [5–8].

Given the favourable properties exhibited by the Magnus expansion in the treatment of the linear problem (1), it comes as no surprise that several generalizations have been proposed along the years. We can mention, in particular, equation (1) when the (in general complex) matrix-valued function $A(t)$ is periodic with period T . In that case, it is possible to combine the Magnus expansion with the Floquet theorem [9] and construct the solution as

$$Y(t) = \exp(\Lambda(t)) \exp(tF), \quad (4)$$

where $\Lambda(t + T) = \Lambda(t)$ and both $\Lambda(t)$ and F are series expansions

$$\Lambda(t) = \sum_{k=1}^{\infty} \Lambda_k(t), \quad F = \sum_{k=1}^{\infty} F_k, \tag{5}$$

with $\Lambda_k(0) = 0$ for all k . This is the so-called Floquet–Magnus expansion [10], and has been widely used in problems of solid state physics and nuclear magnetic resonance [11, 12]. Notice that, due to the periodicity of Λ_k , the constant term F_n can be independently obtained as $F_k = \Omega_k(T)/T$ for all k .

In the general case of a nonlinear ordinary differential equation in \mathbb{R}^n ,

$$\dot{x} = g(x, t), \quad x(0) = x_0 \in \mathbb{R}^n, \tag{6}$$

the usual procedure to construct the Magnus expansion requires first to transform (6) into a certain linear equation involving operators [13]. This is done by introducing the Lie derivative associated with g and the family of linear transformations Φ_t such that $\Phi_t[f] = f \circ \varphi_t$, where φ_t denotes the exact flow defined by (6) and f is any (infinitely) differentiable map $f: \mathbb{R}^n \rightarrow \mathbb{R}$. The operator Φ_t obeys a linear differential equation which is then formally solved with the corresponding Magnus expansion [2]. Once the series is truncated, it corresponds to the Lie derivative of some function $W(x, t)$. Finally, the solution at some given time $t = T$ can be approximated by determining the 1-flow of the autonomous differential equation

$$\dot{y} = W(y, T), \quad y(0) = x_0$$

since, by construction, $y(1) \simeq \varphi_T(x_0)$. Clearly, the whole procedure is different and more involved than in the linear case. It is the purpose of this work to provide a unified framework to derive the Magnus expansion in a simpler way without requiring the apparatus of chronological calculus. This will be possible by applying the continuous transformation theory developed by Dewar in perturbation theory in classical mechanics [14]. In that context, the Magnus series is just the generator of the continuous transformation sending the original system (6) to the trivial one $\dot{X} = 0$. Moreover, the same idea can be applied to the Floquet–Magnus expansion, thus establishing a natural connection with the stroboscopic averaging formalism. In the process, the relation with pre-Lie algebras and other combinatorial objects will appear in a natural way.

The plan of the paper is as follows. We review several procedures to derive the Magnus expansion for the linear equation (1) in section 2 and introduce a binary operator that will play an important role in the sequel. In section 3 we consider continuous changes of variables and their generators in the context of general ordinary differential equations, whereas in sections 4 and 5 we apply this formalism for constructing the Magnus and Floquet–Magnus expansions, respectively, in the general nonlinear setting. There, we also show how they reproduce the classical expansions for linear differential equations. As a result, both expansions can be considered as the output of appropriately continuous changes of variables rendering the original system into a simpler form. Finally, in section 6 we illustrate the techniques developed here by considering two examples: the Van der Pol oscillator and the nonlinear Schrödinger equation with periodic boundary conditions.

2. The Magnus expansion for linear systems

There are many ways to get the terms of the Magnus series (3). If we introduce a (dummy) parameter ε in equation (1), i.e., we replace A by εA , then the successive terms in

$$\Omega(t) = \varepsilon \Omega_1(t) + \varepsilon^2 \Omega_2(t) + \varepsilon^3 \Omega_3(t) + \dots \tag{7}$$

can be determined by inserting $\Omega(t)$ into equation (1) and computing the derivative of the matrix exponential, thus arriving at [15]

$$\varepsilon A = d \exp_{\Omega}(\dot{\Omega}) \equiv \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \text{ad}_{\dot{\Omega}}^k(\Omega) = \dot{\Omega} + \frac{1}{2}[\Omega, \dot{\Omega}] + \frac{1}{3!}[\Omega, [\Omega, \dot{\Omega}]] + \dots \tag{8}$$

where $[A, B]$ denotes the usual Lie bracket (commutator) and $\text{ad}_A^j B = [A, \text{ad}_A^{j-1} B]$, $\text{ad}_A^0 B = B$. At this point it is useful to introduce the linear operator

$$H(t) = (F \triangleright G)(t) := \int_0^t [F(u), G(t)] du \tag{9}$$

so that, in terms of

$$R(t) = \frac{d}{dt} \Omega(t) = \varepsilon R_1(t) + \varepsilon^2 R_2(t) + \varepsilon^3 R_3(t) + \dots, \tag{10}$$

equation (8) can be written as

$$\varepsilon A = R + \frac{1}{2} R \triangleright R + \frac{1}{3!} R \triangleright R \triangleright R + \frac{1}{4!} R \triangleright R \triangleright R \triangleright R + \dots \tag{11}$$

Here we have used the notation

$$F_1 \triangleright F_2 \triangleright \dots \triangleright F_m = F_1 \triangleright (F_2 \triangleright \dots \triangleright F_m).$$

Now the successive terms $R_j(t)$ can be determined by substitution of (10) into (11) and comparing like powers of ε . In particular, this gives

$$\begin{aligned} R_1 &= A, \\ R_2 &= -\frac{1}{2} R_1 \triangleright R_1 = -\frac{1}{2} A \triangleright A, \\ R_3 &= -\frac{1}{2} (R_2 \triangleright R_1 + R_1 \triangleright R_2) - \frac{1}{6} R_1 \triangleright R_1 \triangleright R_1 \\ &= \frac{1}{4} (A \triangleright A) \triangleright A + \frac{1}{12} A \triangleright A \triangleright A \end{aligned}$$

Of course, equation (8) can be inverted, thus resulting in

$$\dot{\Omega} = \sum_{k=0}^{\infty} \frac{B_k}{k!} \text{ad}_{\dot{\Omega}}^k(\varepsilon A(t)), \quad \Omega(0) = 0 \tag{12}$$

where the B_j are the Bernoulli numbers, that is

$$\begin{aligned} \frac{x}{e^x - 1} &= 1 + B_1 x + \frac{B_2}{2!} x^2 + \frac{B_4}{4!} x^4 + \frac{B_6}{6!} x^6 + \dots \\ &= 1 - \frac{1}{2} x + \frac{1}{12} x^2 - \frac{1}{720} x^4 + \dots \end{aligned}$$

In terms of R , equation (12) can be written as

$$\varepsilon^{-1} R = A - B_1 R \triangleright A + \frac{B_2}{2!} R \triangleright R \triangleright A + \frac{B_4}{4!} R \triangleright R \triangleright R \triangleright R \triangleright A + \dots \tag{13}$$

Substituting (10) in equation (13) and working out the resulting expression, one arrives to the following recursive procedure allowing to determine the successive terms $R_j(t)$:

$$\begin{aligned} S_m^{(1)} &= [\Omega_{m-1}, A], & S_m^{(j)} &= \sum_{n=1}^{m-j} [\Omega_n, S_{m-n}^{(j-1)}], \quad 2 \leq j \leq m-1 \\ R_1(t) &= A(t), & R_m(t) &= \sum_{j=1}^{m-1} \frac{B_j}{j!} S_m^{(j)}(t), \quad m \geq 2. \end{aligned} \tag{14}$$

Notice in particular that

$$S_2^{(1)} = A \triangleright A, \quad S_3^{(1)} = -\frac{1}{2} (A \triangleright A) \triangleright A, \quad S_3^{(2)} = A \triangleright A \triangleright A.$$

At this point it is worth remarking that any of the above procedures can be used to write each R_j in terms of the binary operation \triangleright and the original time-dependent linear operator A , which gives in general one term per binary tree, as in [15, 16], or equivalently, one term per planar rooted tree. However, the binary operator \triangleright satisfies, as a consequence of the Jacobi identity of the Lie bracket of vector fields and the integration by parts formula, the so-called pre-Lie relation

$$F \triangleright G \triangleright H - (F \triangleright G) \triangleright H = G \triangleright F \triangleright H - (G \triangleright F) \triangleright H, \tag{15}$$

As shown in [17], this relation can be used to rewrite each R_j as a sum of fewer terms, the number of terms being less than or equal to the number of rooted trees with j vertices. For instance, the formula for R_4 can be written in the simplified form

$$R_4 = -\frac{1}{6} ((A \triangleright A) \triangleright A) \triangleright A - \frac{1}{12} A \triangleright (A \triangleright A) \triangleright A$$

upon using the pre-Lie relation (15) for $F = G \triangleright G$ and $H = G$.

If, on the other hand, one is more interested in getting an explicit expression for $\Omega_j(t)$, the usual starting point is to express the solution of (1) as the series

$$Y(t) = I + \sum_{n=1}^{\infty} \int_{\Delta_n(t)} A(t_1) A(t_2) \dots A(t_n) dt_1 \dots dt_n, \tag{16}$$

where

$$\Delta_n(t) = \{(t_1, \dots, t_n): 0 \leq t_n \leq \dots \leq t_1 \leq t\} \tag{17}$$

and then compute formally the logarithm of (16). Then one gets [18–21]

$$\Omega(t) = \log Y(t) = \sum_{n=1}^{\infty} \Omega_n(t),$$

with

$$\Omega_n(t) = \frac{1}{n} \sum_{\sigma \in S_n} (-1)^{d_\sigma} \frac{1}{\binom{n-1}{d_\sigma}} \int_{\Delta_n(t)} A(t_{\sigma(1)}) A(t_{\sigma(2)}) \cdots A(t_{\sigma(n)}) dt_1 \cdots dt_n. \tag{18}$$

Here $\sigma \in S_n$ denotes a permutation of $\{1, 2, \dots, n\}$. An expression in terms only of independent commutators can be obtained by using the class of bases proposed by Dragt and Forest [22] for the Lie algebra generated by the operators $A(t_1), \dots, A(t_n)$, thus resulting in [23]

$$\begin{aligned} \Omega_n(t) = & \frac{1}{n} \sum_{\sigma \in S_{n-1}} (-1)^{d_\sigma} \frac{1}{\binom{n-1}{d_\sigma}} \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n \\ & [A(t_{\sigma(1)}), [A(t_{\sigma(2)}), \cdots [A(t_{\sigma(n-1)}), A(t_n)] \cdots]], \end{aligned} \tag{19}$$

where now σ is a permutation of $\{1, 2, \dots, n - 1\}$ and d_σ corresponds to the number descents of σ . We recall that σ has a descent in i if $\sigma(i) > \sigma(i + 1), i = 1, \dots, n - 2$.

3. Continuous changes of variables

Our purpose in the sequel is to generalize the previous expansion to general nonlinear differential equations. It turns out that a suitable tool for that purpose is the use of continuous variable transformations generated by operators [14, 24]. We therefore summarize next its main features.

Given a generic ODE system of the form

$$\frac{d}{dt}x = f(x, t), \tag{20}$$

the idea is to apply some near-to-identity change of variables $x \mapsto X$ that transforms the original system (20) into

$$\frac{d}{dt}X = F(X, t), \tag{21}$$

where the vector field $F(X, t)$ adopts some desirable form. In order to do that in a convenient way, we apply a one-parameter family of time-dependent transformations of the form

$$z = \Psi_s(X, t), \quad s \in \mathbb{R},$$

such that $\Psi_0(X, t) \equiv X$, and $x = \Psi_1(X, t)$ is the change of variables that we seek. In this way, one continuously varies s from $s = 0$ to $s = 1$ to move from the trivial change of variables $x = X$ to $x = \Psi_1(X, t)$, so that for each solution $X(t)$ of (21), the function $z(t, s)$ defined by $z(t, s) = \Psi_s(X(t), t)$ satisfies a differential equation

$$\frac{\partial}{\partial t}z = V(z, t, s). \tag{22}$$

In particular, we will have that $F(X, t) = V(X, t, 0)$ and $f(x, t) = V(x, t, 1)$.

Next, the near-to-identity family of maps $X \mapsto z = \Psi_s(X, t)$ is defined in terms of a differential equation in the independent variable s ,

$$\frac{\partial}{\partial s}z(t, s) = W(z(t, s), t, s) \tag{23}$$

by requiring that $z(t, s) = \Psi_s(z(t, 0), t)$ for any solution $z(t, s)$ of (23). The map $\Psi_s(\cdot, t)$ will be near-to-identity if $W(z, t, s)$ is of the form

$$W(z, t, s) = \varepsilon W_1(z, t, s) + \varepsilon^2 W_2(z, t, s) + \cdots,$$

for some small parameter ε .

Proposition 1 ([14]). *Given F and $W = \varepsilon W_1 + \varepsilon^2 W_2 + \cdots$, the right-hand side V of the continuously transformed system (22) can be uniquely determined (as a formal series in powers of ε) from $V(X, t, 0) = F(X, t)$ and*

$$\frac{\partial}{\partial s}V(x, t, s) - \frac{\partial}{\partial t}W(x, t, s) = W'(x, t, s)V(x, t, s) - V'(x, t, s)W(x, t, s), \tag{24}$$

where W' and V' refer to the differentials $\partial_x W$ and $\partial_x V$, respectively.

Proof. By partial differentiation of both sides in (23) with respect to t and partial differentiation of both sides in (22) with respect to s , we conclude that (24) holds for all $x = z(s, t) = \Psi_s(x_0, t)$ with arbitrary x_0 and all (t, s) . One can show that the equality (24) holds for arbitrary (x, t, s) by taking into account that, for given t and s , $x_0 \mapsto x = \Psi_s(x_0, t)$ is one-to-one.

Now, since $V(x, t, 0) = F(x, t)$, we have that

$$V(x, t, s) = F(x, t) + \int_0^s S(x, t, \sigma) d\sigma, \tag{25}$$

where $S = \frac{\partial}{\partial t}W + W'V - V'W$. Clearly, the successive terms of

$$V = F + \varepsilon V_1 + \varepsilon^2 V_2 + \dots$$

are uniquely determined by equating like powers of ε in (25). ■

In the sequel we always assume that the generator W of the change of variables

- (i) does not depend on s , and
- (ii) $W(x, 0, s) \equiv 0$, so that $\Psi_s(x, 0) = x$ and $x(0) = X(0)$.

The successive terms in the expansion of $V(x, t, s)$ in proposition 1 can be conveniently computed with the help of a binary operation \triangleright on maps $\mathbb{R}^{d+1} \rightarrow \mathbb{R}^d$ defined as follows. Given two such maps P and Q , then $P \triangleright Q$ is a new map whose evaluation at $(x, t) \in \mathbb{R}^{d+1}$ takes the value

$$(P \triangleright Q)(x, t) = \int_0^t (P'(x, \tau)Q(x, t) - Q'(x, t)P(x, \tau))d\tau. \tag{26}$$

Under these conditions, from proposition 1, we have that

$$\frac{\partial}{\partial s}V(x, t, s) - \frac{\partial}{\partial t}W(x, t) = [W(x, t), V(x, t, s)] \tag{27}$$

with the notation

$$[W(x, t), V(x, t, s)] := W'(x, t)V(x, t, s) - V'(x, t, s)W(x, t) \tag{28}$$

for the Lie bracket.

Equation (27), in terms of

$$R(x, t) := \frac{\partial}{\partial t}W(x, t), \tag{29}$$

reads

$$\frac{\partial}{\partial s}V(x, t, s) = R(x, t) + \int_0^t (R'(x, \tau)V(x, t, s) - V'(x, t, s)R(x, \tau))d\tau$$

or equivalently

$$V_s = V_0 + sR + R \triangleright \left(\int_0^s V_\sigma d\sigma \right),$$

where we have used the notation $V_s(x, t) := V(x, t, s)$. Since $V(X, t, 0) = F(X, t)$, then

$$\begin{aligned} V(\cdot, \cdot, s) &= s R + \frac{s^2}{2} R \triangleright R + \frac{s^3}{3!} R \triangleright R \triangleright R + \dots \\ &+ F + s R \triangleright F + \frac{s^2}{2} R \triangleright R \triangleright F + \frac{s^3}{3!} R \triangleright R \triangleright R \triangleright F + \dots \end{aligned} \tag{30}$$

with the convention $F_1 \triangleright F_2 \triangleright \dots \triangleright F_m = F_1 \triangleright (F_2 \triangleright \dots \triangleright F_m)$.

We thus have the following result:

Proposition 2. A change of variables $x = \Psi_1(X, t)$ defined in terms of a continuous change of variables $X \mapsto z = \Psi_s(X, t)$ with generator

$$W(x, t) = \varepsilon W_1(x, t) + \varepsilon^2 W_2(x, t) + \dots \tag{31}$$

and $W(x,0) \equiv x$, transforms the system of equations (20) into (21), where f and F are related by

$$f = R + \frac{1}{2} R \triangleright R + \frac{1}{3!} R \triangleright R \triangleright R + \frac{1}{4!} R \triangleright R \triangleright R \triangleright R + \dots$$

$$+ F + R \triangleright F + \frac{1}{2} R \triangleright R \triangleright F + \frac{1}{3!} R \triangleright R \triangleright R \triangleright F + \dots \tag{32}$$

and R is given by (29).

Proposition 2 deals with changes of variables such that $X = \Psi_1(X, 0)$ (as a consequence of $W(X, 0) \equiv X$), so that the initial value problem obtained by supplementing (20) with the initial condition $x(0) = x_0$ is transformed into (21) supplemented with $X(0) = x_0$.

More generally, one may consider generators $W(\cdot, t)$ within some class \mathcal{C} of time-dependent smooth vector fields such that the operator $\partial_t : \mathcal{C} \rightarrow \mathcal{C}$ is invertible. Next result reduces to proposition 2, when one considers some class \mathcal{C} of generators $W(\cdot, t)$ such that $W(x,0) \equiv 0$, so that $\partial_t : \mathcal{C} \rightarrow \mathcal{C}$ is invertible, with inverse defined as $\partial_t^{-1}W(x, t) = \int_0^t W(x, \tau) d\tau$.

Proposition 3. A change of variables $x = \Psi_1(X, t)$ defined in terms of a continuous change of variables $X \mapsto z = \Psi_s(X, t)$ with generator

$$W(x, t) = \varepsilon W_1(x, t) + \varepsilon^2 W_2(x, t) + \dots \tag{33}$$

within some class \mathcal{C} of time dependent smooth vector fields with invertible $\partial_t : \mathcal{C} \rightarrow \mathcal{C}$ transforms the initial value problem

$$\frac{d}{dt}x = f(x, t), \quad x(0) = x_0 \tag{34}$$

into

$$\frac{d}{dt}X = F(X, t), \quad X(0) = \Psi_1^{-1}(x_0), \tag{35}$$

where f, F , and $R = \partial_t W$ are related by (32), and the binary operator $\triangleright : \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$ is defined as

$$P \triangleright Q = (\partial_t^{-1}P')Q - Q'(\partial_t^{-1}P) = [\partial_t^{-1}P, Q]. \tag{36}$$

Notice that the operation \triangleright of (36) satisfies the pre-Lie relation (15), and that this proposition applies, in particular, to the class \mathcal{C} of smooth (2π) -periodic vector fields in \mathbb{R}^d with vanishing average. In that case the operator ∂_t is invertible, with inverse given by

$$\partial_t^{-1}W(x, t) = \sum_{\substack{k \in \mathbb{Z} \\ k \neq 0}} \frac{1}{i k} e^{i k t} \hat{W}_k(x), \quad \text{if } W(x, t) = \sum_{\substack{k \in \mathbb{Z} \\ k \neq 0}} e^{i k t} \hat{W}_k(x).$$

4. Continuous transformations and the Magnus expansion

Consider now an initial value problem of the form

$$\frac{d}{dt}x = \varepsilon g(x, t), \quad x(0) = x_0, \tag{37}$$

where the parameter ε has been introduced for convenience. As stated in the introduction, the solution $x(t)$ of this problem (20) can be approximated at a given t as the solution $y(s)$ at $s = 1$ of the autonomous initial value problem

$$\frac{d}{ds}y = \varepsilon W_1(y, t) := \varepsilon \int_0^t g(z, \tau) d\tau, \quad y(0) = x_0.$$

This is nothing but the first term in the Magnus approximation of $x(t)$. As a matter of fact, the Magnus expansion is a formal series (31) such that, for each fixed value of t , formally $x(t) = y(1)$, where $y(s)$ is the solution of

$$\frac{d}{ds}y = W(y, t), \quad y(0) = x_0. \tag{38}$$

The Magnus expansion (31) can then be obtained by applying a change of variables $x = \Psi_1(X, t)$, defined in terms of a continuous transformation $X \mapsto z = \Psi_s(X, t)$ with generator $W = W(x, t)$ independent of s , that transforms (37) into

$$\frac{d}{dt}X = 0.$$

This can be achieved by applying proposition 2 with $F(X, t) \equiv 0$ and $f(x, t) = \varepsilon g(x, t)$, i.e., solving

$$\varepsilon g = R + \frac{1}{2} R \triangleright R + \frac{1}{3!} R \triangleright R \triangleright R + \frac{1}{4!} R \triangleright R \triangleright R \triangleright R + \dots \tag{39}$$

for

$$R(x, t) = \varepsilon R_1(x, t) + \varepsilon^2 R_2(x, t) + \varepsilon^3 R_3(x, t) + \dots$$

and determining the generator W as

$$W(x, t) = \int_0^t R(x, \tau) d\tau. \tag{40}$$

At this point it is worth analyzing how the usual Magnus expansion for linear systems developed in section 2 is reproduced with this formalism. To do that, we introduce operators $\Omega(t)$ and $B_s(t)$ such that

$$W(x, t) := \Omega(t)x, \quad V(x, t, s) := B_s(t)x, \quad \frac{\partial}{\partial t}W(x, t) := R(t)x.$$

Now equation (27) reads

$$\left(\frac{\partial}{\partial s}B_s\right)x - Rx = \Omega B_s x - B_s \Omega x \tag{41}$$

or equivalently

$$B_s = B_0 + sR + R \triangleright \left(\int_0^s B_\sigma d\sigma\right), \tag{42}$$

where the binary operation \triangleright defined in (26) reproduces (9). Since $B_1(t) = A(t)$ and $B_0 = 0$, then (39) is precisely (11). The continuous change of variables is then given by

$$X \mapsto z = \Psi_s(X, t) = \exp(s\Omega(t))X$$

so that

$$x(t) = \Psi_1(X, t) = e^{\Omega(t)}X(t) = e^{\Omega(t)}X(0) = e^{\Omega(t)}x(0)$$

reproduces the Magnus expansion in the linear case. In consequence, the expression for each term $W_j(x, t)$ in the Magnus series for the ODE (37) can be obtained from the corresponding formula for the linear case with the binary operation (26) and all results collected in section 2 are still valid in the general setting by replacing the commutator by the Lie bracket (28).

5. Continuous transformations and the Floquet–Magnus expansion

The procedure of section 3 can be extended when there is a periodic time dependence in the differential equation. In that case one gets a generalized Floquet–Magnus expansion with agrees with the usual one when the problem is linear.

As usual, the starting point is the initial value problem

$$\frac{d}{dt}x = \varepsilon g(x, t), \quad x(0) = x_0, \tag{43}$$

where now $g(x, t)$ depends periodically on t , with period T . As before, we apply a change of variables $x = \Psi_1(X, t)$, defined in terms of a continuous transformation $X \mapsto z = \Psi_s(X, t)$ with generator $W = W(x, t)$ that removes the time dependence, i.e., that transforms (43) into

$$\begin{aligned} \frac{d}{dt}X &= \varepsilon G(X; \varepsilon) = \varepsilon G_1(X) + \varepsilon^2 G_2(X) + \varepsilon^3 G_3(X) + \dots, \\ X(0) &= x_0. \end{aligned} \tag{44}$$

In addition, the generator W is chosen to be independent of s and periodic in t with the same period T .

This can be achieved by considering proposition 2 with $F(X, t) := \varepsilon G(X; \varepsilon)$ and $f(x, t) := \varepsilon g(x, t)$, solving (32) for the series

$$R(x, t) = \varepsilon R_1(x, t) + \varepsilon^2 R_2(x, t) + \varepsilon^3 R_3(x, t) + \dots$$

Thus, for the first terms one gets

$$\begin{aligned} R_1 &= g - G_1 \\ R_2 &= -\frac{1}{2}R_1 \triangleright R_1 - R_1 \triangleright G_1 - G_2 \\ R_3 &= -\frac{1}{2}(R_1 \triangleright R_2 + R_2 \triangleright R_1) + \frac{1}{3!}R_1 \triangleright R_1 \triangleright R_1 \\ &\quad - R_1 \triangleright G_2 - R_2 \triangleright G_1 - \frac{1}{2}R_1 \triangleright R_1 \triangleright G_1 - G_3 \end{aligned}$$

and, in general,

$$R_j = U_j - G_j, \quad j = 1, 2, \dots, \quad (45)$$

where U_j only contains terms involving g or the vector fields U_m and G_m of a lower order, i.e., with $m < j$. This allows one to solve (45) recursively by taking the average of U_j over one period T , i.e.,

$$G_j(X) = \langle U_j(X, \cdot) \rangle \equiv \frac{1}{T} \int_0^T U_j(X, t) dt,$$

thus ensuring that R_j is periodic. Finally, once G and R are determined, W is obtained from (40), which in turn determines the change of variables.

If we limit ourselves to the linear case, $g(x, t) = A(t)x$, with $A(t + T) = A(t)$, then, by introducing the operators

$$W(x, t) := \Lambda(t)x, \quad V(x, t, s) := B_s(t)x, \quad G(X) := FX,$$

the relevant equation is now

$$B_s = F + s\dot{\Lambda} + \dot{\Lambda} \triangleright \left(\int_0^s B_\sigma d\sigma \right), \quad (46)$$

which, with the additional constraint $B_1(t) = A(t)$, leads to

$$\begin{aligned} \dot{\Lambda}_1 &= A - F_1 \\ \dot{\Lambda}_2 &= -\frac{1}{2}\dot{\Lambda}_1 \triangleright \dot{\Lambda}_1 - \dot{\Lambda}_1 \triangleright F_1 - F_2 \end{aligned}$$

and so on, i.e., exactly the same expressions obtained in [10]. The transformation is now

$$x(t) = \Psi_1(X, t) = e^{\Lambda(t)}X(t) = e^{\Lambda(t)}e^{tF}x(0)$$

thus reproducing the Floquet–Magnus expansion in the periodic case [10].

Several remarks are in order at this point:

1. This procedure has close similarities with several averaging techniques. As a matter of fact, in the quasi-periodic case, it is equivalent to the high order quasi-stroboscopic averaging treatment carried out in [25].
2. Although both the Magnus and the Floquet–Magnus expansions are convergent in the linear case (1) for bounded A , the situation is far more involved in the general nonlinear case. In particular, the formal series expansions carried out in [25] to determine the averaged system corresponding to equation (43) and the change of variables do not converge in general. By the same token, convergence of the Floquet–Magnus expansion cannot be expected in the nonlinear case, although the same type of estimates for error bounds when the series are truncated can be presumably be derived.
3. A different style of high order averaging (that can be more convenient in some practical applications) can be performed by dropping the condition that $W(x, 0) \equiv x$, and requiring instead, for instance, that $W(x, t)$ has vanishing average in t . In that case, the initial condition in (44) must be replaced by $X(0) = \Psi_1^{-1}(x_0)$. The generator $W(x, t)$ of the change of variables and the averaged vector fields $\varepsilon G(x, t)$ can be similarly computed by considering proposition 3 with the class \mathcal{C} of smooth quasi-periodic vector fields (on \mathbb{R}^d or on some smooth manifold) with vanishing average.
4. The Floquet–Magnus expansion in the linear case provides by construction the structure furnished by the Floquet theorem and has found wide application in periodically driven quantum systems, when there is need not only of determining the effective time-independent Hamiltonian H_e , but also the fluctuations around the evolution due to H_e , whereas the standard Magnus expansion only provides H_e when it is determined at multiples of the period. In the nonlinear case the Floquet–Magnus expansion leads both to the averaged system and the required transformation.

6. Examples of application

The nonlinear Magnus expansion has been applied in the context of control theory, namely in non-holonomic motion planning. Several models in robotics, such as mobile platforms, free-floating robots, etc., can be described by the system

$$\dot{x}(t) = A(t)x = \sum_{i=1}^m g_i(x)u_i, \quad (47)$$

where $x = (x_1, \dots, x_n)^T$ is the configuration, $u_i, i = 1, \dots, m$ are the controls, $m < n$, and the g_i are smooth vector fields. A typical motion planning task is to find controls u_i that steer the system from a given initial configuration x_0 to a final one x_f . With these requirements, the (nonlinear) Magnus expansion allows one to express, locally around x_0 , the solution of (47) as $x(t) = \gamma(1)$, where $\gamma(s)$ is the solution of equation (38). In this way, one formally gets all admissible directions of motions in terms of control parameters, so that (a) one can easily select the motion in a desired direction and then (b) determine the control parameters in the optimal way that steer the non-holonomic system into that direction [21, 26, 27]. In this sense, expression (19) and the corresponding one for the generic, nonlinear case, can be very useful in applications, since it only contains independent terms [27, 28]. This problem is also closely connected with the calculation of the logarithm of the Chen–Fliess series [29].

As mentioned previously, the general Floquet–Magnus can also be applied in averaging. A large class of problems where averaging techniques are successfully applied is made of autonomous systems of the form

$$\dot{u} = Au + \varepsilon h(u), \quad u(0) = x_0, \quad (48)$$

where h is a smooth function from \mathbb{R}^n to itself (or more generally from a functional Banach space E space to itself) and A is a skew-symmetric matrix $\mathcal{M}(\mathbb{R}^n)$ (or more generally a linear operator on E) whose spectrum is included in $\frac{2i\pi}{T}\mathbb{Z}$. Denoting $x = e^{-tA}u$ and differentiating leads to

$$\dot{x} = -Ae^{-tA}u + e^{-tA}\dot{u} = \varepsilon e^{-tA}h(e^{tA}x)$$

so that x now satisfies an equation of the very form (43) with

$$g(x, t) = e^{-tA}h(e^{tA}x).$$

The T -periodicity of g with respect to time stems from the fact that $\text{Sp}(A) \subset \frac{2i\pi}{T}\mathbb{Z}$. For this specific case, relation (45) leads to the following expressions

$$G_1(X) = \langle g(X, \cdot) \rangle, \quad (49)$$

$$G_2(X) = -\frac{1}{2} \left\langle \int_0^t [g(X, \tau), g(X, t)] d\tau \right\rangle, \quad (50)$$

$$G_3(X) = \frac{1}{12} \left\langle \int_0^t \left[g(X, \tau), \int_0^\tau [g(X, \sigma), g(X, t)] d\sigma \right] d\tau \right\rangle \\ + \frac{1}{4} \left\langle \int_0^t \left[\int_0^\tau [g(X, \sigma), g(X, \tau)] d\sigma, g(X, t) \right] d\tau \right\rangle. \quad (51)$$

If $g(x, t)$ is a Hamiltonian vector field with Hamiltonian $H(x, t)$, then all G_i 's are Hamiltonian with Hamiltonian H_i 's. These Hamiltonians can be computed through the same formulas with Poisson brackets *in lieu* of Lie brackets (see e.g. [3]).

6.1. Dynamics of the Van der Pol system

As a first and elementary application of previous results, we consider the Van der Pol oscillator, which may be looked at as a perturbation of the simple harmonic oscillator:

$$\begin{cases} \dot{q} = p \\ \dot{p} = -q + \varepsilon(1 - q^2)p \end{cases} \quad (52)$$

Clearly, the system is of the form (48) with $u = (q, p)^T$,

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad h(u) = \begin{pmatrix} 0 \\ (1 - u_1^2)u_2 \end{pmatrix},$$

and is thus amenable to the rewriting (43), where

$$g(x, t) := e^{-tA}h(e^{tA}x) \quad \text{and} \quad e^{tA} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}.$$

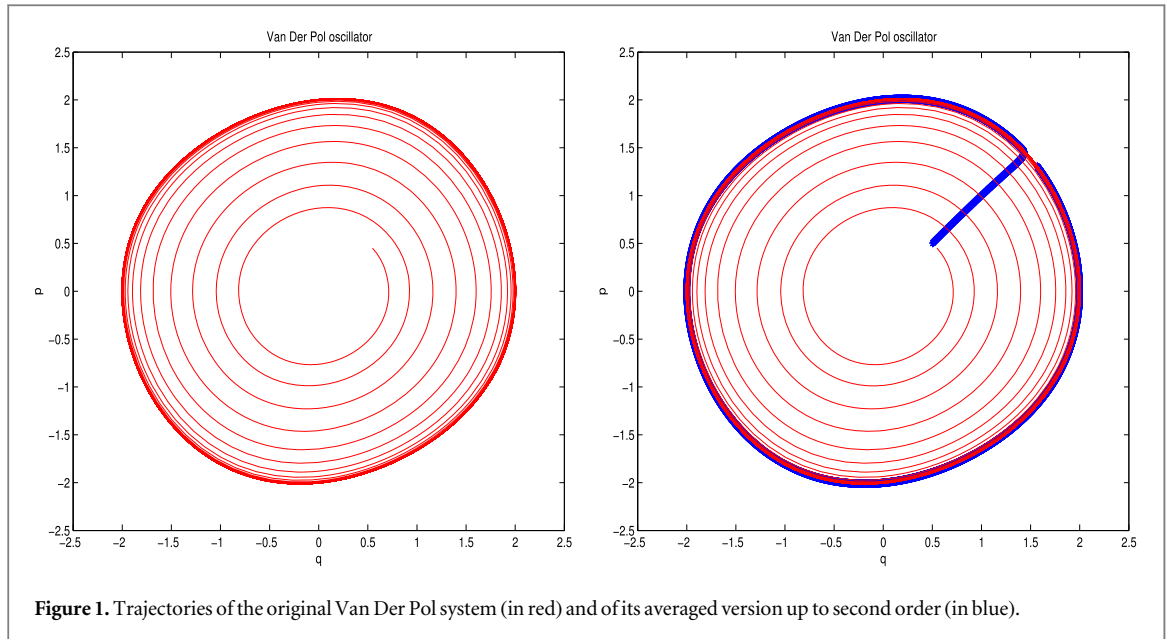


Figure 1. Trajectories of the original Van Der Pol system (in red) and of its averaged version up to second order (in blue).

In short, we have

$$\dot{x} = \varepsilon \xi_t(x) V_t,$$

where

$$V_t = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \quad \text{and} \quad \xi_t(x) = (1 - (\cos(t)x_1 + \sin(t)x_2)^2)(-\sin(t)x_1 + \cos(t)x_2).$$

Previous formulas give for the first term in the procedure

$$G_1(X) = \frac{1}{2\pi} \int_0^{2\pi} \xi_\tau(X) V_\tau d\tau = \begin{pmatrix} -\frac{1}{8}(\|X\|_2^2 - 4)X_1 \\ -\frac{1}{8}(\|X\|_2^2 - 4)X_2 \end{pmatrix} = -\frac{(\|X\|_2^2 - 4)}{8} X$$

and it is then easy to determine an approximate equation of the limit cycle, i.e. $\|X\|_2 = 2$. As for the dynamics of the first-order averaged system, it is essentially governed by the scalar differential equation on $N(X) := \|X\|_2^2$

$$\frac{d}{dt}N(X) = 2(X_1\dot{X}_1 + X_2\dot{X}_2) = -\varepsilon \frac{N(N - 4)}{4},$$

which has two equilibria, namely $\|X\|_2 = 0$ and $\|X\|_2 = 2$. The first one is unstable while the second one is stable. However, the graphical representation (see figure 1) of the solution of (52) soon convinces us that the true limit cycle is not a perfect circle. In order to determine a better approximation of this limit cycle, we thus compute the next term of the average equation from formula (50):

$$\begin{aligned} G_2(X) &= \frac{-1}{4\pi} \int_0^{2\pi} \int_0^t (\xi_t(\nabla_X \xi_\tau, V_t) V_\tau - \xi_\tau(\nabla_X \xi_t, V_\tau) V_t) d\tau \\ &= \begin{bmatrix} -\frac{1}{256} X_2(32 - 24 X_2^2 + 5 X_2^4 - 88 X_1^2 + 21 X_1^4 + 10 X_1^2 X_2^2) \\ \frac{1}{256} X_1(21 X_1^4 + 32 - 88 X_1^2 + 40 X_2^2 + 10 X_1^2 X_2^2 + 5 X_2^4) \end{bmatrix} \\ &= -\frac{\varepsilon^2}{256} D(X) J X, \end{aligned}$$

where

$$D(X) = \begin{pmatrix} 32 - 24 X_2^2 + 5 X_2^4 - 88 X_1^2 + 21 X_1^4 + 10 X_1^2 X_2^2 & 0 \\ 0 & D_{1,1}(X) + 64 X_1^2 \end{pmatrix}$$

and

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{53}$$

Now, considering the new quantity $L(X) = N(X) + \varepsilon Q(X)$ with

$$Q(X) = \nu X_1 X_2^3,$$

we see from the second-order averaged equation

$$\dot{X} = -\varepsilon \frac{(\|X\|_2^2 - 4)}{8} X - \frac{\varepsilon^2}{256} D(X) J X,$$

that

$$\begin{aligned} \frac{dL}{dt} &= \frac{dN}{dt} + \varepsilon (\nabla_X Q, \dot{X}) \\ &= -\varepsilon \frac{N(N-4)}{4} - \frac{\varepsilon^2}{128} (X, D(X) J X) - \varepsilon^2 \frac{N-4}{8} (\nabla_X Q, X) \\ &= -\varepsilon \frac{L(L-4)}{4} - \varepsilon^2 Q + \frac{\varepsilon^2}{2} Q L + \varepsilon^2 \frac{1}{2\nu} Q - \varepsilon^2 \frac{1}{2} (L-4) Q + \mathcal{O}(\varepsilon^3) \\ &= -\varepsilon \frac{L(L-4)}{4} + \mathcal{O}(\varepsilon^3) \end{aligned}$$

for $\nu = -\frac{1}{2}$. A more accurate description of the limit cycle is thus given by the equation

$$\|X\|_2^2 = 4 + \frac{\varepsilon}{2} X_1 X_2^3.$$

6.2. The first two terms of the averaged nonlinear Schrödinger equation (NLS) on the d -dimensional torus

Next, we apply the results of section 5 to the nonlinear Schrödinger equation (for a introduction to the NLS see for instance [30–32])

$$\begin{aligned} i\partial_t \psi &= -\Delta \psi + \varepsilon k(\psi \cdot \bar{\psi}) \psi, \quad t \geq 0, \quad z \in \mathbb{T}_a^d, \\ \psi(0, z) &= \psi_0(z) \in E, \end{aligned}$$

where $\mathbb{T}_a^d = [0, a]^d$ and $E = H^s(\mathbb{T}_a^d)$ is our working space. We hereafter conform to the hypotheses of [33] and assume that k is a real-analytic function and that $s > d/2$, ensuring that E is an algebra⁵. The operator $-\Delta$ is self-adjoint non-negative and its spectrum is

$$\sigma(A) = \left\{ \left(\frac{2\pi}{a} \right)^2 \sum_{j=1}^d l_j^2; l \in \mathbb{Z}^d \right\} \subset \left(\frac{2\pi}{a} \right)^2 \mathbb{N}, \tag{54}$$

so that by Stone's theorem, the group-operator $\exp(it\Delta)$ is periodic with period $T = \frac{a^2}{2\pi}$. We may thus rewrite Schrödinger equation as we did with equation (48). However, we shall instead decompose $\psi(t, z) = q(t, z) + ip(t, z)$ in its real and imaginary parts and derive the corresponding canonical system in the new unknown

$$u(t, \cdot) = \begin{pmatrix} p(t, \cdot) \\ q(t, \cdot) \end{pmatrix} \in H^s(\mathbb{T}_a^d) \times H^s(\mathbb{T}_a^d),$$

that is to say

$$\dot{u} = J^{-1} \text{Diag}(-\Delta, -\Delta) u + \varepsilon k(\|u\|_{\mathbb{R}^2}^2) J^{-1} u, \quad u(0) = \begin{pmatrix} p_0 \\ q_0 \end{pmatrix}, \tag{55}$$

where we have denoted $\dot{u} = \partial_t u$, $\|u\|_{\mathbb{R}^2}^2 = (u^1)^2 + (u^2)^2 = (u, u)_{\mathbb{R}^2}$ (u^1 and u^2 are the two components of u), J is given by (53) and

$$D = \begin{pmatrix} -\Delta & 0 \\ 0 & -\Delta \end{pmatrix}.$$

The operator D is self-adjoint on $L^2(\mathbb{T}_a^d) \times L^2(\mathbb{T}_a^d)$ and an obvious computation shows

$$e^{tJ^{-1}D} = \begin{pmatrix} \cos(t\Delta) & \sin(t\Delta) \\ -\sin(t\Delta) & \cos(t\Delta) \end{pmatrix} := R_t,$$

so that $e^{tJ^{-1}AD}$ is a group of isometries on $L^2(\mathbb{T}_a^d) \times L^2(\mathbb{T}_a^d)$ as well as on $H^s(\mathbb{T}_a^d) \times H^s(\mathbb{T}_a^d)$. Owing to (54) it is furthermore periodic (for all t , $R_{t+T} = R_t$), with period $T = \frac{a^2}{2\pi}$. The very same manipulation as for the prototypical system (48) then leads to

⁵ Under all these assumptions, for all initial value $\psi_0 \in E$ and all $\varepsilon > 0$, there exists a unique solution $\psi \in C([0, b/\varepsilon], E)$ for some $b > 0$ independent of ε [33].

$$\dot{x} = \varepsilon g(x, t), \quad x = \begin{pmatrix} p_0 \\ q_0 \end{pmatrix},$$

with

$$g(x, t) := J^{-1} e^{-tJ^{-1}D} k(\|e^{tJ^{-1}D} x\|_{\mathbb{R}^2}^2) e^{tJ^{-1}D} x = R_{t+T/4} k(\|R_t x\|_{\mathbb{R}^2}^2) R_t x.$$

Now, it can be verified that

$$g(x, t) = J^{-1} \nabla_x H(x, t),$$

where

$$H(x, t) := \frac{1}{2} \int_{\mathbb{T}_a^d} K(\|R_t x\|(t, z)\|_{\mathbb{R}^2}^2) dz \quad \text{with} \quad K(r) = \int_0^r k(\sigma) d\sigma. \tag{56}$$

Remark 1. Recall that the gradient is defined w.r.t. the scalar product (\cdot, \cdot) on $L^2(\mathbb{T}_a^d) \times L^2(\mathbb{T}_a^d)$ that we redefine for the convenience of the reader: for all pair of functions x_1 and x_2 in $L^2(\mathbb{T}_a^d) \times L^2(\mathbb{T}_a^d)$,

$$(x_1, x_2) = \int_{\mathbb{T}_a^d} (x_1^1(z)x_2^1(z) + x_1^2(z)x_2^2(z)) dz = \int_{\mathbb{T}_a^d} (x_1(z), x_2(z))_{\mathbb{R}^2} dz.$$

where x_1^1 and x_1^2 are the two components of x_1 and similarly for x_2 . Hence, by definition of the gradient, we have that

$$\forall (t, x_1, x_2) \in \mathbb{T} \times E^3, \quad (\nabla_x H(x_1, t), x_2) = \partial_x H(x_1, t) x_2.$$

Furthermore,

$$\forall (t, x_1, x_2) \in \mathbb{T} \times E^3, \quad (R_t x_1, x_2) = (x_1, R_{-t} x_2)$$

and

$$\forall (x_1, x_2, x_3) \in E^3 \quad (J \partial_x g(x_1, t) x_2, x_3) = (x_2, J \partial_x g(x_1, t) x_3).$$

Finally, if ϕ_1 and ϕ_2 are hamiltonian vector fields, with Hamiltonians Φ_1 and Φ_2 , then

$$[\phi_1, \phi_2] = \partial_x \phi_1 \phi_2 - \partial_x \phi_2 \phi_1 = J^{-1} \nabla_x \{\Phi_1, \Phi_2\}$$

where the Poisson bracket is defined by

$$\{\Phi_1, \Phi_2\} = (J \phi_1, \phi_2).$$

Now, the first term of the averaged vector field $G(X, \varepsilon)$ is simply

$$G_1(X) = \langle R_{+T/4} k(\|R_t X\|_{\mathbb{R}^2}^2) R_t X \rangle.$$

In order to obtain the second term, we use the simple fact that for any $\delta \in H^s(\mathbb{T}_a^d)$ the derivatives w.r.t. x in the direction δ may be computed as

$$\begin{aligned} \partial_x (k(\|R_t x\|_{\mathbb{R}^2}^2)) \cdot \delta &= k'(\|R_t x\|_{\mathbb{R}^2}^2) (\partial_x \|R_t x\|_{\mathbb{R}^2}^2) \cdot \delta \\ &= 2k'(\|R_t x\|_{\mathbb{R}^2}^2) (R_t x, R_t \delta)_{\mathbb{R}^2} \end{aligned}$$

so that

$$\begin{aligned} \partial_x (g(x, t)) \cdot \delta &= R_{t+T/4} k(\|R_t x\|_{\mathbb{R}^2}^2) R_t \delta \\ &\quad + 2R_{t+T/4} k'(\|R_t x\|_{\mathbb{R}^2}^2) (R_t x, R_t \delta)_{\mathbb{R}^2} R_t x. \end{aligned}$$

Inserted in the expression of G_2 we thus obtain the following expression for the ε^2 -term of the averaged equation

$$G_2(X) = -\frac{1}{2} \left\langle \int_0^t [g(X, \tau), g(X, t)] d\tau \right\rangle = I_1(X) + I_2(X)$$

with

$$\begin{aligned} I_1(X) &= -\frac{1}{2} \left\langle \int_0^t R_{\tau+T/4} k(\|R_\tau x\|_{\mathbb{R}^2}^2) R_{\tau+t+T/4} k(\|R_\tau x\|_{\mathbb{R}^2}^2) R_\tau x d\tau \right\rangle \\ &\quad + \frac{1}{2} \left\langle \int_0^t R_{t+T/4} k(\|R_t x\|_{\mathbb{R}^2}^2) R_{\tau+t+T/4} k(\|R_\tau x\|_{\mathbb{R}^2}^2) R_\tau x d\tau \right\rangle \end{aligned}$$

and

$$I_2(X) = - \left\langle \int_0^t R_{\tau+T/4} k'(\|R_\tau x\|_{\mathbb{R}^2}^2) (R_\tau x, R_{\tau+t+T/4} k(\|R_t x\|_{\mathbb{R}^2}^2) R_t x)_{\mathbb{R}^2} R_\tau x d\tau \right\rangle \\ + \left\langle \int_0^t R_{t+T/4} k'(\|R_t x\|_{\mathbb{R}^2}^2) (R_t x, R_{\tau+t+T/4} k(\|R_\tau x\|_{\mathbb{R}^2}^2) R_\tau x)_{\mathbb{R}^2} R_t x d\tau \right\rangle.$$

As already mentioned, both G_1 and G_2 are Hamiltonian with Hamiltonian H_1 and H_2 which could have been equivalently computed from $H(x, t)$ in (56) (see remark 1).

Acknowledgments

The work of FC and AM has been supported by Ministerio de Economía y Competitividad (Spain) through project MTM2016-77660-P (AEI/FEDER, UE). AM is also supported by the consolidated group IT1294-19 of the Basque Government. PC acknowledges funding by INRIA through its Sabbatical program and thanks the University of the Basque Country for its hospitality.

ORCID iDs

Fernando Casas  <https://orcid.org/0000-0002-6445-279X>

References

- [1] Magnus W 1954 On the exponential solution of differential equations for a linear operator *Comm. Pure and Appl. Math.* **VII** 649–73
- [2] Blanes S, Casas F, Oteo J and Ros J 2009 The Magnus expansion and some of its applications *Phys. Rep.* **470** 151–238
- [3] Blanes S and Casas F 2016 *A Concise Introduction to Geometric Numerical Integration* (Boca Raton: CRC Press)
- [4] Hairer E, Lubich C and Wanner G 2006 *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations* 2nd edn (Berlin: Springer)
- [5] Blanes S, Casas F, Oteo J and Ros J 1998 Magnus and Fer expansions for matrix differential equations: the convergence problem *J. Phys. A: Math. Gen.* **22** 259–68
- [6] Casas F 2007 Sufficient conditions for the convergence of the Magnus expansion *J. Phys. A: Math. Theor.* **40** 15001–17
- [7] Lakos G 2017 Convergence estimates for the Magnus expansion *Tech. Rep.* arXiv:1709.01791
- [8] Moan P and Niesen J 2008 Convergence of the Magnus series *Found. Comput. Math.* **8** 291–301
- [9] Coddington E and Levinson N 1955 *Theory of Ordinary Differential Equations* (New York: McGraw Hill)
- [10] Casas F, Oteo J A and Ros J 2001 Floquet theory: exponential perturbative treatment *J. Phys. A: Math. Gen.* **34** 3379–88
- [11] Kuwahara T, Mori T and Saito K 2016 Floquet-Magnus theory and generic transient dynamics in periodically driven many-body quantum systems *Ann. Phys.* **367** 96–124
- [12] Mananga E and Charpentier T 2016 On the floquet-magnus expansion: applications in solid-state nuclear magnetic resonance and physics *Phys. Rep.* **609** 1–50
- [13] Agrachev A and Gamkrelidze R 1979 The exponential representation of flows and the chronological calculus *Math. USSR-Sb.* **35** 727–85
- [14] Dewar R 1976 Renormalized canonical perturbation theory for stochastic propagators *J. Phys. A: Math. Gen.* **9** 2043–57
- [15] Iserles A, Munthe-Kaas H, Nørsett S and Zanna A 2000 Lie-group methods *Acta Numerica* **9** 215–365
- [16] Iserles A and Nørsett S 1999 On the solution of linear differential equations in Lie groups *Phil. Trans. Royal Soc. A* **357** 983–1019
- [17] Ebrahimi-Fard K and Manchon D 2009 A Magnus- and Fer-type formula in dendriform algebras *Found. Comp. Math.* **9** 295–316
- [18] Agrachev A and Gamkrelidze R 1994 The shuffle product and symmetric groups *Differential equations, Dynamical Systems, and Control Science* ed K Elworthy, W Everitt and E Lee (New York: Marcel Dekker) pp 365–82
- [19] Bialynicki-Birula I, Mielnik B and Plebański J 1969 Explicit solution of the continuous Baker-Campbell-Hausdorff problem and a new expression for the phase operator *Ann. Phys.* **51** 187–200
- [20] Mielnik B and Plebański J 1970 Combinatorial approach to Baker-Campbell-Hausdorff exponents *Ann. Inst. Henri Poincaré* **XII** 215–54
- [21] Strichartz RS 1987 The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations *J. Funct. Anal.* **72** 320–45
- [22] Dragt A and Forest E 1983 Computation of nonlinear behavior of Hamiltonian systems using Lie algebraic methods *J. Math. Phys.* **24** 2734–44
- [23] Arnal A, Casas F and Chiralt C 2018 A general formula for the Magnus expansion in terms of iterated integrals of right-nested commutators *J. Phys. Commun.* **2** 035024
- [24] Cary J 1981 Lie transform perturbation theory for Hamiltonian systems *Phys. Rep.* **79** 129–59
- [25] Chartier P, Murua A and Sanz-Serna J 2012 A formal series approach to averaging: exponentially small error estimates *Disc. Cont. Dyn. Syst.* **32** 3009–27
- [26] Duleba I 1997 Locally optimal motion planning of nonholonomic systems *Int. J. Robotic Systems* **14** 767–88
- [27] Duleba I and Kheifi W 2006 Pre-control form of the generalized Campbell-Baker-Hausdorff-Dynkin formula for affine nonholonomic systems *Syst. Control Lett.* **55** 146–57
- [28] Duleba I 1998 On a computationally simple form of the generalized campbell-baker-hausdorff-dynkin formula *Systems Control Letters* **34** 191–202
- [29] Kawski M 2002 The combinatorics of nonlinear controllability and noncommuting flows *Mathematical Control Theory, ICTP Lect. Notes, VIII, Abdus Salam Int. Cent. Theoret. Phys.* (Trieste) pp 223–311
- [30] Carles R 2008 *Semi-Classical Analysis for Nonlinear Schrödinger equations* (Singapore: World Scientific)
- [31] Cazenave T 2003 *Semilinear Schrödinger equations* (American Mathematical Society)

- [32] Cazenave T and Haraux A 1998 *An Introduction to Semilinear Evolution equations* (Oxford: Clarendon)
- [33] Castella F, Chartier P, Méhats F and Murua A 2015 Stroboscopic averaging for the nonlinear Schrödinger equation *Found. Comput. Math.* **15** 519–59