



HAL
open science

Stroboscopic Averaging for the Nonlinear Schrödinger Equation

François Castella, Philippe Chartier, Florian Méhats, Ander Murua

► **To cite this version:**

François Castella, Philippe Chartier, Florian Méhats, Ander Murua. Stroboscopic Averaging for the Nonlinear Schrödinger Equation. *Foundations of Computational Mathematics*, 2015, 15 (2), pp.519-559. 10.1007/s10208-014-9235-7. hal-00732850

HAL Id: hal-00732850

<https://hal.science/hal-00732850>

Submitted on 17 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stroboscopic averaging for the nonlinear Schrödinger equation

F. Castella*, Ph. Chartier†, F. Méhats‡ and A. Murua§

September 17, 2012

Abstract

In this paper, we are concerned with an averaging procedure, –namely *Stroboscopic averaging* [SVM07, CMSS10]–, for highly-oscillatory evolution equations posed in a (possibly infinite dimensional) Banach space, typically partial differential equations (PDEs) in a high-frequency regime where only one frequency is present. We construct a high-order averaged system whose solution remains *exponentially* close to the exact one over long time intervals, possesses the same geometric properties (structure, invariants, . . .) as compared to the original system, and is non-oscillatory. We then apply our results to the nonlinear Schrödinger equation on the d -dimensional torus \mathbb{T}^d , or in \mathbb{R}^d with a harmonic oscillator, for which we obtain a hierarchy of Hamiltonian averaged models. Our results are illustrated numerically on several examples borrowed from the recent literature.

Keywords: highly-oscillatory evolution equation, stroboscopic averaging, Hamiltonian PDEs, invariants, nonlinear Schrödinger.

MSC numbers: 34K33, 37L05, 35Q55.

1 Introduction

In this text we consider general, highly-oscillatory evolution equations, posed in a Banach space X , of the form

$$\frac{d}{dt}u^\varepsilon(t) = g_{t/\varepsilon}(u^\varepsilon(t)), \quad u^\varepsilon(0) = u_0 \in X, \quad t \in [0, T], \quad (\mathcal{P}^\varepsilon)$$

where $(\theta, u) \mapsto g_\theta(u)$ is P -periodic in θ , smooth in $\theta \in \mathbb{T}$ (\mathbb{T} denotes the torus $\mathbb{R}/(P\mathbb{Z})$), smooth in $u \in X$, and it is assumed that the above problem is well-posed on a *fixed* time interval $[0, T]$ independent of $\varepsilon \rightarrow 0$. This is a high-frequency system, with one frequency.

*IRMAR, Université de Rennes 1 and INRIA-Rennes, Campus de Beaulieu, 35042 Rennes Cedex, France. Email: Francois.Castella@univ-rennes1.fr

†INRIA-Rennes, IRMAR and ENS Bruz, Campus de Beaulieu, 35042 Rennes Cedex, France. Email: Philippe.Chartier@inria.fr

‡IRMAR, Université de Rennes 1 and INRIA-Rennes, Campus de Beaulieu, 35042 Rennes Cedex, France. Email: Florian.Mehats@univ-rennes1.fr

§Konputazio Zientziak eta A. A. Saila, Informatika Fakultatea, UPV/EHU, E-20018 Donostia-San Sebastián, Spain. Email: Ander.Murua@ehu.es

It is posed in an infinite-dimensional setting. The question we address is that of *high-order* averaging. In other words, we look for approximations of $u^\varepsilon(t)$, say $v^\varepsilon(t)$, such that $v^\varepsilon(t)$ is close to $u^\varepsilon(t)$ over $[0, T]$ (in some sense we detail below) to within $\mathcal{O}(\varepsilon^{n+1})$ where n is arbitrary large, *and* such that $v^\varepsilon(t)$ satisfies an *autonomous* evolution equation of the form $dv^\varepsilon(t)/dt = G^{[n]}(v^\varepsilon(t))$, say, where the fast variable t/ε has been averaged out. Note that for analytic convenience, the whole discussion below is actually developed for a rescaled time variable. We consider indeed the following rescaled version of $(\mathcal{P}^\varepsilon)$,

$$\frac{d}{dt}u^\varepsilon(t) = \varepsilon g_t(u^\varepsilon(t)), \quad u^\varepsilon(0) = u_0 \in X, \quad t \in [0, T/\varepsilon], \quad (\mathcal{Q}^\varepsilon)$$

where the unknown is still denoted u^ε . The problem is now assumed to be well-posed on the long time interval $[0, T/\varepsilon]$, the nonlinearity is rescaled by ε , and the question is to approximate u^ε over $[0, T/\varepsilon]$. The typical example we have in mind is that of nonlinear PDE's in a high-frequency regime, where only one frequency is present, and the averaged effect of the oscillations in the nonlinearities is to be computed at high order. We fully treat at the end of this text the case of the nonlinear Schrödinger equation on the torus \mathbb{T}^d , or on \mathbb{R}^d with a harmonic potential.

To be a bit more technical, the purpose of high-order averaging is, setting the error term $\mathcal{O}(\varepsilon^{n+1})$ to zero in order to fix the ideas, to find a *periodic, near-identity* and smooth change of variable Φ_θ^ε , together with a flow Ψ_t^ε , the flow map of an *autonomous* differential equation on X , such that the solution of the original equation $(\mathcal{Q}^\varepsilon)$ takes the *composed form*

$$u^\varepsilon(t) = \Phi_t^\varepsilon \circ \Psi_t^\varepsilon(u_0).$$

We refer e.g. to Lochak-Meunier [LM88] or Sanders-Verhulst [SVM07] for textbooks on these issues. The point is, such a form completely separates the dependence of $u^\varepsilon(t)$ upon the fast, periodic variable θ (through a given and smooth change of unknown), and the dependence of $u^\varepsilon(t)$ upon the slow variable t (through the flow of a given and autonomous vector field). The factorized form $u^\varepsilon(t) = \Phi_t^\varepsilon \circ \Psi_t^\varepsilon(u_0)$ is somehow analogous to the two-scale expansions (or WKB expansions as well – see [Wen26, Kra26, Bri26] on that point), where $u^\varepsilon(t)$ is sought in the two-scale form $u^\varepsilon(t) = U^\varepsilon(\theta, t)|_{\theta=t}$ as is more usual in the context of high-frequency PDE's. The present factorized form yet involves a deeper structure (beyond the mere separation of a fast periodic variable and a slow variable). This is one aspect we investigate. Besides, in the particular case of *stroboscopic averaging*, which is the framework we retain here, we impose that the near-identity mapping Φ_θ^ε satisfies $\Phi_{\theta=0}^\varepsilon = \text{Id}$ (the other, actually more standard, normalisation consists in imposing at variance a vanishing mean value, namely $\int_{\mathbb{T}} \Phi_\theta^\varepsilon d\theta = 0$). One interest of stroboscopic averaging relies, amongst others, in the fact that periodicity readily provides $\Phi_{\theta=nP}^\varepsilon = \text{Id}$, for all $n \in \mathbb{Z}$, i.e. the change of variables Φ_θ^ε reduces to identity at any multiple of the fast period, and we have $u^\varepsilon(t) = \Psi_t^\varepsilon(u_0)$ at these times (called stroboscopic times). The more crucial interest of stroboscopic averaging is, it allows to preserve the structure of the original problem along the averaging process, as already pointed out in Chartier et al. [CMSS10, CMSS]. This is a second aspect we investigate here.

Traditionally, the flow Ψ_t^ε and the change of variables Φ_θ^ε are sought by performing power expansions in ε , setting the Ansatz $\Phi_\theta^\varepsilon(u) = \text{Id} + \varepsilon \Phi_\theta^{(1)}(u) + \varepsilon^2 \Phi_\theta^{(2)}(u) + \dots$, and so on, and identifying like powers of ε , while separating the fast variable θ and the slow variable t

as well. This leads to a hierarchy of equations, so-called homological equations, that is to be solved iteratively. It allows to reconstruct in parallel the change of unknown Φ_θ^ε and the flow Ψ_t^ε . In the present one-frequency context, this procedure goes back to the work of Perko in the finite-dimensional case [Per69], who gives general analytic formulae at any order, as well as an elegant iterative procedure.

One key aspect of the present work is, we *do not* expand the various unknowns in ε (nor do we separate θ and t at once). We actually identify a new, *closed* equation on the change of variables Φ_t^ε . We show how to solve this equation using a fixed point procedure in an analytic framework, taking advantage of the fact that the – to be inverted – nonlinear relation is an ε -perturbation of the identity. This equation, called *transport equation* in the sequel (see (2.8)) is a nonlinear transport equation, with nonlocal nonlinearities. It is fully nonlinear in that the transport term involves derivatives of the unknown Φ_t^ε with respect to $u \in X$. We circumvent this latter difficulty by requiring analytic regularity, which allows to absorb the higher order derivatives. Eventually, we show that this boils down to requiring g (in (Q^ε)) is analytic with respect to u , an assumption that turns out to be naturally satisfied in the PDE context associated with the present considerations. We show in passing that Φ_t^ε is unique in some sense, so that, for instance, the n -th order Taylor expansion of Φ_t^ε in ε necessarily coincides with the formulae provided by any formula given by any other perturbative method, provided Φ_0^ε is the identity map (stroboscopic averaging). Once the change of variables Φ_t^ε is obtained, we show that the flow Ψ_t^ε is easily reconstructed. The final result we establish in that direction is, general high-order averaging may be achieved, with error terms of size $\mathcal{O}(e^{-c/\varepsilon})$ for some $c > 0$. Note that the transport equation, whose importance is pointed out here, plays a role that is somehow similar to the so-called singular equation of geometric optics, in the context of high-frequency hyperbolic PDE's (see [JMR95]).

It is furthermore interesting to note that we require only continuity of the mapping $(\theta, u) \mapsto g_\theta(u)$ with respect to θ , in complete agreement with previous works of Neishtadt [Nei84] or Chartier et al. [CMSS10] in the finite-dimensional context. In the presence of several non-resonant frequencies, analyticity with respect to θ becomes a necessary assumption (see for instance [Sim94, CMSS12] for the finite-dimensional case) and the extension of the present work to this situation would probably raise no serious difficulty, under a similar assumption of analyticity on θ . This comes in sharp contrast with the Magnus expansion that we outline in Section 5 with the specific aim of deriving explicit formulae for the high-order averaged vector fields: an averaging technique based on Magnus series is by essence strictly restricted to the one-frequency situation.

A second key aspect of the present work is, we prove that the so-obtained stroboscopic averaging preserves the structure of the original equation. For instance we show that the autonomous flow Ψ_t^ε is symplectic whenever the original oscillatory problem has a Hamiltonian structure. In the similar spirit, we show that the autonomous flow Ψ_t^ε possesses an invariant as soon as the original oscillatory problem does.

Finally, we apply our results to general nonlinear Schrödinger equations (NLS) in a high-frequency regime. We show that, up to a simple and standard filtering procedure, the obtained results apply directly. They provide a hierarchy of models that approximate the original equations to within arbitrary order, and whose precise analytic expression may be explicitly computed. These models essentially possess the same structure (invariants, Hamiltonian structure) than the original one.

Averaging as developed in this paper constitutes in our view an alternative to other existing techniques, amongst which stand most prominently Birkhoff’s forms and more recently Modulated Fourier Expansions. Birkhoff’s forms technique has been elaborated in the context of infinite dimensional systems by Bambusi [Bam03, Bam05, BG06, Bam08], Bourgain [Bou96, Bou07], Colliander [CKS⁺10] and Grébert [GVB11, GT12], to mention just a few authors. It is based on successive non-linear changes of variables which gradually transform the system into its final *normal form*. Modulated Fourier expansions have been introduced in the finite dimensional context by Hairer and Lubich [HL00], and later on extended to some Hamiltonian PDEs (long-time behaviour of solutions to the semi-linear Klein-Gordon equation [CHL08b], cubic Schrödinger equations with small initial data (or small nonlinearity) [GL10]). The approach has the advantage of a possible application to numerical discretizations of the original equation (see for instance [CHL08a]), in contrast with the approach via nonlinear coordinate transformations bringing the equation to a normal form. Note that unlike the situation we consider, most of the papers quoted above deal with the case of non-resonant frequencies. Finally, we wish to stress the fact that efficient numerical methods, which use only the original equations, can be built upon the existence of an averaged vector field (see [CCMSS11, CMSS10] for the ODE case): this will be the subject of another paper by the same authors.

Our first main theorem (Theorem 2.7), an averaging result in an abstract Banach space, is stated in Subsection 2.3, once preliminary assumptions have been settled in Subsection 2.1 and ideas sustaining stroboscopic averaging outlined in Subsection 2.2. The remaining of Section 2 is devoted to technical proofs and intermediate theorems, with the exception of Subsection 2.5 which is concerned with the linear case, where expansions are shown to converge (in contrast with the general non-linear case). Geometric aspects are dealt with in Section 3: invariants in Subsection 3.2 (see Theorem 3.7) and the more involved situation of Hamiltonian equations in the context of Hilbert spaces in Subsection 3.1 (see Theorem 3.5). Section 4 considers the instantiation for NLS of Theorems 2.7, 3.5 and 3.7. In the NLS context, Theorem 4.2 can be considered as the main input of this article. Finally, Section 5 is devoted to the explicit computation of the averaged vector fields with aid of *formal* Magnus expansions and Section 6 presents numerical illustrations of the behavior of some specific solutions to NLS described in the literature and interpreted “in the light of stroboscopic averaging”.

2 High-order averaging in a Banach space

Let X a *real* Banach space, equipped with the norm $\|\cdot\|_X$. We consider the following highly-oscillatory evolution equation, posed in X , namely

$$\begin{aligned} \frac{d}{dt}u^\varepsilon(t) &= \varepsilon g_t(u^\varepsilon(t)), & u(t) &\in X, \\ u^\varepsilon(0) &= u_0, & u_0 &\in X, \end{aligned} \tag{2.1}$$

where the initial datum u_0 is given. Here the function

$$(\theta, u) \in \mathbb{T} \times X \mapsto g_\theta(u) \in X$$

is given, smooth with respect to $u \in X$, smooth and periodic¹ in $\theta \in \mathbb{T}$. We actually require $u \mapsto g_\theta(u)$ to be real analytic in a sense we define later. We refer to Assumption 2.3 below for the precise hypotheses.

Our aim is to compute high-order (in ε) approximations of the solution $u^\varepsilon(t)$ to (2.1), on time intervals of size $\mathcal{O}(1/\varepsilon)$. We therefore readily introduce the following basic assumption.

Assumption 2.1 *The Cauchy problem (2.1) is uniformly well-posed in the following sense. There exist*

$$T > 0, \varepsilon^* > 0, \text{ and a bounded open subset } K \subset X,$$

such that, for all $\varepsilon \in]0, \varepsilon^]$, the problem (2.1) admits a unique solution $u^\varepsilon \in C^1([0, T/\varepsilon], X)$, $u^\varepsilon(t)$ remaining in K for all $t \leq T/\varepsilon$.*

The above assumption freezes the meaning of the quantities T , ε^* and K throughout the remaining part of Section 2.

2.1 Preliminaries

The analysis we present in this paper heavily relies on the use of analytic functions: this allows, ultimately, to derive optimal error estimates with remainder terms of the form $\mathcal{O}(\exp(-c/\varepsilon))$ (for some $c > 0$). It also allows to point out the intrinsic structures involved in the analysis.

For that reason, given the *real* Banach space X , we introduce the complexification of X , defined as

$$X_{\mathbb{C}} = \{U := u + i\tilde{u}, (u, \tilde{u}) \in X^2\}.$$

We denote $u = \Re(U) \in X$ and $\tilde{u} = \Im(U) \in X$ the real and imaginary parts of U . The space $X_{\mathbb{C}}$ is a Banach space when endowed with the norm

$$\|U\|_{X_{\mathbb{C}}} := \sup_{\lambda \in \mathbb{C}^*} \frac{\|\Re(\lambda U)\|_X}{|\lambda|}.$$

If X is a Hilbert space, it is more convenient to equip $X_{\mathbb{C}}$ with the Hermitian norm associated to the complex scalar product

$$(u + i\tilde{u}, v + i\tilde{v})_{X_{\mathbb{C}}} = (u, v)_X + (\tilde{u}, \tilde{v})_X + i((\tilde{u}, v)_X - (u, \tilde{v})_X).$$

Note that for $u \in X$, we have $\|u\|_{X_{\mathbb{C}}} = \|u\|_X$.

Now, given any $\rho > 0$, we consider the open enlargement of K in $X_{\mathbb{C}}$ given by

$$K_\rho = \{u + \tilde{u} : (u, \tilde{u}) \in K \times X_{\mathbb{C}}, \|\tilde{u}\|_{X_{\mathbb{C}}} < \rho\}.$$

We define analytic functions on K_ρ as follows.

Definition 2.2 (Analytic functions on a Banach space – see [PT87])

Consider a continuous function $(\theta, u) \in \mathbb{T} \times K_\rho \mapsto f_\theta(u) \in X_{\mathbb{C}}$. The map f_θ is said to

¹Here and below, "periodic" refers to " P -periodic", and the normalisation that we retain is such that \mathbb{T} is the torus $\mathbb{R}/(P\mathbb{Z})$

be analytic on $K_\rho \subset X_{\mathbb{C}}$ whenever it is continuously differentiable on K_ρ , i.e. there exists a continuous map

$$\begin{aligned} \mathbb{T} \times K_\rho &\rightarrow \mathcal{L}(X_{\mathbb{C}}) \\ (\theta, u) &\mapsto (\partial_u f_\theta)(u) \end{aligned}$$

where $\mathcal{L}(X_{\mathbb{C}})$ is the set of bounded linear maps from $X_{\mathbb{C}}$ to $X_{\mathbb{C}}$, which satisfies

$$\begin{aligned} \forall u \in K_\rho, \exists \delta > 0, \forall h \in X_{\mathbb{C}} \text{ with } \|h\|_{X_{\mathbb{C}}} \leq \delta, \\ \sup_{\theta \in \mathbb{T}} \|f_\theta(u+h) - f_\theta(u) - (\partial_u f_\theta)(u)h\|_{X_{\mathbb{C}}} = o(\|h\|). \end{aligned}$$

When $(\theta, u) \mapsto f_\theta(u)$ is a bounded analytic function on $\mathbb{T} \times K_\rho$, we denote

$$\|f\|_\rho = \sup_{(\theta, u) \in \mathbb{T} \times K_\rho} \|f_\theta(u)\|_{X_{\mathbb{C}}}.$$

Analytic functions in the sense of Definition 2.2 share many of the properties of standard analytic functions from \mathbb{C} to \mathbb{C} . For instance, if f_θ is analytic on K_ρ , Cauchy's formula holds. Indeed, whenever $0 < \delta < \rho$, for any u and \tilde{u} such that $u \in K_{\rho-\delta}$ and $\tilde{u} \in X_{\mathbb{C}}$ with $\|\tilde{u}\|_{X_{\mathbb{C}}} = 1$, for any $\mu \in \mathbb{C}$ with $|\mu| < \delta$, one has

$$f_\theta(u + \mu\tilde{u}) = \frac{1}{2i\pi} \int_{|\xi|=\delta} \frac{f_\theta(u + \xi\tilde{u})}{\xi - \mu} d\xi.$$

This allows to use Cauchy's estimates in a neighborhood of any point $u \in K_\rho$. For instance, given $0 < \delta < \rho$, one can estimate the norm of the first derivative $\partial_u f_\theta$ as

$$\|\partial_u f\|_{\rho-\delta} \leq \frac{1}{\delta} \|f\|_\rho. \quad (2.2)$$

This comes from the relation $(\partial_u f_\theta)(u)\tilde{u} = \left. \frac{d}{d\mu} f_\theta(u + \mu\tilde{u}) \right|_{\mu=0} = \frac{1}{2i\pi} \int_{|\xi|=\delta} \frac{f_\theta(u + \xi\tilde{u})}{\xi^2} d\xi$, which implies, in passing, that $\partial_u f_\theta$ is also analytic in the sense of Definition 2.2, as a function from K_ρ to $\mathcal{L}(X_{\mathbb{C}})$.

We are now ready to state the assumptions on g_θ in (2.1) required by our analysis.

Assumption 2.3 *The function $(\theta, u) \mapsto g_\theta(u)$ is C^0 and periodic in θ . Besides, $(\theta, u) \mapsto g_\theta(u)$ is real-analytic in u , in the following sense. There exist*

$$R > 0, C_K > 0,$$

such that for all $\theta \in \mathbb{T}$, $u \mapsto g_\theta(u)$ is analytic on K_{2R} in the sense of Definition 2.2, while $(\theta, u) \mapsto g_\theta(u)$ is bounded by C_K on $\mathbb{T} \times K_{2R}$, i.e.

$$\|g\|_{2R} = \sup_{(\theta, u) \in \mathbb{T} \times K_{2R}} \|g_\theta(u)\|_{X_{\mathbb{C}}} \leq C_K.$$

The above assumption defines once for all the quantities R and C_K throughout the remaining part of Section 2.

2.2 The equations of stroboscopic averaging

As explained in the introduction, ideally, the purpose of high-order averaging is to find a periodic, near-identity, and smooth change of variable Φ_θ^ε ,

$$(\theta, u) \in \mathbb{T} \times K \mapsto \Phi_\theta^\varepsilon(u) \in X$$

together with a flow Ψ_t^ε , the flow map of an autonomous differential equation with smooth vector field G^ε on X , written as

$$\frac{d}{dt} \Psi_t^\varepsilon(u_0) = \varepsilon G^\varepsilon(\Psi_t^\varepsilon(u_0)), \quad (2.3)$$

and such that the solution of the original equation (2.1) takes the *composed form*²

$$u^\varepsilon(t) = \Phi_t^\varepsilon \circ \Psi_t^\varepsilon(u_0). \quad (2.4)$$

Besides, in the framework of stroboscopic averaging that we retain here, we impose that the near-identity mapping Φ_θ satisfies

$$\Phi_{\theta=0}^\varepsilon = \text{Id}.$$

Let us now formally seek the equations satisfied by Φ_θ^ε and Ψ_t^ε . By differentiating both sides of (2.4) w.r.t. t and using (2.3) we readily get

$$\frac{\partial \Phi_t^\varepsilon}{\partial t}(\Psi_t^\varepsilon(u_0)) + \varepsilon \frac{\partial \Phi_t^\varepsilon}{\partial u}(\Psi_t^\varepsilon(u_0)) G^\varepsilon(\Psi_t^\varepsilon(u_0)) = \varepsilon g_t(\Phi_t^\varepsilon \circ \Psi_t^\varepsilon(u_0)). \quad (2.5)$$

Considering (2.5) for $u_0 = \Psi_{-t}(u)$ and replacing t by $\theta \in \mathbb{T}$ (as g_t and Φ_t^ε are periodic with respect to t) we obtain

$$\frac{\partial \Phi_\theta^\varepsilon}{\partial \theta}(u) + \varepsilon \frac{\partial \Phi_\theta^\varepsilon}{\partial u}(u) G^\varepsilon(u) = \varepsilon g_\theta(\Phi_\theta^\varepsilon(u)). \quad (2.6)$$

Let us formally solve the equation (2.6). To that aim, we introduce the (standard) notation for averages with respect to the variable θ , namely, for $(\theta, u) \in \mathbb{T} \times X \mapsto f_\theta(u) \in X$, the average $\langle f \rangle : X \rightarrow X$ denotes the mapping

$$\langle f \rangle(u) := \frac{1}{P} \int_{\mathbb{T}} f_\theta(u) d\theta.$$

Firstly, and with this notation at hand, taking averages in θ on both sides of (2.6) eliminates the first term $\partial_\theta \Phi_\theta^\varepsilon$ owing to periodicity, so that we obtain

$$\frac{\partial \langle \Phi^\varepsilon \rangle}{\partial u}(u) G^\varepsilon(u) = \langle g \circ \Phi^\varepsilon \rangle(u).$$

Assuming for the time being that the linear operator $v \mapsto \frac{\partial \langle \Phi^\varepsilon \rangle}{\partial u}(u) v$ is invertible for any u , we get

$$G^\varepsilon(u) := \left(\frac{\partial \langle \Phi^\varepsilon \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \Phi^\varepsilon \rangle(u). \quad (2.7)$$

²*Stricto sensu*, we have frozen the initial datum u_0 up to now, and relation (2.4) only holds for this value of u_0 . Needless to say, averaging aims at establishing such a factorization whenever u_0 belongs to some open set.

In other terms, we have here derived the value of the vector field G^ε (hence that of Ψ_t^ε).

Secondly, inserting previous relation in equation (2.6), we are led to the relation

$$\frac{\partial \Phi_\theta^\varepsilon}{\partial \theta}(u) + \varepsilon \frac{\partial \Phi_\theta^\varepsilon}{\partial u}(u) \left(\frac{\partial \langle \Phi^\varepsilon \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \Phi^\varepsilon \rangle(u) = \varepsilon g_\theta \circ \Phi_\theta^\varepsilon(u), \quad (2.8)$$

i.e., in an integral version,

$$\Phi_\theta^\varepsilon(u) = u + \varepsilon \int_0^\theta \left(g_\xi \circ \Phi_\xi^\varepsilon(u) - \frac{\partial \Phi_\xi^\varepsilon}{\partial u}(u) \left(\frac{\partial \langle \Phi^\varepsilon \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \Phi^\varepsilon \rangle(u) \right) d\xi. \quad (2.9)$$

This is a closed equation on Φ_θ^ε , though nonlinear and nonlocal.

In our perspective, there remains to solve equation (2.9) for Φ_θ^ε . Once this is done, the autonomous vector field G^ε and the associated flow Ψ_t^ε are immediately deduced along formula (2.7), and the averaging procedure is complete.

We now need to turn the above formal computations into a rigorous analytic procedure. In particular, the question arises whether the main equation (2.9) possesses a solution in the class of periodic smooth functions. It is known, see e.g. counter-examples constructed in [Nei84, CMSS12], that in general, the answer is negative: equation (2.8) can only be solved *up to an error term* of size $\mathcal{O}(\exp(-c/\varepsilon))$ for some $c > 0$. Accordingly, our aim in the sequel is to exhibit such a (quasi-) solution, and to prove that it provides a rigorous averaging procedure to within $\mathcal{O}(\exp(-c/\varepsilon))$ small terms.

2.3 Main result: averaging to within exponentially small remainder terms

It is useful to introduce the following nonlinear operator. Given any periodic and smooth mapping $(\theta, u) \in \mathbb{T} \times K_\rho \mapsto \varphi_\theta(u)$ with invertible partial derivative $\partial_u \langle \varphi \rangle$, we associate the mapping $(\theta, u) \in \mathbb{T} \times K_\rho \mapsto \Lambda(\varphi)_\theta(u)$ defined as

$$\Lambda(\varphi)_\theta(u) = g_\theta \circ \varphi_\theta(u) - \frac{\partial \varphi_\theta}{\partial u}(u) \left(\frac{\partial \langle \varphi \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \varphi \rangle(u). \quad (2.10)$$

We also introduce the nonlinear, nonlocal operator Γ^ε which associates to φ_θ the mapping $(\theta, u) \in \mathbb{T} \times K_\rho \mapsto \Gamma^\varepsilon(\varphi)_\theta(u)$ defined as

$$\Gamma^\varepsilon(\varphi)_\theta(u) = u + \varepsilon \int_0^\theta \Lambda(\varphi)_\xi(u) d\xi. \quad (2.11)$$

Remark 2.4 *Note that if φ_θ is periodic, then $\Lambda(\varphi)_\theta$ has zero average, hence $\Gamma^\varepsilon(\varphi)_\theta$ is periodic as well, and our definitions make sense.*

With this notation at hand, the basic equation of averaging (2.9) reads, in its integral form,

$$\Phi_\theta^\varepsilon(u) = \Gamma^\varepsilon(\Phi^\varepsilon)_\theta(u), \quad (2.12)$$

a fixed point equation that we solve below by considering iterates $(\Gamma^\varepsilon)^k$, $k = 0, 1, \dots, n$. Equivalently, the equation of averaging (2.8) reads, in its differential form,

$$\partial_\theta \Phi_\theta^\varepsilon(u) = \varepsilon \Lambda(\Phi^\varepsilon)_\theta(u), \quad (2.13)$$

with $\Phi_{\theta=0}^\varepsilon(u) = u$. At this point, we thus consider the sequence of functions

$$\Phi_\theta^{[0]} = \text{Id}, \quad \Phi_\theta^{[k+1]} = \Gamma^\varepsilon(\Phi^{[k]})_\theta, \quad k = 0, 1, 2, \dots, n, \quad (2.14)$$

together with the vector fields

$$G^{[k]}(u) := \left(\frac{\partial \langle \Phi^{[k]} \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \Phi^{[k]} \rangle(u). \quad (2.15)$$

In addition, we introduce the following terms occurring in the expansion of $G^{[k]}$, for $k \geq 0$, namely

$$G_{k+1}(u) = \frac{1}{k!} \left. \frac{d^k G^{[k]}}{d\varepsilon^k} \right|_{\varepsilon=0}(u). \quad (2.16)$$

Remark 2.5 *By construction (this will be justified rigorously in Theorem 2.13), for all $n, k \geq 0$,*

$$\frac{1}{k!} \left. \frac{d^k G^{[n+k]}}{d\varepsilon^k} \right|_{\varepsilon=0}(u) = G_{k+1}(u).$$

For the convenience of the reader, we now sum up the notations used later on in Section 2:

- $\varepsilon_0 := \frac{R}{8C_K P}$, $\varepsilon_1 := \varepsilon_0/2$, $\varepsilon_2 = \varepsilon_0 \min\left(\frac{1}{28}, \frac{\varepsilon_0 P}{T}\right)$,
- $C_0 = 16C_K$, $C_1 = 2C_K$, $\eta = 2/\varepsilon_0$,
- $r_n := \frac{R}{n+1}$, $R_k = 2R - kr_n$ for $k = 1, \dots, n-1$, so that $R_0 = 2R$ and $R_{n+1} = R$.

Lemma 2.6 *Given $n \in \mathbb{N}$, the maps $\Phi_\theta^{[k]}(u)$ and $G_{k+1}(u)$ for $0 \leq k \leq n+1$ are well-defined for any $\varepsilon \in \mathbb{C}$ with $|\varepsilon| \leq \varepsilon_0/(n+1)$, they are C^1 in $\theta \in \mathbb{T}$, analytic in u for $u \in K_{R_k}$, and analytic in ε for $|\varepsilon| < \varepsilon_0/(n+1)$. Moreover, the following estimate holds true for any $\varepsilon \in \mathbb{C}$ such that $(n+1)|\varepsilon| < \varepsilon_0$, namely*

$$\|\Phi^{[k]} - \text{Id}\|_{R_k} \leq \frac{r_n}{2}. \quad (2.17)$$

From now on, let us fix μ such that $0 < \mu < 1$. Then, for $0 < \varepsilon < \mu\varepsilon_2$, define the integer $n_\varepsilon \in \mathbb{N}^*$ as $(n_\varepsilon + 1) = \lfloor \mu/(\eta\varepsilon) \rfloor$ and denote

$$\tilde{G}^\varepsilon(u) = \tilde{G}^{[n_\varepsilon]}(u), \quad \text{with} \quad \tilde{G}^{[n]}(u) = \sum_{k=0}^n \varepsilon^k G_{k+1}(u), \quad (2.18)$$

where the vector fields $G_{k+1}(u)$ are defined in (2.16).

Theorem 2.7 *For $0 < \varepsilon < \min(\varepsilon^*, \mu\varepsilon_2)$, let $(n_\varepsilon + 1) = \lfloor \mu/(\eta\varepsilon) \rfloor$. Consider $\tilde{\Phi}_\theta^\varepsilon = \Phi_\theta^{[n_\varepsilon]}$ defined in Lemma 2.6 and \tilde{G}^ε defined by (2.18). Then the following holds:*

(i) *The solution $u_\varepsilon(t)$ of (2.1) satisfies*

$$\forall t \in [0, T/\varepsilon], \quad u_\varepsilon(t) = \tilde{\Phi}_t^\varepsilon(U(t)). \quad (2.19)$$

where $U(t)$ is the solution of the (quasi-autonomous) equation

$$\frac{dU}{dt} = \varepsilon \tilde{G}^\varepsilon(U) + \varepsilon \mathcal{R}_t^\varepsilon(U), \quad (2.20)$$

where $(\theta, u) \mapsto \mathcal{R}_\theta^\varepsilon(u)$ is periodic and smooth in θ , analytic on K_R , and admits the upper bound

$$\|\mathcal{R}^\varepsilon\|_R \leq \frac{4C_K}{(1-\mu)\mu^2} \exp\left(-\frac{\mu|\log(\mu)|}{\eta\varepsilon}\right). \quad (2.21)$$

(ii) Introduce $\tilde{\Psi}_t^\varepsilon$ the t -flow of the autonomous differential equation

$$\frac{dU}{dt} = \varepsilon \tilde{G}^\varepsilon(U). \quad (2.22)$$

If in addition, $\mu \leq 1/2$, then the solution $u_\varepsilon(t)$ to the original initial value problem (2.1) is exponentially close to the factorized form $\tilde{\Phi}_t^\varepsilon \circ \tilde{\Psi}_t^\varepsilon(u_0)$, i.e. we have

$$\forall t \in [0, T/\varepsilon], \quad \left\| u^\varepsilon(t) - \tilde{\Phi}_t^\varepsilon \circ \tilde{\Psi}_t^\varepsilon(u_0) \right\|_{X_{\mathbb{C}}} \leq 12 R \mu^{-2} \exp\left(-\frac{\mu|\log(\mu)|\varepsilon_0}{4\varepsilon}\right). \quad (2.23)$$

(iii) If $T = +\infty$ in Assumption 2.1 and $\mu \leq 1/2$, then (2.19) holds for all $t > 0$, and (2.23) holds on $[0, \tilde{T}/\varepsilon^{1+\alpha}]$, for any $\tilde{T} > 0$ and any $0 < \alpha < 1$, provided

$$0 < \varepsilon < \min\left(\varepsilon^*, \frac{\mu\varepsilon_0}{28}, \left(\frac{\mu\varepsilon_0^2 P}{\tilde{T}}\right)^{\frac{1}{1-\alpha}}\right).$$

2.4 Technical proofs and intermediate Theorems

In this Subsection, we give the technical details of the proof of Theorem 2.7 and state a few intermediate results which are of interest by their own in specific contexts (fixed-order truncation, non-expanded version of the vector field).

2.4.1 Two basic lemmas

To begin with, the next lemma gives sufficient conditions for the quantities $\Lambda(\varphi)$ and $\Gamma^\varepsilon(\varphi)$ to be well-defined in a clean analytic fashion.

Lemma 2.8 *Let $0 < \delta < \rho \leq 2R$. Assume that the function $(\theta, u) \in \mathbb{T} \times K_\rho \mapsto \varphi_\theta(u) \in X_{\mathbb{C}}$ is analytic in the sense of Definition 2.2, and that φ_θ is a near-identity mapping, in that*

$$\|\varphi - \text{Id}\|_\rho \leq \frac{\delta}{2}.$$

Then, the following holds.

(i) *The mapping $\partial_u \langle \varphi \rangle^{-1}$ is well-defined and analytic on $K_{\rho-\delta}$, and it satisfies*

$$\|\partial_u \langle \varphi \rangle^{-1}\|_{\rho-\delta} \leq 2.$$

(ii) *The mappings $(\theta, u) \in \mathbb{T} \times K_{\rho-\delta} \mapsto \Lambda(\varphi)_\theta(u)$ and $(\theta, u) \in \mathbb{T} \times K_{\rho-\delta} \mapsto \Gamma^\varepsilon(\varphi)_\theta(u)$ are well-defined and analytic, and they satisfy, for any $|\varepsilon| \geq 0$,*

$$\|\Lambda(\varphi)\|_{\rho-\delta} \leq 4C_K \quad \text{and} \quad \|\Gamma^\varepsilon(\varphi) - \text{Id}\|_{\rho-\delta} \leq 4C_K P |\varepsilon|.$$

Proof of Lemma 2.8. From the assumption $\|\varphi - \text{Id}\|_\rho \leq \delta/2$ and the Cauchy estimate (2.2) applied to $f_\theta(u) = \varphi_\theta(u) - u$, we get

$$\|\partial_u \varphi - \text{Id}\|_{\rho-\delta} \leq \frac{1}{\delta} \|\varphi - \text{Id}\|_\rho \leq \frac{1}{2}.$$

Consequently, using Neumann series, we obtain

$$\|\partial_u \varphi\|_{\rho-\delta} \leq \frac{3}{2} \quad \text{and} \quad \|(\partial_u \langle \varphi \rangle)^{-1}\|_{\rho-\delta} \leq \sum_{k=0}^{\infty} \|\partial_u \langle \varphi \rangle - \text{Id}\|_{\rho-\delta}^k \leq 2.$$

Now, for $\theta \in \mathbb{T}$, $u \in K$, $\tilde{u} \in X_{\mathbb{C}}$ with $\|\tilde{u}\|_{X_{\mathbb{C}}} < \rho - \delta$, one has

$$\|\varphi_\theta(u + \tilde{u}) - u\|_{X_{\mathbb{C}}} \leq \|\varphi - \text{Id}\|_{\rho-\delta} + \|\tilde{u}\|_{X_{\mathbb{C}}} < \frac{\delta}{2} + \rho - \delta < \rho \leq 2R.$$

Hence $\varphi_\theta(K_{\rho-\delta}) \subset K_{2R}$ and, by Assumption 2.3, we recover $\|g \circ \varphi\|_{\rho-\delta} \leq C_K$ together with $\|\langle g \circ \varphi \rangle\|_{\rho-\delta} \leq C_K$. Eventually, we obtain

$$\|\Lambda(\varphi)\|_{\rho-\delta} \leq 4C_K.$$

Integration in θ next provides $\|\Gamma^\varepsilon(\varphi) - \text{Id}\|_{\rho-\delta} \leq 4C_K P |\varepsilon|$. Besides, by composition theorems, the functions $\Lambda(\varphi)_\theta$ and $\Gamma^\varepsilon(\varphi)_\theta$ are analytic on $K_{\rho-\delta}$ in the sense of Definition 2.2. \square

Lemma 2.8 shows that, starting from a function $(\theta, u) \in \mathbb{T} \times K_{2R} \mapsto \varphi_\theta(u) \in X_{\mathbb{C}}$, we can consider iterates $(\Gamma^\varepsilon)^k(\varphi)_\theta$ at the cost of a gradual thinning of their domains of analyticity, provided ε is small enough. In order to estimate the ‘‘convergence’’ of this sequence, we also establish the following contraction property.

Lemma 2.9 *Let $0 < \delta < \rho \leq 2R$ and consider two periodic, near-identity mappings $(\theta, u) \in \mathbb{T} \times K_\rho \mapsto \varphi_\theta(u)$ and $(\theta, u) \in \mathbb{T} \times K_\rho \mapsto \hat{\varphi}_\theta(u)$, analytic on K_ρ and satisfying*

$$\|\varphi - \text{Id}\|_\rho \leq \frac{\delta}{2} \quad \text{and} \quad \|\hat{\varphi} - \text{Id}\|_\rho \leq \frac{\delta}{2}.$$

Then the following estimates hold true whenever $|\varepsilon| \geq 0$, namely

$$\|\Lambda(\varphi) - \Lambda(\hat{\varphi})\|_{\rho-\delta} \leq \frac{C_0}{\delta} \|\varphi - \hat{\varphi}\|_\rho \quad \text{and} \quad \|\Gamma^\varepsilon(\varphi) - \Gamma^\varepsilon(\hat{\varphi})\|_{\rho-\delta} \leq \frac{C_0 P |\varepsilon|}{\delta} \|\varphi - \hat{\varphi}\|_\rho.$$

Proof of Lemma 2.9. For the sake of brevity, let us denote $A_\theta(u) = \partial_u \varphi_\theta(u)$ and $\hat{A}_\theta(u) = \partial_u \hat{\varphi}_\theta(u)$. We have

$$\begin{aligned} \|\Lambda(\varphi) - \Lambda(\hat{\varphi})\|_{\rho-\delta} &\leq \|g \circ \varphi - g \circ \hat{\varphi}\|_{\rho-\delta} \\ &\quad + \|A\|_{\rho-\delta} \|\langle A \rangle^{-1}\|_{\rho-\delta} \|g \circ \varphi - g \circ \hat{\varphi}\|_{\rho-\delta} \\ &\quad + \|A\|_{\rho-\delta} \|\langle A \rangle^{-1} - \langle \hat{A} \rangle^{-1}\|_{\rho-\delta} \|g \circ \hat{\varphi}\|_{\rho-\delta} \\ &\quad + \|A - \hat{A}\|_{\rho-\delta} \|\langle \hat{A} \rangle^{-1}\|_{\rho-\delta} \|g \circ \hat{\varphi}\|_{\rho-\delta} \end{aligned}$$

Proceeding as in Lemma 2.8, we get $\|A\|_{\rho-\delta} \leq \frac{3}{2}$, together with $\|\langle A \rangle^{-1}\|_{\rho-\delta} \leq 2$, and similarly for \hat{A} . Besides, using the relation $\langle A \rangle^{-1} - \langle \hat{A} \rangle^{-1} = \langle A \rangle^{-1} \langle \hat{A} - A \rangle \langle \hat{A} \rangle^{-1}$, a Cauchy estimate provides

$$\|\langle A \rangle^{-1} - \langle \hat{A} \rangle^{-1}\|_{\rho-\delta} \leq \|\langle A \rangle^{-1}\|_{\rho-\delta} \|\langle \hat{A} \rangle^{-1}\|_{\rho-\delta} \|A - \hat{A}\|_{\rho-\delta} \leq \frac{4}{\delta} \|\varphi - \hat{\varphi}\|_{\rho}.$$

Finally, whenever $u \in K_{\rho-\delta}$, since $\|\varphi - \text{Id}\|_{\rho} \leq \delta/2$ and similarly for $\hat{\varphi}$, we recover in particular $\varphi_{\theta}(u) \in K_{\rho-\delta/2}$ and $\hat{\varphi}_{\theta}(u) \in K_{\rho-\delta/2}$. We deduce

$$\|\langle g \circ \hat{\varphi} \rangle\|_{\rho-\delta} \leq C_K$$

and

$$\|g \circ \varphi - g \circ \hat{\varphi}\|_{\rho-\delta} \leq \|\partial_u g\|_{\rho-\delta/2} \|\varphi - \hat{\varphi}\|_{\rho-\delta} \leq \frac{2C_K}{\delta} \|\varphi - \hat{\varphi}\|_{\rho-\delta}.$$

Collecting all terms, we have

$$\|\Lambda(\varphi) - \Lambda(\hat{\varphi})\|_{\rho-\delta} \leq \left(\frac{2C_K}{\delta} + \frac{6C_K}{\delta} + \frac{6C_K}{\delta} + \frac{2C_K}{\delta} \right) \|\varphi - \hat{\varphi}\|_{\rho}$$

and the corresponding bound for Γ^{ε} is obtained by integration in θ . \square

2.4.2 Existence and uniqueness of quasi-solutions to (2.8) with polynomial remainder

In this section we exhibit (quasi-) solutions to the equations of averaging that we have pointed out, namely equation (2.8) or equivalently (2.9) or (2.12) for the mapping $\Phi_{\theta}^{\varepsilon}$ and equation (2.7) for the autonomous vector field G^{ε} . The exhibited functions are only quasi-solutions, in that the corresponding equations are only solved to within a polynomial error term of size $\mathcal{O}(\varepsilon^{n+1})$ for any $n \geq 0$. We begin with the proof of Lemma 2.6.

Proof of Lemma 2.6. The function $\Phi_{\theta}^{[0]} = \text{Id}$ is clearly analytic in the sense of Definition 2.2 on K_{R_0} , analytic for all ε , smooth in θ , and clearly satisfies (2.17). Assume now that, for an integer $k \leq n$, the function $\Phi_{\theta}^{[k]}$ is analytic on K_{R_k} , analytic in ε for $|\varepsilon| < \varepsilon_0/(n+1)$, smooth in θ , and satisfies (2.17). Then, according to Lemma 2.8 with $\rho = R_k$ and $\delta = \frac{R}{n+1} = r_n$, the function $\Phi_{\theta}^{[k+1]}$ is well-defined by (2.14), analytic on $K_{R_{k+1}}$ and satisfies for $0 \leq |\varepsilon| < \varepsilon_0/(n+1)$ the estimate

$$\|\Phi^{[k+1]} - \text{Id}\|_{R_{k+1}} \leq 4C_K P |\varepsilon| < \frac{4C_K P \varepsilon_0}{n+1} = \frac{R}{2(n+1)} = \frac{r_n}{2}.$$

As a composition of analytic functions, it is furthermore analytic in ε again for $|\varepsilon| < \varepsilon_0/(n+1)$. Smoothness in θ is also clear. This finishes the induction argument for $\Phi^{[k]}$. By definition (2.15), $G^{[k]}$ is then also analytic in K_{R_k} and so is G_{k+1} , defined by (2.16). \square

We are now in position to establish the existence of quasi-solutions to (2.8).

Theorem 2.10 [Existence of quasi-solutions to (2.8)] For $n \in \mathbb{N}$, consider the sequence of functions

$$\Phi_\theta^{[0]} = \text{Id}, \quad \Phi_\theta^{[k+1]} = \Gamma^\varepsilon(\Phi^{[k]})_\theta, \quad k = 0, \dots, n,$$

and the associated sequence of defects $(\delta_\theta^{[k]})_{k=0, \dots, n}$, defined as

$$\begin{aligned} \varepsilon \delta_\theta^{[k]}(u) &:= \frac{\partial \Phi_\theta^{[k]}}{\partial \theta}(u) - \varepsilon \Lambda(\Phi^{[k]})_\theta(u) \\ &= \frac{\partial \Phi_\theta^{[k]}}{\partial \theta}(u) + \varepsilon \frac{\partial \Phi_\theta^{[k]}}{\partial u}(u) \left(\frac{\partial \langle \Phi^{[k]} \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \Phi^{[k]} \rangle(u) - \varepsilon g_\theta \circ \Phi_\theta^{[k]}(u). \end{aligned} \quad (2.24)$$

Then, the following holds true:

(i) The mappings $\Phi_\theta^{[n]}$ and $\delta_\theta^{[n]}$ are C^1 in θ , analytic on respectively K_{R+r_n} and K_R , and analytic in $\varepsilon \in \mathbb{C}$ whenever $|\varepsilon| < \varepsilon_0/(n+1)$.

(ii) The mappings $\Phi_\theta^{[n]}$ and $\delta_\theta^{[n]}$ satisfy the following estimates for all $|\varepsilon| < \varepsilon_0/(n+1)$

$$\|\Phi^{[n]} - \text{Id}\|_{R+r_n} \leq \frac{r_n}{2}, \quad (2.25)$$

$$\|\delta^{[n]}\|_R \leq C_1 (\eta(n+1)|\varepsilon|)^n. \quad (2.26)$$

(iii) For all $\theta \in \mathbb{T}$, the mapping $u \mapsto \Phi_\theta^{[n]}(u)$ has an inverse defined on K_R with values in K_{R+r_n} . This inverse $(\Phi^{[n]})^{-1}$ is analytic. Moreover, we have $\|(\Phi^{[n]})^{-1}\|_{R-r_n} \leq R$.

Remark 2.11 In other words, the n -th iterate $\Phi_\theta^{[n]}$ satisfies equation (2.8) up to a remainder term of size $C_1(\eta(n+1))^n |\varepsilon|^{n+1}$.

Proof of Theorem 2.10. Statement (i) and estimate (2.25) are obvious consequences of Lemma 2.6, up to a possible singularity at $\varepsilon = 0$ which is ruled out by estimate (2.26), which we now prove. On the one hand, by differentiation of $\Phi_\theta^{[n]} = \Gamma^\varepsilon(\Phi^{[n-1]})_\theta$, we recover for $n \geq 1$

$$\varepsilon \delta_\theta^{[n]} = \varepsilon \left(\Lambda(\Phi^{[n-1]})_\theta - \Lambda(\Phi^{[n]})_\theta \right).$$

Hence, the contraction property stated in Lemma 2.9 yields with $\delta = r_n$ and $\rho = R + \delta = R_n$.

$$\begin{aligned} |\varepsilon| \|\delta^{[n]}\|_R &\leq \frac{C_0 |\varepsilon|}{r_n} \|\Phi^{[n]} - \Phi^{[n-1]}\|_{R_n} = \frac{1}{P} \frac{C_0 P |\varepsilon|}{r_n} \left\| \Gamma^\varepsilon(\Phi^{[n-1]}) - \Gamma^\varepsilon(\Phi^{[n-2]}) \right\|_{R_n} \\ &\leq \frac{1}{P} \left(\frac{C_0 P |\varepsilon|}{r_n} \right)^2 \left\| \Gamma^\varepsilon(\Phi^{[n-2]}) - \Gamma^\varepsilon(\Phi^{[n-3]}) \right\|_{R_{n-1}} \leq \dots \\ &\leq \frac{1}{P} \left(\frac{C_0 P |\varepsilon|}{r_n} \right)^n \|\Phi^{[1]} - \Phi^{[0]}\|_{R_1}. \end{aligned}$$

On the other hand, by definition of $\Phi_\theta^{[1]}$, we have $\|\Phi^{[1]} - \Phi^{[0]}\|_{R_1} \leq 2|\varepsilon| P C_K$. This proves (2.26).

As for Statement (iii), it is clear that if we take $u_1, u_2 \in K_R$ such that $\Phi_\theta^{[n]}(u_1) = \Phi_\theta^{[n]}(u_2)$, we have

$$\begin{aligned} \|u_1 - u_2\|_{X_{\mathbb{C}}} &\leq \|\partial_u \Phi^{[n]} - \text{Id}\|_R \|u_1 - u_2\|_{X_{\mathbb{C}}} \leq \frac{1}{r_n} \|\Phi^{[n]} - \text{Id}\|_{R+r_n} \|u_1 - u_2\|_{X_{\mathbb{C}}} \\ &\leq \frac{1}{2} \|u_1 - u_2\|_{X_{\mathbb{C}}} \end{aligned}$$

so that $u_1 = u_2$. As for the existence part, given $(u, \tilde{u}) \in K \times X_{\mathbb{C}}$ with $\rho := \|\tilde{u}\|_{X_{\mathbb{C}}} < R$, we consider the sequence v_k defined by

$$v_0 = u + \tilde{u} \in K_R, \quad v_{k+1} = v_k - \Phi_\theta^{[n]}(v_k) + u + \tilde{u}.$$

Firstly, whenever $v_k \in K_{(R+\rho+r_n)/2}$, we have

$$\|v_{k+1} - u\|_{X_{\mathbb{C}}} \leq \rho + \|\Phi_\theta^{[n]} - \text{Id}\|_{R+r_n} < \frac{R + \rho + r_n}{2},$$

from which we deduce by induction that the whole the sequence $(v_k)_{k \in \mathbb{N}}$ belongs to $K_{(R+\rho+r_n)/2}$. Moreover, we may write

$$\begin{aligned} \|v_{k+1} - v_k\|_{X_{\mathbb{C}}} &= \|(\Phi_\theta^{[n]} - \text{Id})(v_k) - (\Phi_\theta^{[n]} - \text{Id})(v_{k-1})\|_{X_{\mathbb{C}}} \\ &\leq \|\partial_u \Phi^{[n]} - \text{Id}\|_{(R+\rho+r_n)/2} \|v_k - v_{k-1}\|_{X_{\mathbb{C}}} \leq \frac{1}{1 + (R - \rho)/r_n} \|v_k - v_{k-1}\|_{X_{\mathbb{C}}}, \end{aligned}$$

hence the sequence v_k converges towards some $v \in \overline{K_{(R+\rho+r_n)/2}} \subset K_{R+r_n}$. If $\rho < R - r_n$, we have furthermore $\|v\| < R$. The analyticity of $(\Phi_\theta^{[n]})^{-1}$ is a direct consequence of the *Inverse Function Theorem*. \square

Remark 2.12 *From the definition of the sequence $(\Phi_\theta^{[k]})_{k=0, \dots, n+1}$, it can be inductively inferred that, if u belongs to $K_R \cap X$ and $\varepsilon \in \mathbb{R}$, then $\Phi_\theta^{[n]}(u)$, $G^{[n]}(u)$ and $\delta_\theta^{[n]}(u)$ belong to X as well.*

The next Theorem establishes that the first n terms in the expansions of $\Phi_\theta^{[n]}$ in powers of ε , are *independent* of the construction used (a similar property naturally holds also for $G^{[n]}$). This validates the fact that equations (2.8) and (2.7) are ‘‘canonical’’ in some sense.

Theorem 2.13 [Uniqueness of quasi-solutions to (2.8)] *Fix $n \in \mathbb{N}$ and consider a function $(\theta, u) \mapsto \widehat{\Phi}_\theta(u)$, which is C^1 in $\theta \in \mathbb{T}$, analytic on K_{R+r_n} , analytic in ε for $|\varepsilon| < \varepsilon_0/(n+1)$, and satisfies*

$$\widehat{\Phi}_0 = \text{Id}, \quad \|\widehat{\Phi} - \text{Id}\|_{R+r_n} \leq \frac{r_n}{2}. \quad (2.27)$$

Assume that the defect associated with $\widehat{\Phi}_\theta$, defined as

$$\varepsilon \widehat{\delta}_\theta(u) := \frac{\partial \widehat{\Phi}_\theta}{\partial \theta} - \varepsilon \Lambda \left(\widehat{\Phi} \right)_\theta$$

satisfies, for all $|\varepsilon| < \varepsilon_0/(n+1)$, the estimate $\|\widehat{\delta}\|_R \leq \widehat{C} |\varepsilon|^n$ for some constant $\widehat{C} > 0$ independent of ε . Then we necessarily have, whenever $|\varepsilon| < \varepsilon_0/(4(n+1))$, the estimate

$$\|\widehat{\Phi} - \Phi^{[n]}\|_{r_n} \leq C_3(n) |\varepsilon|^{n+1},$$

where $\Phi_\theta^{[n]}$ is the function defined in Theorem 2.10 and $C_3(n) = P \left(2\widehat{C} + 20C_K (\eta(n+1))^n \right)$.

Proof of Theorem 2.13. The result stems from successive applications of the contraction Lemma 2.9. First, by Lemma 2.8 and starting from (2.27) and (2.17), we have for any $k \leq n$,

$$\|(\Gamma^\varepsilon)^k(\widehat{\Phi}) - \text{Id}\|_{R+r_n-kr_n} \leq \frac{r_n}{2}, \quad \|(\Gamma^\varepsilon)^k(\Phi^{[n]}) - \text{Id}\|_{R+r_n-kr_n} \leq \frac{r_n}{2}, \quad (2.28)$$

where we used the fact that $4C_K P |\varepsilon| \leq \frac{r_n}{2}$. These estimates allow for the application of Lemma 2.9 and we get, for all for $k \leq n$,

$$\|(\Gamma^\varepsilon)^{k+1}(\widehat{\Phi}) - (\Gamma^\varepsilon)^k(\widehat{\Phi})\|_{R-kr_n} \leq \left(\frac{C_0 P |\varepsilon|}{r_n} \right)^k \|\Gamma^\varepsilon(\widehat{\Phi}) - \widehat{\Phi}\|_R \leq \frac{1}{2^k} \|\Gamma^\varepsilon(\widehat{\Phi}) - \widehat{\Phi}\|_R$$

and similarly for $\Phi^{[n]}$, where we have used $C_0 P |\varepsilon|/r_n \leq 1/2$. Summation provides

$$\|(\Gamma^\varepsilon)^{n+1}(\widehat{\Phi}) - \widehat{\Phi}\|_{r_n} \leq \sum_{k=0}^n \|(\Gamma^\varepsilon)^{k+1}(\widehat{\Phi}) - (\Gamma^\varepsilon)^k(\widehat{\Phi})\|_{r_n} \leq 2 \|\Gamma^\varepsilon(\widehat{\Phi}) - \widehat{\Phi}\|_R, \quad (2.29)$$

and similarly for $\Phi^{[n]}$. On the other hand, using (2.28) and applying again $n+1$ times the contraction Lemma 2.9, we obtain

$$\|(\Gamma^\varepsilon)^{n+1}(\widehat{\Phi}) - (\Gamma^\varepsilon)^{n+1}(\Phi^{[n]})\|_{r_n} \leq \left(\frac{C_0 P |\varepsilon|}{r_n} \right)^{n+1} \|\widehat{\Phi} - \Phi^{[n]}\|_{R+r_n} \leq \left(\frac{C_0 P |\varepsilon|}{r_n} \right)^{n+1} r_n, \quad (2.30)$$

where the last inequality uses (2.27) and the similar estimate for $\Phi^{[n]}$. Finally, from estimate (2.29) on $\widehat{\Phi}$ and the similar bound on $\Phi^{[n]}$, and from (2.30), we deduce

$$\begin{aligned} & \|\widehat{\Phi} - \Phi^{[n]}\|_{r_n} \\ & \leq \|\widehat{\Phi} - (\Gamma^\varepsilon)^{n+1}(\widehat{\Phi})\|_{r_n} + \|\Phi^{[n]} - (\Gamma^\varepsilon)^{n+1}(\Phi^{[n]})\|_{r_n} + \|(\Gamma^\varepsilon)^{n+1}(\widehat{\Phi}) - (\Gamma^\varepsilon)^{n+1}(\Phi^{[n]})\|_{r_n} \\ & \leq 2\|\widehat{\Phi} - \Gamma^\varepsilon(\widehat{\Phi})\|_R + 2\|\Phi^{[n]} - \Gamma^\varepsilon(\Phi^{[n]})\|_R + C_0 P \left(\frac{C_0 P}{R} (n+1) \right)^n |\varepsilon|^{n+1} \\ & \leq 2\widehat{C} P |\varepsilon|^{n+1} + 2C_1 P (\eta(n+1)|\varepsilon|)^n \varepsilon + C_0 P \left(\frac{C_0 P}{R} (n+1) \right)^n |\varepsilon|^{n+1}, \end{aligned} \quad (2.31)$$

where we have used

$$\Phi_\theta^{[n]} - \Gamma^\varepsilon(\Phi^{[n]})_\theta = \varepsilon \int_0^\theta \delta_\xi^{[n]} d\xi, \quad \text{and} \quad \widehat{\Phi}_\theta - \Gamma^\varepsilon(\widehat{\Phi})_\theta = \varepsilon \int_0^\theta \widehat{\delta}_\xi d\xi,$$

together with estimate (2.26) on $\delta^{[n]}$ and the assumption on $\widehat{\delta}$. Gathering the various constants gives the result. \square

2.4.3 Proof of Theorem 2.7

Up to now, we have obtained a collection of mappings $\Phi_\theta^{[n]}$ which produce a defect in equation (2.8) (or (2.13)) that is of size $\mathcal{O}((n+1)^n \varepsilon^{n+1})$. By optimizing the choice of the parameter n , we now produce a mapping $\Phi^{[n_\varepsilon]}$ associated with a defect of order $\mathcal{O}(\exp(-c/\varepsilon))$ for some $c > 0$.

Proof. Part (i). Since $(n_\varepsilon + 1)\varepsilon \leq \frac{\mu}{\eta} = \frac{\mu\varepsilon_0}{2} < \varepsilon_0$, Theorem 2.10 applies with $n = n_\varepsilon$ and we have that

$$\partial_\theta \tilde{\Phi}_\theta^\varepsilon(u) = \varepsilon g_\theta \circ \tilde{\Phi}_\theta^\varepsilon(u) - \varepsilon \partial_u \tilde{\Phi}_\theta^\varepsilon(u) G^{[n_\varepsilon]}(u) + \varepsilon \delta_\theta^{[n_\varepsilon]}(u),$$

whenever $u \in K_R \subset K_{R+r_{n_\varepsilon}}$. Therefore, introducing the exact solution $U(t)$ of the equation

$$\frac{dU(t)}{dt} = \varepsilon \left(\tilde{G}^\varepsilon(U(t)) + \mathcal{R}_t^\varepsilon(U(t)) \right), \quad U(0) = u_0,$$

with $\mathcal{R}_t^\varepsilon(u) = G^{[n_\varepsilon]}(u) - \tilde{G}^\varepsilon(u) - \left(\partial_u \tilde{\Phi}_t^\varepsilon(u) \right)^{-1} \delta_t^{[n_\varepsilon]}(u)$, the function $u^\varepsilon(t) := \tilde{\Phi}_t^\varepsilon(U(t))$ clearly satisfies $u^\varepsilon(0) = u_0$ together with

$$\begin{aligned} \frac{du^\varepsilon}{dt}(t) &= \varepsilon (g_t \circ \tilde{\Phi}_t^\varepsilon)(U(t)) - \varepsilon \partial_u \tilde{\Phi}_t^\varepsilon(U(t)) \cdot G^{[n_\varepsilon]}(U(t)) + \varepsilon \delta_t^{[n_\varepsilon]}(U(t)) \\ &\quad + \varepsilon \partial_u \tilde{\Phi}_t^\varepsilon(U(t)) \cdot \left(G^{[n_\varepsilon]}(U(t)) - \left(\partial_u \tilde{\Phi}_t^\varepsilon(U(t)) \right)^{-1} \delta_t^{[n_\varepsilon]}(U(t)) \right) \\ &= \varepsilon (g_t \circ \tilde{\Phi}_t^\varepsilon)(U(t)) = \varepsilon g_t(u^\varepsilon(t)), \end{aligned}$$

as desired. Hence $u^\varepsilon(t)$ coincides for any time $t \in [0, T/\varepsilon]$ with the solution of (2.1). Now, on the one hand, Theorem 2.10 and the choice of n^ε ensure that the defect $\delta_\theta^{[n_\varepsilon]}$ satisfies

$$\|\delta^{[n_\varepsilon]}\|_R \leq C_1 (\eta(n_\varepsilon + 1)\varepsilon)^{n_\varepsilon} \leq C_1 \mu^{n_\varepsilon},$$

and on the other hand, the analyticity of $G^{[n_\varepsilon]}$ w.r.t. ε and Cauchy's formulae allow to write for all $u \in K_R$ and $\delta := \frac{\varepsilon_0}{2(n_\varepsilon + 1)}$, the estimate

$$\begin{aligned} \|G^{[n_\varepsilon]}(u) - \tilde{G}^\varepsilon(u)\|_{X_C} &= \left\| \sum_{k \geq n_\varepsilon + 1} \frac{\varepsilon^k}{k!} \frac{d^k G^{[n_\varepsilon]}}{d\varepsilon^k} \Big|_{\varepsilon=0} (u) \right\|_{X_C} \\ &\leq \sum_{k \geq n_\varepsilon + 1} \frac{\varepsilon^k}{k!} \frac{k!}{\delta^k} \sup_{|\varepsilon| < \delta} \|G^{[n_\varepsilon]}(u)\|_{X_C} \leq \frac{(\varepsilon/\delta)^{n_\varepsilon + 1}}{1 - (\varepsilon/\delta)} \sup_{|\varepsilon| < \delta} \|G^{[n_\varepsilon]}\|_R \leq C_1 \frac{\mu^{n_\varepsilon + 1}}{1 - \mu}, \end{aligned}$$

where we have used $\left\| \left(\partial_u \langle \Phi^{[n_\varepsilon]} \rangle \right)^{-1} \right\|_R \leq 2$ to bound³ $G^{[n_\varepsilon]}$ on K_R by C_1 and $|\varepsilon|/\delta \leq \mu$. To conclude, it remains to write

$$\begin{aligned} \|\mathcal{R}\|_R &\leq \|G^{[n_\varepsilon]} - \tilde{G}^\varepsilon\|_R + \left\| \left(\partial_u \Phi^{[n_\varepsilon]} \right)^{-1} \right\|_R \|\delta^{[n_\varepsilon]}\|_R \\ &\leq \frac{(2 - \mu)C_1}{1 - \mu} \mu^{n_\varepsilon} \leq \frac{(2 - \mu)C_1}{(1 - \mu)\mu^2} \exp\left(-\frac{\mu|\log(\mu)|}{\eta\varepsilon}\right), \end{aligned}$$

³This stems from Lemma 2.8-(i) together with the known estimate $\|\Phi^{[n_\varepsilon]} - \text{Id}\|_{R+r_{n_\varepsilon}} \leq r_{n_\varepsilon}/2$ (where $r_{n_\varepsilon} = R/(n_\varepsilon + 1)$), as established in Theorem 2.10.

where we have used $0 < \mu < 1$ and $n_\varepsilon \geq (\mu/(\eta|\varepsilon|)) - 2$.

Parts (ii) and (iii). Let $\tilde{\Psi}_t^\varepsilon$ be the flow of the autonomous equation

$$\frac{dU}{dt} = \varepsilon \tilde{G}^\varepsilon(U).$$

There exists $T_1 > 0$ such that $\tilde{\Psi}_t^\varepsilon(u_0)$ is well-defined for all $0 \leq t \leq T_1/\varepsilon$, given that \tilde{G}^ε is analytic on K_R , hence Lipschitz continuous on any K_ρ with $\rho < R$. Now, we have, on the one hand,

$$\frac{du^\varepsilon(t)}{dt} = \varepsilon g_t(u^\varepsilon(t))$$

and, on the other hand with $\tilde{u}^\varepsilon(t) = \tilde{\Phi}_t^\varepsilon \circ \tilde{\Psi}_t^\varepsilon(u_0)$

$$\frac{d\tilde{u}^\varepsilon(t)}{dt} = \varepsilon g_t(\tilde{u}^\varepsilon(t)) - \varepsilon \left(\partial_u \tilde{\Phi}_t^\varepsilon \circ (\tilde{\Phi}_t^\varepsilon)^{-1} \right) (\tilde{u}^\varepsilon(t)) \cdot \left(\mathcal{R}_t^\varepsilon \circ (\tilde{\Phi}_t^\varepsilon)^{-1} \right) (\tilde{u}^\varepsilon(t))$$

as long as \tilde{u}^ε and $(\tilde{\Phi}_t^\varepsilon)^{-1}(\tilde{u}^\varepsilon(t))$ remain in K_R . If $L = \frac{C_K}{R}$ denotes a Lipschitz constant for g on K_R , a variant of Gronwall Lemma then gives

$$\|u^\varepsilon(t) - \tilde{u}^\varepsilon(t)\|_{X_C} \leq \frac{3}{2} \|\mathcal{R}^\varepsilon\|_R \frac{e^{\varepsilon L t} - 1}{L} \leq \frac{6R}{\mu^2(1-\mu)} e^{\varepsilon L t - \frac{\mu|\log(\mu)|}{\varepsilon\eta}} := M(t, \varepsilon, \mu)$$

where we have used the bound $\|\partial_u \tilde{\Phi}_t^\varepsilon - \text{Id}\|_R \leq 1/2$.

Now we recall that, by assumption of the Theorem, $u^\varepsilon(t)$ exists and belongs to K for $0 < \varepsilon < \varepsilon^*$ and $0 \leq t \leq T/\varepsilon^{1+\alpha}$ (for Part (ii), we have $\alpha = 0$ and, for Part (iii), we have $0 < \alpha < 1$). In particular, choosing $\varepsilon < \mu\varepsilon_2$ with ε_2 small enough so that $M(T/\varepsilon, \varepsilon, \mu) \leq R - r_{n_\varepsilon}$, ensures that \tilde{u}^ε and $(\tilde{\Phi}_t^\varepsilon)^{-1}(\tilde{u}^\varepsilon(t))$ remain in K_R whenever $t \leq T/\varepsilon^{1+\alpha}$ (hence $T_1 \geq T/\varepsilon^\alpha$). Now we claim that, with the choice

$$0 < \varepsilon < \min \left(\varepsilon^*, \left(\frac{\mu\varepsilon_0^2 P}{T} \right)^{\frac{1}{1-\alpha}}, \frac{\mu\varepsilon_0}{28} \right), \quad (2.32)$$

(which means $0 < \varepsilon < \min(\varepsilon^*, \mu\varepsilon_2)$ in the case $\alpha = 0$), we obtain estimate (2.23).

Proof of the claim. For $\varepsilon^{1-\alpha} \leq \mu \frac{|\log(\mu)|}{2\eta L T} = \mu\varepsilon_0^2 |\log(\mu)| \frac{2P}{T}$, we have $e^{\varepsilon L t - \frac{\mu|\log(\mu)|}{\varepsilon\eta}} \leq e^{-\frac{\mu|\log(\mu)|}{2\varepsilon\eta}}$ on $[0, T/\varepsilon^{1+\alpha}]$, so that for all $0 \leq t \leq T/\varepsilon^{1+\alpha}$

$$M(t, \varepsilon, \mu) \leq \frac{6R}{\mu^2(1-\mu)} e^{-\frac{\mu|\log(\mu)|}{2\varepsilon\eta}}.$$

The quantity $M(T/\varepsilon^{1+\alpha}, \varepsilon, \mu)$ is then bounded by $R/2 \leq R - r_{n_\varepsilon}$ (note that, by definition of n_ε and ε_2 , we have $n_\varepsilon \geq 1$) if furthermore

$$\varepsilon < \mu \frac{\varepsilon_0}{4} \frac{-\log(\mu)}{\log \left(\frac{12}{\mu^2(1-\mu)} \right)}.$$

Imposing for instance that $0 < \mu < 1/2$, a combined bound on ε is given by (2.32), a condition under which

$$\forall t \in [0, T/\varepsilon^{1+\alpha}], \quad \|u^\varepsilon(t) - \tilde{u}^\varepsilon(t)\|_{X_C} \leq 12 R \mu^{-2} e^{-\frac{\mu|\log(\mu)|\varepsilon_0}{4\varepsilon}}.$$

This proves the claim and the proof of Theorem 2.7 is complete. \square

2.5 The linear case

In this section, we consider the particular case

$$g_\theta(u) \equiv A_\theta u,$$

where A_θ is a bounded linear operator on X . In this case, it turns out the change of variable solution of (2.8) can be exactly constructed, due to the fact that our iterative procedure actually converges. Let us prove this statement. Naturally, Assumption 2.3 is replaced here by the following assumption.

Assumption 2.14 *The map $\theta \mapsto A_\theta \in \mathcal{L}(X)$ is continuous and P -periodic.*

The initial value problem reads

$$\begin{aligned} \frac{d}{dt} u^\varepsilon(t) &= \varepsilon A_t u^\varepsilon(t), & u^\varepsilon(t) &\in X, \\ u^\varepsilon(0) &= u_0, & u_0 &\in X. \end{aligned} \quad (2.33)$$

We denote by $\|\cdot\|_{\mathcal{L}(X)}$ the operator norm on X and the space $C(\mathbb{T}, \mathcal{L}(X))$ is equipped with the norm

$$\|\Phi\| = \sup_{\theta \in \mathbb{T}} \|\Phi_\theta\|_{\mathcal{L}(X)} = \sup_{(\theta, u) \in \mathbb{T} \times X, \|u\|_X=1} \|\Phi_\theta\|_X.$$

In the present linear setting, Theorem 2.7 takes the following form.

Theorem 2.15 [Exact averaging in the linear case] *Consider $u^\varepsilon(t)$ the solution of (2.33) and denote $\varepsilon_l = \frac{1}{\alpha}(\frac{3}{2} - \sqrt{2})$, with $\alpha = \int_{\mathbb{T}} \|A_\theta\|_{\mathcal{L}(X)} d\theta$. Then for all $0 \leq \varepsilon < \varepsilon_l$, there exists a map $\theta \mapsto \Phi_\theta^\varepsilon \in \mathcal{L}(X)$, such that*

- (i) *The function Φ_θ^ε is P -periodic and C^1 in θ , and satisfies $\Phi_{\theta=0} = \text{Id}$.*
- (ii) *For all $\theta \in \mathbb{T}$, the operator Φ_θ^ε is invertible in $\mathcal{L}(X)$.*
- (iii) *For all $u_0 \in X$, the solution of (2.33) admits the factorized form*

$$\forall t \in \mathbb{R}_+ \quad u^\varepsilon(t) = \Phi_t^\varepsilon e^{\varepsilon t G^\varepsilon} u_0,$$

where $G^\varepsilon \in \mathcal{L}(X)$ is defined by

$$G^\varepsilon = \langle \Phi^\varepsilon \rangle^{-1} \langle A \Phi^\varepsilon \rangle. \quad (2.34)$$

Proof of Theorem 2.15. In the linear framework, equation (2.8) becomes

$$\frac{d\Phi_\theta^\varepsilon}{d\theta} + \varepsilon \Phi_\theta^\varepsilon \langle \Phi^\varepsilon \rangle^{-1} \langle A \Phi^\varepsilon \rangle = \varepsilon A \Phi_\theta^\varepsilon, \quad \text{or equivalently} \quad \Phi_\theta^\varepsilon = \Gamma^\varepsilon(\Phi^\varepsilon)_\theta,$$

where the nonlinear map Γ^ε acts on the set of functions in $C(\mathbb{T}, \mathcal{L}(X))$ which are invertible for all θ , and is defined for any such function Φ by

$$\Gamma^\varepsilon(\Phi)_\theta = \text{Id} + \varepsilon \int_0^\theta (A_\xi \Phi_\xi - \Phi_\xi \langle \Phi \rangle^{-1} \langle A \Phi \rangle) d\xi. \quad (2.35)$$

We introduce the sequence $\Phi_\theta^{[0]} = \text{Id}$, with $\Phi_\theta^{[n+1]} = \Gamma^\varepsilon(\Phi^{[n]})_\theta$ for all $n \in \mathbb{N}$. We claim that this sequence is well-defined for all $n \in \mathbb{N}$, with the following estimate

$$\|\Phi^{[n]} - \text{Id}\| < d_* := \frac{1}{2} - \varepsilon\alpha - \sqrt{\varepsilon^2\alpha^2 - 3\varepsilon\alpha + \frac{1}{4}}. \quad (2.36)$$

Note that the term $\varepsilon^2\alpha^2 - 3\varepsilon\alpha + \frac{1}{4}$ is positive due to the assumption $\varepsilon < \varepsilon_l$. We now prove (2.36) by induction (it clearly holds true in the case $n = 0$). Assume the claim is proved for some n , and set $d_n := \|\Phi^{[n]} - \text{Id}\|$. From the fact that $0 < \varepsilon < \varepsilon_l$, and from the induction assumption, we first deduce $d_n < d_* < \frac{1}{2} + \varepsilon\alpha \leq 2 - \sqrt{2} < 1$. Then, by Neumann series, we deduce that $\Phi_\theta^{[n]}$ is invertible for all θ , and that $(\Phi_\theta^{[n]})^{-1} \in C(\mathbb{T}, \mathcal{L}(X))$. In particular, $\Phi_\theta^{[n+1]} = \Gamma^\varepsilon(\Phi^{[n]})_\theta$ is well-defined. Moreover, from (2.35), we estimate

$$\|\Gamma^\varepsilon(\Phi^{[n]}) - \text{Id}\| \leq \varepsilon \|\Phi^{[n]}\| \left(1 + \frac{\|\Phi^{[n]}\|}{1 - \|\Phi^{[n]} - \text{Id}\|} \right) \int_{\mathbb{T}} \|A_\theta\|_{\mathcal{L}(X)} d\theta,$$

from which we deduce $d_{n+1} \leq 2\varepsilon\alpha \frac{1 + d_n}{1 - d_n}$. Since d_* is the smallest root of $d(1 - d) - 2\varepsilon\alpha(1 + d)$, thanks to the assumption $0 < \varepsilon < \varepsilon_l$, a direct analysis of the function $f(d) = 2\varepsilon\alpha \frac{1+d}{1-d}$ shows that $0 \leq d \leq d_*$ implies $0 \leq f(d) \leq f(d_*) = d_*$. Consequently $d_{n+1} \leq f(d_n) < d_*$ and (2.36) is proved.

To prove the Theorem, there only remains to prove that the Lipschitz constant of Γ^ε is less than one on the domain $\{\Phi \text{ s.t. } \|\Phi - \text{Id}\| \leq d_*\}$. To do so, we take $\Phi \in C(\mathbb{T}, \mathcal{L}(X))$ and $\widehat{\Phi} \in C(\mathbb{T}, \mathcal{L}(X))$ satisfying $\|\Phi - \text{Id}\| < d_*$ and $\|\widehat{\Phi} - \text{Id}\| < d_*$. As before, we readily deduce that for all $\theta \in \mathbb{T}$, both operators Φ_θ and $\widehat{\Phi}_\theta$ are invertible on X , with inverse operators satisfying $\|\Phi^{-1}\| \leq 1/(1 - d_*)$ and $\|\widehat{\Phi}^{-1}\| \leq 1/(1 - d_*)$. Hence, as in Lemma 2.9, we write

$$\begin{aligned} \|\Gamma^\varepsilon(\Phi) - \Gamma^\varepsilon(\widehat{\Phi})\| &\leq \varepsilon P \langle \|A(\Phi - \widehat{\Phi})\|_{\mathcal{L}(X)} \rangle + \varepsilon P \|\Phi\| \|\Phi^{-1}\| \langle \|A(\Phi - \widehat{\Phi})\|_{\mathcal{L}(X)} \rangle \\ &\quad + \varepsilon P \|\Phi\| \|\Phi^{-1} - \widehat{\Phi}^{-1}\| \langle \|A\widehat{\Phi}\|_{\mathcal{L}(X)} \rangle + \varepsilon P \|\Phi - \widehat{\Phi}\| \|\widehat{\Phi}^{-1}\| \langle \|A\widehat{\Phi}\|_{\mathcal{L}(X)} \rangle \\ &\leq \varepsilon\alpha \left(1 + \|\Phi\| \|\Phi^{-1}\| + \|\Phi\| \|\widehat{\Phi}\| \|\Phi^{-1}\| \|\widehat{\Phi}^{-1}\| + \|\widehat{\Phi}\| \|\widehat{\Phi}^{-1}\| \right) \|\Phi - \widehat{\Phi}\| \\ &\leq \varepsilon\alpha \left(1 + \|\Phi\| \|\Phi^{-1}\| \right) \left(1 + \|\widehat{\Phi}\| \|\widehat{\Phi}^{-1}\| \right) \|\Phi - \widehat{\Phi}\| \\ &\leq \frac{4\varepsilon\alpha}{(1 - d_*)^2} \|\Phi - \widehat{\Phi}\|. \end{aligned}$$

Introducing d^* , the largest root of $d(1 - d) - 2\varepsilon\alpha(1 + d)$, we now observe (using $d^*d_* = 2\varepsilon\alpha$ and $d^* + d_* = 1 - 2\varepsilon\alpha$) that

$$\frac{4\varepsilon\alpha}{(1 - d_*)^2} < \frac{4\varepsilon\alpha}{(1 - d_*)(1 - d^*)} = 1.$$

Convergence of the sequence $\Phi^{[n]}$ in $C(\mathbb{T}, \mathcal{L}(X))$ follows, and the Theorem is easily deduced. \square

3 Geometric aspects

One of the advantages of stroboscopic averaging is that it preserves the geometric properties of the initial equation. In this spirit, we now prove that stroboscopic averaging preserves both the Hamiltonian structure and the invariants of the original equation (if any).

3.1 Preservation of the Hamiltonian structure

We assume here that X is a Hilbert space, i.e. the norm $\|\cdot\|_X$ stems from a real scalar product $(\cdot, \cdot)_X$. Moreover, for further application to the case of the nonlinear Schrödinger equation, we assume that X is a dense subspace continuously embedded in some ambient Hilbert space Z , with real scalar product $(\cdot, \cdot)_Z$. The dual space X' is then identified through the duality given by the scalar product $(\cdot, \cdot)_Z$. In practice, the workspace X will be a Sobolev space $H^s(\mathbb{R}^d)$ (for some "large" $s > 0$) and the ambient space Z will be $L^2(\mathbb{R}^d)$, so that the dual space X' is $H^{-s}(\mathbb{R}^d)$.

In this context, we introduce the following notions.

Definition 3.1 *The vector field $(\theta, u) \mapsto g_\theta(u)$ in Assumption 2.3 is said to be Hamiltonian if there exists a bounded invertible linear map $J : X \rightarrow X$, skew-symmetric with respect to $\langle \cdot, \cdot \rangle_Z$, and a function $(\theta, u) \mapsto H_\theta(u)$ analytic in the sense of Definition 2.2, such that*

$$\forall(\theta, u, v) \in \mathbb{T} \times K \times X, \quad (\partial_u H_\theta)(u) v = (Jg_\theta(u), v)_Z. \quad (3.1)$$

A smooth map $(\theta, u) \mapsto \Phi_\theta(u)$ is said to be symplectic if

$$\forall(\theta, u, v, w) \in \mathbb{T} \times K \times X^2, \quad (J\partial_u \Phi_\theta(u)v, \partial_u \Phi_\theta(u)w)_Z = (Jv, w)_Z.$$

Remark 3.2 *Recall that this definition can be also written $g_\theta(u) = J^{-1}\nabla_u H_\theta(u)$, where the gradient is taken with respect to the scalar product $(\cdot, \cdot)_Z$ and is defined by*

$$\forall(\theta, u, v) \in \mathbb{T} \times K \times X^2, \quad (\nabla_u H_\theta(u), v)_Z = \partial_u H_\theta(u) v.$$

Moreover, the two following classical properties hold true. First, if $(\theta, u) \mapsto g_\theta(u)$ is Hamiltonian, then

$$\forall(u, v, w) \in K \times X^2 \quad (J\partial_u g_\theta(u)v, w)_Z = (v, J\partial_u g_\theta(u)w)_Z.$$

Second, if f_1 and f_2 are Hamiltonian, with Hamiltonian respectively given by F_1 and F_2 , then the Lie-Jacobi bracket

$$f(u) = [f_1(u), f_2(u)] = \partial_u f_1(u) f_2(u) - \partial_u f_2(u) f_1(u)$$

is also Hamiltonian, with Hamiltonian given by the Poisson bracket

$$F(u) = \{F_1, F_2\}(u) = (Jf_1(u), f_2(u))_Z.$$

Definition 3.3 An analytic vector field f (depending on ε) is said to be Hamiltonian up to an ε^{k+1} perturbation if there exists an analytic function F such that

$$\forall (u, v) \in K \times X, \quad (\partial_u F)(u) v = (Jf(u), v)_Z + \mathcal{O}(\varepsilon^{k+1} \|v\|_X).$$

A smooth map $(\theta, u) \mapsto \Phi_\theta(u)$ (depending on ε) is said to be symplectic up to an ε^{k+1} perturbation if

$$\forall (\theta, u, v, w) \in \mathbb{T} \times K \times X^2, \quad (J\partial_u \Phi_\theta(u)v, \partial_u \Phi_\theta(u)w)_Z = (Jv, w)_Z + \mathcal{O}(\varepsilon^{k+1} \|v\|_X \|w\|_X).$$

We now establish that whenever $(\theta, u) \mapsto g_\theta(u)$ is Hamiltonian, the associated averaged vector field $G^{[n]}$ obtained in Theorem 2.10 is Hamiltonian as well, up to an ε^{n+1} perturbation. Prior to that, we state the following lemma.

Lemma 3.4 Under the assumptions of Theorem 2.10 and provided $(\theta, u) \mapsto g_\theta(u)$ is Hamiltonian, suppose that $\Phi_\theta^{[n]}$ is symplectic up to an ε^{k+1} perturbation term, with $0 \leq k \leq n$, i.e. that for all $(\theta, u) \in \mathbb{T} \times K$ and $v, w \in X$

$$\left(J\partial_u \Phi_\theta^{[n]}(u)v, \partial_u \Phi_\theta^{[n]}(u)w \right)_Z = (Jv, w)_Z + \mathcal{O}(\varepsilon^{k+1} \|v\|_X \|w\|_X). \quad (\mathcal{S}_k)$$

Then $G^{[n]}$ is Hamiltonian up to an ε^{k+1} perturbation term. More precisely, we have, for all $u \in K$,

$$G^{[n]}(u) = J^{-1} \nabla_u H^{[n]}(u) + \mathcal{O}(\varepsilon^{k+1}), \quad (\mathcal{H}_k)$$

with Hamiltonian

$$H^{[n]}(u) = \left\langle H \circ \Phi^{[n+1]}(u) \right\rangle - \frac{1}{2\varepsilon} \left\langle \left(J\partial_\theta \Phi^{[n+1]}(u), \Phi^{[n+1]}(u) \right)_Z \right\rangle. \quad (3.2)$$

Proof of Lemma 3.4. We first compute $\partial_u H^{[n]}(u)$ when $u \in K$. We define for convenience

$$H_a^{[n]}(u) := \left\langle H \circ \Phi^{[n+1]}(u) \right\rangle \quad \text{and} \quad H_b^{[n]}(u) := -\frac{1}{2\varepsilon} \left\langle \left(J\partial_\theta \Phi^{[n+1]}(u), \Phi^{[n+1]}(u) \right)_Z \right\rangle,$$

so that $H^{[n]}(u) = H_a^{[n]}(u) + H_b^{[n]}(u)$. On the one hand, using that g_θ is Hamiltonian (according to Definition 3.1). For any $u \in K$ and $v \in X$, we recover

$$\begin{aligned} \partial_u H_a^{[n]}(u) v &= \left\langle \partial_u H_\theta \left(\Phi^{[n+1]}(u) \right) \left(\partial_u \Phi^{[n+1]}(u) \right) v \right\rangle = - \left\langle \left(g \left(\Phi^{[n+1]}(u) \right), J\partial_u \Phi^{[n+1]}(u) v \right)_Z \right\rangle. \end{aligned}$$

On the other hand, computing $\partial_u H_b^{[n]}$ and next using an integration by parts in θ , we have

$$\begin{aligned} \partial_u H_b^{[n]}(u) v &= -\frac{1}{2\varepsilon P} \int_{\mathbb{T}} \left(\left(J\partial_u \partial_\theta \Phi_\theta^{[n+1]}(u) v, \Phi_\theta^{[n+1]}(u) \right)_Z + \left(J\partial_\theta \Phi_\theta^{[n+1]}(u), \partial_u \Phi_\theta^{[n+1]}(u) v \right)_Z \right) d\theta \\ &= \frac{1}{2\varepsilon P} \int_{\mathbb{T}} \left(\left(J\partial_u \Phi_\theta^{[n+1]}(u) v, \partial_\theta \Phi_\theta^{[n+1]}(u) \right)_Z - \left(J\partial_\theta \Phi_\theta^{[n+1]}(u), \partial_u \Phi_\theta^{[n+1]}(u) v \right)_Z \right) d\theta \\ &= \frac{1}{\varepsilon P} \int_{\mathbb{T}} \left(\partial_\theta \Phi_\theta^{[n+1]}(u), J\partial_u \Phi_\theta^{[n+1]}(u) v \right)_Z d\theta. \end{aligned}$$

These results eventually provide the relation

$$\partial_u H^{[n]}(u) v = - \left\langle \left(g \circ \Phi^{[n+1]}(u) - \frac{1}{\varepsilon} \partial_\theta \Phi^{[n+1]}(u), J \partial_u \Phi^{[n+1]}(u) v \right)_Z \right\rangle. \quad (3.3)$$

The right-hand-side of (3.3) may now be simplified. From the very construction of $G^{[n+1]}$ and $\Phi^{[n+1]}$ we have

$$\partial_\theta \Phi_\theta^{[n+1]}(u) + \varepsilon \partial_u \Phi_\theta^{[n+1]}(u) \left(\partial_u \langle \Phi^{[n+1]} \rangle(u) \right)^{-1} \langle g \circ \Phi^{[n+1]} \rangle(u) - \varepsilon g_\theta \circ \Phi_\theta^{[n+1]}(u) = \mathcal{O}(\varepsilon^{n+2})$$

and $G^{[n+1]}(u) := \left(\partial_u \langle \Phi^{[n+1]} \rangle(u) \right)^{-1} \langle g \circ \Phi^{[n+1]} \rangle(u)$, for any $(\theta, u) \in \mathbb{T} \times K$. Here the term $\mathcal{O}(\varepsilon^{n+2})$ is meant in the sense of functions that are analytic in u , with value in X (or $X_{\mathbb{C}}$), see Theorem 2.10. This provides, by picking up $v \in X$ and taking the scalar product with $J \partial_u \Phi_\theta(u) v$, the relation

$$\begin{aligned} & \left(g_\theta \circ \Phi_\theta^{[n+1]}(u) - \frac{1}{\varepsilon} \partial_\theta \Phi_\theta^{[n+1]}(u), J \partial_u \Phi_\theta^{[n+1]}(u) v \right)_Z \\ &= \left(\partial_u \Phi_\theta^{[n+1]}(u) G^{[n+1]}(u), J \partial_u \Phi_\theta^{[n+1]}(u) v \right)_Z + \mathcal{O}(\varepsilon^{n+1} \|v\|_X). \end{aligned}$$

Hence, taking the average in θ , and using (3.3) yields in any circumstance

$$\partial_u H^{[n]}(u) v = - \left\langle \left(\partial_u \Phi^{[n+1]}(u) G^{[n+1]}(u), J \partial_u \Phi^{[n+1]}(u) v \right)_Z \right\rangle + \mathcal{O}(\varepsilon^{n+1} \|v\|_X). \quad (3.4)$$

This is where assumption (S_k) is used. It provides, using $k+1 \leq n+1$,

$$\left(\partial_u \Phi_\theta^{[n]}(u) G^{[n]}(u), J \partial_u \Phi_\theta^{[n]}(u) v \right)_Z = \left(G^{[n]}(u), J v \right)_Z + \mathcal{O}(\varepsilon^{k+1} \|v\|_X).$$

Inserting this identity in (3.4) and using the fact that $\Phi^{[n+1]} = \Phi^{[n]} + \mathcal{O}(\varepsilon^{n+1})$ and $G^{[n+1]} = G^{[n]} + \mathcal{O}(\varepsilon^{n+1})$ (see Theorem 2.13) gives

$$\partial_u H^{[n]}(u) v = \left(J G^{[n]}(u), v \right)_Z + \mathcal{O}(\varepsilon^{k+1} \|v\|_X).$$

The proof of the lemma is complete. \square

Lemma 3.4 allows to establish the

Theorem 3.5 [Stroboscopic averaging preserves the Hamiltonian structure]

Under the assumptions of Theorem 2.10 and assuming that g_θ is Hamiltonian, for all $n \in \mathbb{N}$, the functions $\Phi_\theta^{[n]}$ and $G^{[n]}$ are respectively symplectic and Hamiltonian up to ε^{n+1} -perturbation terms, namely for all $(\theta, u) \in \mathbb{T} \times K$, $v, w \in X$, we have (here $H^{[n]}$ is defined by (3.2)),

$$\left(J \partial_u \Phi_\theta^{[n]}(u) v, \partial_u \Phi_\theta^{[n]}(u) w \right)_Z = (J v, w)_Z + \mathcal{O}(\varepsilon^{n+1} \|v\|_X \|w\|_X), \quad (3.5)$$

$$G^{[n]}(u) = J^{-1} \nabla_u H^{[n]}(u) + \mathcal{O}(\varepsilon^{n+1}). \quad (3.6)$$

Remark 3.6 Consider the truncated Hamiltonian

$$\tilde{H}^{[n]}(u) = \sum_{k=0}^n \frac{\varepsilon^k}{k!} \left. \frac{d^k H^{[n]}}{d\varepsilon^k} \right|_{\varepsilon=0} (u). \quad (3.7)$$

A direct consequence of Theorem 3.5, and of the smoothness of $H^{[n]}$ and $G^{[n]}$ in ε , is that the truncated averaged vector field defined by (2.18) is exactly Hamiltonian, namely

$$\tilde{G}^{[n]}(u) = J^{-1} \nabla_u \tilde{H}^{[n]}(u).$$

(both functions are n -th order polynomials in ε and coincide to within $\mathcal{O}(\varepsilon^{n+1})$). This equality holds in the strong sense of X -valued functions.

Proof of Theorem 3.5. By construction, $\Phi_\theta^{[n]}(u) = u + \mathcal{O}(\varepsilon)$, so that (\mathcal{S}_0) and (\mathcal{H}_0) (as denoted in Lemma 3.4) hold. Now, assume that (\mathcal{S}_k) holds for some $0 \leq k \leq n-1$. By Lemma 3.4, we know that (\mathcal{H}_k) holds. Consider the flow $\Psi_t^{[n]}$ associated to the vector field $\varepsilon G^{[n]}$, which is defined for all $t \in [-P, P]$, at least for small ε :

$$\forall u \in K, \quad \partial_t \Psi_t^{[n]}(u) = \varepsilon G^{[n]} \circ \Psi_t^{[n]}(u), \quad \Psi_0^{[n]}(u) = u.$$

We claim that

$$\forall v, w \in X, \quad \left(J \partial_u \Psi_t^{[n]}(u) v, \partial_u \Psi_t^{[n]}(u) w \right)_Z = (Jv, w)_Z + \mathcal{O}(\varepsilon^{k+2} \|v\|_X \|w\|_X). \quad (3.8)$$

In order to prove (3.8), let us differentiate the left-hand side of this equation. By using the antisymmetry of J , we get

$$\begin{aligned} \frac{d}{dt} \left(J \partial_u \Psi_t^{[n]}(u) v, \partial_u \Psi_t^{[n]}(u) w \right)_Z &= -\varepsilon \left((\partial_u G^{[n]} \circ \Psi_t^{[n]}(u)) \partial_u \Psi_t^{[n]}(u) v, J \partial_u \Psi_t^{[n]}(u) w \right)_Z \\ &\quad + \varepsilon \left(J \partial_u \Psi_t^{[n]}(u) v, (\partial_u G^{[n]} \circ \Psi_t^{[n]}(u)) \partial_u \Psi_t^{[n]}(u) w \right)_Z \end{aligned}$$

Besides, by differentiating \mathcal{H}_k , we obtain, for all $v, w \in X$,

$$\left(\partial_u G^{[n]}(u) w, v \right)_Z = -\partial_u^2 H^{[n]}(u) (J^{-1} v, w) + \mathcal{O}(\varepsilon^{k+1} \|v\|_X \|w\|_X),$$

thus, by symmetry of $\partial_u^2 H^{[n]}$,

$$\begin{aligned} \frac{d}{dt} \left(J \partial_u \Psi_t^{[n]}(u) v, \partial_u \Psi_t^{[n]}(u) w \right)_Z &= \varepsilon \partial_u^2 H^{[n]} \circ \Psi_t^{[n]}(u) \left(\partial_u \Psi_t^{[n]} w, \partial_u \Psi_t^{[n]} v \right) \\ &\quad - \varepsilon \partial_u^2 H^{[n]} \circ \Psi_t^{[n]}(u) \left(\partial_u \Psi_t^{[n]} w, \partial_u \Psi_t^{[n]} v \right) + \mathcal{O}(\varepsilon^{k+2} \|v\|_X \|w\|_X) \\ &= \mathcal{O}(\varepsilon^{k+2} \|v\|_X \|w\|_X). \end{aligned}$$

Integrating this equation yields (3.8). Now, for all $(\theta, u) \in \mathbb{T} \times K$, denote $\chi_t(u) = \Phi_t^{[n]} \circ \Psi_t^{[n]}(u)$. By Theorem 2.10 and (3.1), we have

$$\begin{aligned} \partial_t \chi_t(u) &= (\partial_t \Phi_t^{[n]}) \circ \Psi_t^{[n]}(u) + \varepsilon (\partial_u \Phi_t^{[n]}) \circ \Psi_t^{[n]}(u) G^{[n]} \circ \Psi_t^{[n]}(u) \\ &= \varepsilon J^{-1} \nabla H_t(\chi_t(u)) + \mathcal{O}(\varepsilon^{n+1}), \end{aligned}$$

so that the map χ_t is quasi-symplectic, i.e., proceeding as for $\Psi_t^{[n]}$, we have

$$\forall v, w \in X, \quad (J\partial_u\chi_t(u)v, \partial_u\chi_t(u)w)_Z = (Jv, w)_Z + \mathcal{O}(\varepsilon^{n+1}\|v\|_X\|w\|_X). \quad (3.9)$$

Finally, since $\Psi_t^{[n]}$ is the flow of an autonomous equation, one has $\Phi_t^{[n]} = \chi_t \circ \Psi_{-t}^{[n]}$ so, from (3.8) and (3.9), one gets (\mathcal{S}_{k+1}) . An induction argument finishes the proof. \square

3.2 Preservation of the invariants

Assume that the solution of (2.1), associated with the field $(\theta, u) \mapsto \varepsilon g_\theta(u)$, admits an invariant. More precisely, assume that the smooth function $Q : \mathbb{T} \times X \rightarrow \mathbb{R}$, which possibly depends on ε , satisfies

$$Q_t(u^\varepsilon(t)) \equiv Q_0(u_0).$$

For instance, in the framework of Hilbert spaces $X \subset Z$ presented in Section 3.1 and considering the nonlinear Schrödinger equation, the quantity $Q(u) = \|u\|_Z^2$ is an invariant provided $(g_\theta(u), u)_Z = 0$ whenever $u \in Z$. The question is whether the averaged field $G^{[n]}$ possesses Q_θ as an (almost) invariant as well. It turns out that the answer is positive, and the proof is strikingly simple. The crucial fact is that $Q_\theta \circ \Phi_\theta$ is (almost) independent of θ , while Q_0 is an (almost) invariant of the averaged system, up to small perturbation terms.

Before going on, let us make the invariance assumption more precise. Differentiating the relation $Q_t(u^\varepsilon(t)) \equiv Q_0(u_0)$ provides

$$\partial_\theta Q_\theta(u) + \varepsilon \partial_u Q_\theta(u) g_\theta(u) = 0, \quad (3.10)$$

whenever $(\theta, u) = (t, u^\varepsilon(t))$. In the sequel, we shall require that this relation actually holds true for any $\theta \in \mathbb{T}$ and any $u \in K$.

Theorem 3.7 [Stroboscopic averaging preserves the invariants]

Under the assumptions of Theorem 2.10, assume that the function $(\theta, u) \in \mathbb{T} \times X \mapsto Q_\theta(u) \in \mathbb{R}$ is an invariant of the field εg_θ , in that (3.10) holds for any $(\theta, u) \in \mathbb{T} \times K$. Assume that $(\theta, u) \mapsto Q_\theta(u)$ is analytic on K_ρ for some $0 < \rho \leq R$.

Then, for all $n \in \mathbb{N}$, the change of variable $\Phi_\theta^{[n]}$ and the averaged vector field $G^{[n]}$ satisfy, whenever $u \in K$, $\theta \in \mathbb{T}$,

$$Q_\theta(\Phi_\theta^{[n]}(u)) = Q_0(u) + \mathcal{O}(\varepsilon^{n+1}), \quad \text{and} \quad (\partial_u Q_0)(u) G^{[n]}(u) = \mathcal{O}(\varepsilon^n). \quad (3.11)$$

In particular, we have $(d/dt)Q_0(\Psi_t^{[n]}(u_0)) = \mathcal{O}(\varepsilon^{n+1})$, whenever $t \in [0, T/\varepsilon]$.

Remark 3.8 *If the invariant Q_θ does not depend on ε , then, from $G^{[n+1]} - \tilde{G}^{[n]} = \mathcal{O}(\varepsilon^{n+1})$ and remarking that $\tilde{G}^{[n]}$ is a polynomial of degree n in ε , one deduces from (3.11) that we have*

$$(\partial_u Q_0)(u) \tilde{G}^{[n]}(u) = 0,$$

so that Q_0 is exactly preserved by the autonomous equation (2.22).

Proof of Theorem 3.7. Relation (2.24), written in the form

$$\partial_\theta \Phi_\theta^{[n]}(u) + \varepsilon \partial_u \Phi_\theta^{[n]}(u) G^{[n]}(u) = \varepsilon g_\theta \circ \Phi_\theta^{[n]}(u) + \mathcal{O}(\varepsilon^{n+1}),$$

provides after premultiplying by $(\partial_u Q_\theta) \circ \Phi_\theta^{[n]}$ and using the fact that Q is an invariant of εg_θ , the relation

$$\begin{aligned} & (\partial_u Q_\theta) \circ \Phi_\theta^{[n]}(u) \partial_\theta \Phi_\theta^{[n]}(u) + \varepsilon \partial_u (Q_\theta \circ \Phi_\theta^{[n]})(u) G^{[n]}(u) \\ &= \varepsilon \left(\partial_u Q_\theta \circ \Phi_\theta^{[n]} \right) (u) \left(g_\theta \circ \Phi_\theta^{[n]} \right) (u) + \mathcal{O}(\varepsilon^{n+1}) \\ &= -(\partial_\theta Q_\theta) \circ \Phi_\theta^{[n]}(u) + \mathcal{O}(\varepsilon^{n+1}), \end{aligned}$$

whenever $u \in K$. Note that the term $\mathcal{O}(\varepsilon^{n+1})$ is meant in the sense of analytic functions, which means in the $\|\cdot\|_{\rho'}$ norm, whenever $0 < \rho' < \rho$, say. Therefore, we arrive at

$$\partial_\theta \left(Q_\theta \circ \Phi_\theta^{[n]} \right) (u) + \varepsilon \partial_u \left(Q_\theta \circ \Phi_\theta^{[n]} \right) (u) G^{[n]}(u) = \mathcal{O}(\varepsilon^{n+1}). \quad (3.12)$$

In particular, taking averages on both sides yields

$$\partial_u \left\langle Q \circ \Phi^{[n]} \right\rangle (u) G^{[n]}(u) = \mathcal{O}(\varepsilon^n). \quad (3.13)$$

The Theorem now comes from an induction argument. Assume that, for some $k < n$, and for all $\theta \in \mathbb{T}$, we have $Q_\theta \circ \Phi_\theta^{[n]}(u) = Q_0(u) + \mathcal{O}(\varepsilon^{k+1})$. Note that this property is clearly true when $k = 0$ since $\Phi^{[n]} = \text{Id} + \mathcal{O}(\varepsilon)$ and since $Q_\theta(u) = Q_0(u) + \mathcal{O}(\varepsilon)$ by integrating (3.10). It comes $Q_\theta \circ \Phi_\theta^{[n]} = Q_0 + \mathcal{O}(\varepsilon^{k+1}) = \langle Q \circ \Phi^{[n]} \rangle + \mathcal{O}(\varepsilon^{k+1})$, hence after differentiation $\partial_u (Q_\theta \circ \Phi_\theta^{[n]}) = \partial_u \langle Q \circ \Phi^{[n]} \rangle + \mathcal{O}(\varepsilon^{k+1})$, and eventually we recover in (3.13)

$$\partial_u (Q_\theta \circ \Phi_\theta^{[n]})(u) G^{[n]}(u) = \mathcal{O}(\varepsilon^n + \varepsilon^{k+1}).$$

This provides in (3.12)

$$\partial_\theta (Q_\theta \circ \Phi_\theta^{[n]})(u) = \mathcal{O}(\varepsilon^{n+1} + \varepsilon^{k+2}),$$

which, by integration in θ , provides $Q_\theta \circ \Phi_\theta^{[n]} = Q_0 \circ \Phi_0^{[n]} + \mathcal{O}(\varepsilon^{k+2}) = Q_0 + \mathcal{O}(\varepsilon^{k+2})$, whenever $k < n$. The recursion is complete. \square

4 Application to the nonlinear Schrödinger equation

In this section, we apply the results of Sections 2 and 3 to the nonlinear Schrödinger (NLS) equation, written as

$$\begin{aligned} i\partial_t \psi^\varepsilon(t, x) &= (A\psi^\varepsilon)(t, x) + \varepsilon f(|\psi^\varepsilon(t, x)|^2) \psi^\varepsilon(t, x), \quad t \geq 0, \quad \psi^\varepsilon(t, \cdot) \in X, \\ \psi^\varepsilon(0, x) &= \psi_0(x) \in X. \end{aligned} \quad (4.1)$$

Here f is a real-analytic function. We set $Z = L^2(\Omega)$ where $\Omega \subset \mathbb{R}^d$ is open, and $A : D(A) \subset Z \rightarrow Z$ is a linear unbounded self-adjoint operator with dense domain $D(A) \subset Z$.

We assume A is of the form $A = -\Delta_x + V(x)$, for some potential $V(x)$. We also assume A is non-negative, which allows to take fractional powers of A thanks to the functional calculus for self-adjoint operators. Lastly, we set

$$X = \{u \in Z \text{ s.t. } (1 + A)^{s/2}u \in Z\},$$

for some $s \geq 1$ (to have a bounded energy, see (4.12) below), i.e. X is chosen in the Sobolev scale induced by A (with the obvious norm). We assume that $X \subset L^\infty(\Omega)$ continuously, and we also assume that X is an algebra (i.e. $\|uv\|_X \lesssim \|u\|_X \|v\|_X$ whenever u and v belong to X), to deal with the nonlinear term $f(|\psi|^2)\psi$.

The first key assumption is, we assume A has compact resolvent. This imposes both mild regularity assumptions on the potential $V(x)$, and (more importantly) compactness in the variable x (typically Ω is bounded, or $\Omega = \mathbb{R}^d$ with $V(x) \rightarrow +\infty$ as $|x| \rightarrow \infty$ to cut-off large values of x). Compactness of the resolvent of A ensures that the spectrum of A is discrete.

The second key assumption on A , and actually the most restrictive one, is

$$\text{the spectrum of } A \text{ is a subset of } \lambda\mathbb{N} \text{ for some } \lambda > 0. \quad (4.2)$$

In other words, while the Stone theorem ensures that the propagator $\exp(i\theta A)$ is well-defined as a strong group of unitary operators on Z whenever $\theta \in \mathbb{R}$ (this is due to the fact that A is self-adjoint), we are here assuming that $\theta \mapsto \exp(i\theta A)$ is *periodic* (with period $2\pi/\lambda$). In general, when A has compact resolvent, the function $\theta \mapsto \exp(i\theta A)$ is *almost-periodic* only, in that it entails an infinite, countable, number of independent frequencies.

A last, more technical, functional analytic assumption is in order, to deal with the nonlinear term $f(|\psi|^2)\psi$ in (4.1). Namely, we need a tame estimate, in that for any smooth and nonlinear function $G : \mathbb{C} \rightarrow \mathbb{C}$ satisfying $G(0) = 0$, there exists a nondecreasing C^1 function $C_G : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that, for all $u \in X$, we have

$$\|G(u)\|_X \leq C_G(\|u\|_{L^\infty}) \|u\|_X. \quad (4.3)$$

This statement completes and refines the assumed fact that X is an algebra.

Under all these assumptions, the local in time existence of strong solutions to (4.1), for any fixed value $\varepsilon > 0$, is standard, see for instance [Car08, CH98, Caz03], and it becomes feasible to deal with averaging issues in this equation.

Note that two paradigms are covered by our analysis.

Case 1: NLS on the d -dimensional torus. Let $\mathbb{T}_a^d = [0, a]^d$, with $a > 0$. We consider the equation

$$\begin{aligned} i\partial_t \psi^\varepsilon(t, x) &= -\Delta_x \psi^\varepsilon(t, x) + \varepsilon f(|\psi^\varepsilon(t, x)|^2) \psi^\varepsilon(t, x), & t \geq 0, \\ \psi^\varepsilon(0, x) &= \psi_0(x), & x \in \mathbb{T}_a^d, \end{aligned} \quad (4.4)$$

with periodic boundary conditions. In this case we set $A = -\Delta_x$ with domain $D(A) = \{u \in H^2(\mathbb{T}_a^d) \text{ s.t. } u|_{x_j=0} = u|_{x_j=a} \text{ for } j = 1, \dots, d\}$, where H^2 is the usual Sobolev space $\{u(x) \in L^2 \text{ s.t. } \Delta_x u(x) \in L^2\}$. The operator $A : D(A) \rightarrow Z = L^2(\mathbb{T}_a^d; \mathbb{C})$ is self-adjoint non-negative with compact resolvent, the embedding $D(A) \subset Z$ is dense, and the spectrum

of A is $\sigma(A) = \{(2\pi/a)^2|k|^2 = (2\pi/a)^2(k_1^2 + \dots + k_d^2); k \in \mathbb{Z}^d\} \subset (2\pi/a)^2\mathbb{N}$. We take $\psi_0 \in X$, where we set $X = H^s(\mathbb{T}_a^d; \mathbb{C})$, for some $s > d/2$, and H^s is the Sobolev space of periodic functions associated with the norm $\|\cdot\|_{H^s}$ defined by

$$\|u\|_{H^s}^2 = \|(1 - \Delta)^{s/2}u\|_{L^2}^2 = a^d \sum_{k \in \mathbb{Z}^d} (1 + (2\pi/a)^2|k|^2)^s |u_k|^2, \quad (4.5)$$

$$\text{where we write } u(x) = \sum_{k \in \mathbb{Z}^d} u_k e^{i(2\pi/a)k \cdot x} \text{ whenever } u \in L^2.$$

In that context, the tame estimate (4.3) is clear, since the assumption $s > d/2$ immediately ensures, for any smooth $G : \mathbb{C} \rightarrow \mathbb{C}$ with $G(0) = 0$, the estimate

$$\|G(u)\|_{H^s} \leq C_G (\|u\|_{L^\infty}) \|u\|_{H^s}, \quad (4.6)$$

for some nondecreasing function $C_G : \mathbb{R}^+ \rightarrow \mathbb{R}^+$. The constraint $s > d/2$ also ensures that X is an algebra. \square

Case 2: the Gross-Pitaevskii equation. Take $\omega > 0$. We consider the equation

$$\begin{aligned} i\partial_t \psi^\varepsilon(t, x) &= (-\Delta_x + \omega^2|x|^2 - d\omega) \psi^\varepsilon(t, x) + \varepsilon f(|\psi^\varepsilon(t, x)|^2) \psi^\varepsilon(t, x), & t \geq 0, \\ \psi^\varepsilon(0, x) &= \psi_0(x). & x \in \mathbb{R}^d. \end{aligned} \quad (4.7)$$

In this case we set $A = -\Delta + \omega^2|x|^2 - d\omega$, with domain $D(A) = \{u \in L^2(\mathbb{R}^d) \text{ s.t. } Au \in L^2(\mathbb{R}^d)\} = \{u \in H^2(\mathbb{R}^d) : |x|^2 u \in L^2(\mathbb{R}^d)\}$. The operator $A : D(A) \rightarrow Z = L^2(\mathbb{R}^d; \mathbb{C})$ is self-adjoint non-negative with compact resolvent, the embedding $D(A) \subset Z$ is dense, and the spectrum of A is $\sigma(A) = \{2k\omega; k \in \mathbb{N}\} = 2\omega\mathbb{N}$. We take $\psi_0 \in X$ where we set $X = \Sigma^s(\mathbb{R}^d; \mathbb{C})$, for some $s > d/2$, and Σ^s is the space $\{u \in H^s(\mathbb{R}^d) \text{ s.t. } |x|^s u \in L^2(\mathbb{R}^d)\}$, associated with the norm

$$\|u\|_{\Sigma^s} = \|(1 - \Delta + \omega^2|x|^2 - d\omega)^{s/2}u\|_{L^2}. \quad (4.8)$$

The following crucial equivalence of norms holds (see e.g. [BACM08])

$$\|u\|_{\Sigma^s} \sim \|u\|_{H^s} + \||x|^s u\|_{L^2}. \quad (4.9)$$

This ingredient immediately provide the desired tame estimate (4.3) in the present context. Indeed, taking G as in (4.3) we have

$$\begin{aligned} \|G(u)\|_X &\leq C (\|G(u)\|_{H^s} + \||x|^s G(u)\|_{L^2}) && \text{(for some constant } C > 0) \\ &\leq C \left(C_G (\|u\|_{L^\infty}) \|u\|_{H^s} + \left(\max_{|u| \leq \|u\|_{L^\infty}} |G'(u)| \right) \||x|^s u\|_{L^2} \right) && \text{(we use (4.6))} \\ &\leq C \left(C_G (\|u\|_{L^\infty}) + \max_{|u| \leq \|u\|_{L^\infty}} |G'(u)| \right) \|u\|_X. \end{aligned}$$

The same argument shows that X is an algebra. \square

The well-posedness of the Cauchy problem for (4.1) can be formulated as follows.

Proposition 4.1 *Let $\psi_0 \in X$. Then, for all $\varepsilon > 0$, there exists $T_{\max}^\varepsilon \in]0, +\infty]$ and a unique maximal solution $\psi^\varepsilon \in C([0, T_{\max}^\varepsilon[, X)$ to (4.1). This solution is maximal in the sense that*

$$\text{if } T_{\max}^\varepsilon < +\infty, \text{ then } \limsup_{t \rightarrow T_{\max}^\varepsilon} \|\psi^\varepsilon(t)\|_{L^\infty} = +\infty.$$

Moreover, for all $\kappa > 1$, there exists $T_\kappa > 0$ such that for any $\varepsilon > 0$ we have $T_{\max}^\varepsilon > T_\kappa/\varepsilon$, and

$$\forall t \in [0, T_\kappa/\varepsilon], \quad \|\psi^\varepsilon(t)\|_X \leq \kappa \|\psi_0\|_X. \quad (4.10)$$

Furthermore, one has the following conservation laws, valid whenever $t < T_{\max}^\varepsilon$,

$$\|\psi^\varepsilon(t)\|_{L^2}^2 = \|\psi_0\|_{L^2}^2 \quad (\text{conservation of mass}), \quad (4.11)$$

$$\begin{aligned} \frac{1}{2} (A\psi^\varepsilon(t), \psi^\varepsilon(t))_{L^2} + \frac{\varepsilon}{2} \int F(|\psi^\varepsilon|^2)(t, x) dx \\ = \frac{1}{2} (A\psi_0, \psi_0)_{L^2} + \frac{\varepsilon}{2} \int F(|\psi_0|^2) dx, \end{aligned} \quad (\text{conservation of energy}), \quad (4.12)$$

where $F(u) = \int_0^u f(v) dv$.

Proof of Proposition 4.1. We refer to [Car08] for the proof of the existence and uniqueness result for fixed $\varepsilon > 0$, as well as the proof of the standard relations (4.11) and (4.12). We only prove here the a priori estimate (4.10). Without loss of generality, we assume $\psi_0 \neq 0$. Let $\kappa > 1$ and define

$$T_\kappa^\varepsilon = \varepsilon \max \{ T < T_{\max}^\varepsilon \text{ s.t. } \forall t \in [0, T], \|\psi^\varepsilon(t)\|_X \leq \kappa \|\psi_0\|_X \}.$$

Let us prove the existence of T_κ independent of ε such that $T_\kappa^\varepsilon \geq T_\kappa$. The Duhamel formulation of (4.1) provides

$$\psi^\varepsilon(t) = e^{-itA} \psi_0 - i\varepsilon \int_0^t e^{-i(t-\tau)A} f(|\psi^\varepsilon(\tau)|^2) \psi^\varepsilon(\tau) d\tau.$$

Hence, using the fact that $\exp(-itA)$ is unitary on X , which comes from the definition of the space X and its associated norm, and from the unitarity of $\exp(-itA)$ on Z , we have

$$\begin{aligned} \|\psi^\varepsilon(t)\|_X &\leq \|\psi_0\|_X + \varepsilon \int_0^t \|f(|\psi^\varepsilon(\tau)|^2) \psi^\varepsilon(\tau)\|_X d\tau \\ &\leq \|\psi_0\|_X + \varepsilon \int_0^t C_{\tilde{f}}(\|\psi^\varepsilon(\tau)\|_{L^\infty}) \|\psi^\varepsilon(\tau)\|_X d\tau \\ &\leq \|\psi_0\|_X + \varepsilon \int_0^t C_{\tilde{f}}(c_0 \|\psi^\varepsilon(\tau)\|_X) \|\psi^\varepsilon(\tau)\|_X d\tau, \end{aligned}$$

where $C_{\tilde{f}}$ is the nondecreasing function in (4.3) associated to $\tilde{f}(u) = f(|u|^2)u$ and c_0 is the norm of the continuous embedding $X \subset L^\infty$. As a consequence, whenever $t \leq T_\kappa^\varepsilon/\varepsilon$ we recover

$$\|\psi^\varepsilon(t)\|_X \leq \|\psi_0\|_X + \varepsilon C_{\tilde{f}}(c_0 \kappa \|\psi_0\|_X) \int_0^t \|\psi^\varepsilon(\tau)\|_X d\tau,$$

and the Gronwall lemma asserts, whenever $0 \leq t \leq T_\kappa^\varepsilon/\varepsilon$, the estimate

$$\|\psi^\varepsilon(t)\|_X \leq \|\psi_0\|_X e^{T_\kappa^\varepsilon C_{\tilde{f}}(c_0\kappa\|\psi_0\|_X)}.$$

This in turn ensures

$$T_\kappa^\varepsilon \geq T_\kappa := \frac{\log \kappa}{C_{\tilde{f}}(c_0\kappa\|\psi_0\|_X)}. \quad (4.13)$$

□

Let us now put the NLS equation (4.1) under the form (2.1). To this aim, we first pass to canonical coordinates setting $q^\varepsilon = \Re(\psi^\varepsilon)$, $p^\varepsilon = \Im(\psi^\varepsilon)$. The unknown $y^\varepsilon(t) = \begin{pmatrix} p^\varepsilon(t) \\ q^\varepsilon(t) \end{pmatrix} \in Z \times Z$ satisfies

$$\partial_t y^\varepsilon(t) = J^{-1} A y^\varepsilon(t) + \varepsilon f(y^\varepsilon(t)^2) J^{-1} y^\varepsilon(t), \quad y^\varepsilon(0) = u_0 := \begin{pmatrix} \Im(\psi_0) \\ \Re(\psi_0) \end{pmatrix}, \quad (4.14)$$

where we have denoted $y^\varepsilon(t)^2 := p^\varepsilon(t)^2 + q^\varepsilon(t)^2$ and $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. We also make a slight abuse of notation, in that we denote $Ay^\varepsilon \equiv \begin{pmatrix} A p^\varepsilon \\ A q^\varepsilon \end{pmatrix}$, which makes the operator A self-adjoint on $Z = L^2 \times L^2$. In the same spirit, we denote in the sequel $\|y^\varepsilon(t)\|_Z^2 \equiv \|p^\varepsilon(t)\|_Z^2 + \|q^\varepsilon(t)\|_Z^2$, and similarly for the X -norm. An obvious computation shows

$$e^{\theta J^{-1} A} = \begin{pmatrix} \cos(\theta A) & -\sin(\theta A) \\ \sin(\theta A) & \cos(\theta A) \end{pmatrix},$$

hence $e^{\theta J^{-1} A}$ is a group of isometries on $Z \times Z$ and on $X \times X$. Moreover, assumption (4.2) shows $\theta \rightarrow e^{\theta J^{-1} A}$ is periodic (with period $2\pi/\lambda$). Consider now the function

$$u^\varepsilon(t) = e^{-tJ^{-1}A} y^\varepsilon(t). \quad (4.15)$$

Inserting (4.15) in (4.14) immediately yields

$$\partial_t u^\varepsilon(t) = \varepsilon g_t(u^\varepsilon(t)), \quad u^\varepsilon(0) = u_0, \quad (4.16)$$

provided we define, whenever $u \in X \times X$,

$$g_\theta(u) := J^{-1} e^{-\theta J^{-1} A} f\left((e^{\theta J^{-1} A} u)^2\right) e^{\theta J^{-1} A} u. \quad (4.17)$$

As desired, equation (4.16) is of the form (2.1) considered in the previous paragraphs. In order to apply the results we obtained in the previous sections, there remains to check that the nonlinear function g_θ acting on the Banach space $X \times X$ satisfies Assumptions 2.1 (well-posedness of the problem on a fixed times interval of size $\mathcal{O}(1/\varepsilon)$) and 2.3 (analyticity of g_θ). We also prove that g_θ is Hamiltonian, in the sense of Definition 3.1.

Assumption 2.1 holds true. Take any $\kappa > 1$ and any $\psi_0 \in X$, $\psi_0 \neq 0$. The precise constraints on these two choices are made precise later. Proposition 4.1 and the fact that $\|u^\varepsilon(t)\|_X =$

$\|y^\varepsilon(t)\|_X$ imply in any circumstance that $\|u^\varepsilon(t)\|_X \leq \kappa\|\psi_0\|_X$ whenever $t \leq T_\kappa/\varepsilon$, hence Assumption 2.1 holds true, where $\varepsilon^* > 0$ can be chosen arbitrarily and we may take $T = T_\kappa$ as well as (recall that $\kappa - 1 > 0$)

$$K = \{u \in X \times X : \|u\|_X < (2\kappa - 1)M\} \quad (\text{where } \|\psi_0\|_X = M). \quad (4.18)$$

□

Assumption 2.3 holds true. Periodicity of g_θ in θ is obvious, thanks to assumption (4.2). Now, take $R_0 < R_a$, where R_a is the radius of analyticity of the function f and denote $\tilde{f}(z) = f(z^2)z$. The function \tilde{f} is clearly analytic and bounded for $|z| < \sqrt{R_0}$. On the other hand, take c_0 as the norm of the (assumed) continuous embedding $X \subset L^\infty$. We clearly have

$$\forall u \in X_{\mathbb{C}}, \quad \|u\|_{L^\infty} \leq c_0\|u\|_{X_{\mathbb{C}}}.$$

With these observations in mind, we choose the parameters $M = \|\psi_0\|_X > 0$, $R > 0$, and $\kappa > 1$, such that

$$M < \frac{\sqrt{R_0}}{c_0}, \quad \text{and} \quad (2\kappa - 1)M + 2R \leq \frac{\sqrt{R_0}}{c_0}. \quad (4.19)$$

For all $u \in K_{2R}$, where K is given by (4.18), one has clearly $\|u\|_{L^\infty} \leq c_0\|u\|_{X_{\mathbb{C}}} < \sqrt{R_0}$, so that for all $x \in \Omega$, the quantity $u(x)$ belongs to the domain of analyticity of \tilde{f} . Hence one may write, for any function $h \in X$ such that $\|h\|_{X_{\mathbb{C}}}$ (hence $\|h\|_{L^\infty}$) is small enough, the relation

$$\tilde{f}(u(x) + h(x)) - \tilde{f}(u(x)) - \partial_u \tilde{f}(u(x)) h(x) = \int_0^1 (1-t) \partial_u^2 \tilde{f}(u(x) + th(x)) h(x)^2 dt,$$

and the assumed fact that X is an algebra, together with the fact that \tilde{f} is here computed in a fixed subset of its domain of analyticity, allows to upper bound the right-hand-side as an $o(\|h\|_{X_{\mathbb{C}}})$. Hence the function $u \mapsto \tilde{f}(u)$ is analytic on K_{2R} in the sense of Definition 2.2. Moreover, we remark that for all $\theta \in \mathbb{T}$ and $u \in K_{2R}$,

$$\left\| e^{\theta J^{-1}A} u \right\|_{L^\infty_{\mathbb{C}}} \leq c_0 \left\| e^{\theta J^{-1}A} u \right\|_{X_{\mathbb{C}}} = c_0 \|u\|_{X_{\mathbb{C}}} < \sqrt{R_0}.$$

Hence, by standard composition theorems, the function g_θ defined by (4.17) is analytic on K_{2R} (we use that $e^{\theta J^{-1}A}$ and J^{-1} are bounded and linear on $X_{\mathbb{C}}$). Finally, we have, for all $\theta \in \mathbb{T}$, $u \in K_{2R}$,

$$\begin{aligned} \|g_\theta(u)\|_{X_{\mathbb{C}}} &= \|J^{-1} \circ e^{-\theta J^{-1}A} \circ \tilde{f} \circ e^{\theta J^{-1}A}(u)\|_{X_{\mathbb{C}}} \\ &= \|\tilde{f} \circ e^{\theta J^{-1}A}(u)\|_{X_{\mathbb{C}}} \leq C_K := \max_{|z| \leq \sqrt{R_0}} |\tilde{f}(z)|. \end{aligned}$$

We have proved that g_θ satisfies Assumption 2.3. □

The vector field g_θ is Hamiltonian. If J is the above defined matrix and if F is defined as $F(u) = \int_0^u f(v)dv$, it is clear that J is skew-symmetric with respect to the scalar product on $Z \times Z$, and one can check the identity

$$g_\theta(u) = J^{-1} \nabla_u H_\theta(u), \quad \text{where } H_\theta(u) := \frac{1}{2} \int F \left((e^{\theta J^{-1}A} u)^2 \right) (x) dx. \quad (4.20)$$

In other words, g_θ is Hamiltonian in the sense of Definition 3.1. □

As a consequence of all these considerations, the results of Sections 2 and 3 can be applied in this context. We summarize these results in the following Theorem. For simplicity, we identify the initial complex-valued function $\psi^\varepsilon(t)$ and the function $y^\varepsilon(t)$. Similarly, we also identify $u^\varepsilon(t) = e^{tJ^{-1}A}y^\varepsilon(t)$ and $e^{-itA}\psi^\varepsilon(t)$.

Theorem 4.2 *Let M, κ, R and $T := T_\kappa$ be chosen as above. There exists $\varepsilon_0 > 0$, a function $G^\varepsilon(u)$, analytic in $u \in X$, and a function $\Phi_\theta^\varepsilon(u)$, analytic in $u \in X$, continuously differentiable and periodic in $\theta \in \mathbb{T}$ with $\Phi_0^\varepsilon = \text{id}$, such that the following holds.*

For all $\psi_0 \in X$ such that $\|\psi_0\|_X \leq M$, the unique maximal solution to (4.1) given by Proposition 4.1 satisfies

$$\sup_{0 \leq t \leq \frac{T}{\varepsilon}} \left\| \psi^\varepsilon(t) - e^{-itA} \Phi_t^\varepsilon \left(\tilde{\psi}^\varepsilon(t) \right) \right\|_X \leq C \exp \left(-\frac{C^{-1}}{\varepsilon} \right), \quad (4.21)$$

for some $C > 0$ independent of ε and of ψ_0 , while $\tilde{\psi}^\varepsilon \in C^1([0, T/\varepsilon], X)$ solves the autonomous equation

$$\frac{\partial \tilde{\psi}^\varepsilon}{\partial t} = \varepsilon G^\varepsilon(\tilde{\psi}^\varepsilon), \quad \tilde{\psi}^\varepsilon(0) = \psi_0. \quad (4.22)$$

Moreover, the autonomous equation (4.22) is Hamiltonian, i.e. there exists a real-analytic function $H^\varepsilon(u)$ such that $G^\varepsilon(u) = J^{-1} \nabla H^\varepsilon(u)$.

Lastly, the following two conservation laws are satisfied. For all $t \leq T/\varepsilon$, we have the exact conservation of mass

$$\|\tilde{\psi}(t)\|_{L^2}^2 = \|\psi_0\|_{L^2}^2 \quad (4.23)$$

and the almost conservation of energy

$$\begin{aligned} & \frac{1}{2} \left(A \tilde{\psi}^\varepsilon(t), \tilde{\psi}^\varepsilon(t) \right)_{L^2} + \frac{\varepsilon}{2} \int F \left(|\tilde{\psi}^\varepsilon|^2 \right) (t, x) dx \\ &= \frac{1}{2} (A \psi_0, \psi_0)_{L^2} + \frac{\varepsilon}{2} \int F \left(|\psi_0|^2 \right) dx + \mathcal{O} \left(e^{-C/\varepsilon} \right). \end{aligned} \quad (4.24)$$

Proof of Theorem 4.2. Due to the fact that the original Schrödinger equation may be put under the form $\partial_t u^\varepsilon(t) = \varepsilon g_t(u^\varepsilon(t))$ as in (4.16), this Theorem is a direct consequence of Theorem 2.7, of Theorem 3.5 and of Theorem 3.7, provided we set $G^\varepsilon = \tilde{G}^{[n_\varepsilon]}$ and $\Phi_\theta^\varepsilon = \Phi_\theta^{[n_\varepsilon]}$ (and n_ε is as in Theorem 2.7). In that context, the Hamiltonian H^ε is the function $\tilde{H}^{[n_\varepsilon]}$ given according to Remark 3.6. Lastly, the original equation (4.16) preserves the mass $\|u\|_{L^2}^2$ as well as the filtered energy

$$\begin{aligned} Q_\theta(u) &:= \frac{1}{2} (Au, u)_{L^2} + \frac{\varepsilon}{2} \int F \left((e^{\theta J^{-1}A} u)^2 \right) (x) dx \\ &= \frac{1}{2} (Au, u)_{L^2} + \varepsilon H_\theta(u). \end{aligned} \quad (4.25)$$

Hence the exact mass conservation (4.23) is a consequence of Theorem 3.7 and of Remark 3.8, since the mass invariant does not depend on ε . Besides, the almost conservation of energy (4.24) is a consequence of Theorem 3.7, recalling that the $\mathcal{O}(\varepsilon^n)$ in this Theorem naturally becomes an $\mathcal{O}(e^{-C/\varepsilon})$ given the optimal choice $n = n_\varepsilon$ of the integer n . \square

Remark 4.3 Surprisingly enough, one also deduces from this result a new almost invariant for the initial problem (4.1). Indeed, the invariance of Q_θ under the autonomous evolution equation (4.22), when written in the form $\partial_u Q_0(u) G^\varepsilon(u) = \mathcal{O}(e^{-C/\varepsilon})$ (see (3.11)), provides $\partial_u Q_0(u) J^{-1} \nabla_u H^\varepsilon(u) = \mathcal{O}(e^{-C/\varepsilon})$. The point is, Q_0 coincides with the Hamiltonian of the original Schrödinger equation (4.1). Hence, reading the above almost invariance in the reverse order $\partial_u H^\varepsilon(u) J^{-1} \nabla_u Q_0^\varepsilon(u) = \mathcal{O}(e^{-C/\varepsilon})$ immediately provides

$$\forall t \leq T/\varepsilon, \quad H^\varepsilon(\psi^\varepsilon(t)) = H^\varepsilon(\psi_0) + \mathcal{O}\left(e^{-C/\varepsilon}\right).$$

Remark 4.4 In dimension $d = 1$, we consider the Schrödinger equation (4.1) with initial datum $\psi_0 \in X = \{u \in Z \text{ s.t. } (1 + A)^{1/2} u \in Z\}$. In other words, we choose $s = 1$ in the above notation. In that case, it is known that for ε small enough, the solution of (4.1) is global in time and uniformly bounded in X . Therefore Theorem 2.7-part (iii) applies and the estimates of Theorem 4.2 hold true on longer time intervals $[0, \frac{T}{\varepsilon^{1+\alpha}}]$, for any $\alpha \in]0, 1[$.

Note in passing that the global existence of $\psi^\varepsilon(t)$ in that case comes from the following simple argument. Consider a pair $M > 0, \kappa > 1$. We claim that there exists $\varepsilon(M, \kappa)$ such that, if $\|\psi_0\|_X \leq M$ and $\varepsilon \leq \varepsilon(M, \kappa)$, then the solution is global, with $\|\psi^\varepsilon(t)\|_X \leq \kappa M$ for all $t \geq 0$. To prove the claim, denote

$$C_{\kappa, M} = \max_{\|u\|_X \leq 2\kappa M} \int |F(|u|^2)| dx < \infty \quad \text{and} \quad \varepsilon(M, \kappa) = \frac{(\kappa^2 - 1)M^2}{2C_{\kappa, M}}.$$

Here we used the embedding $X \subset L^\infty$ to have $C_{\kappa, M} < +\infty$. Recalling that $\|\psi^\varepsilon\|_X^2 = (\psi^\varepsilon, \psi^\varepsilon)_{L^2} + (A\psi^\varepsilon, \psi^\varepsilon)_{L^2}$ in the present case ($s = 1$), the conservation laws (4.11) and (4.12) give, as long as $\|\psi^\varepsilon(t)\|_X \leq 2\kappa M$ and provided $\varepsilon \leq \varepsilon(M, \kappa)$, the estimate

$$\|\psi^\varepsilon(t)\|_X^2 \leq \|\psi_0\|_X^2 + 2\varepsilon C_{\kappa, M} \leq M^2 + 2\varepsilon C_{\kappa, M} \leq \kappa^2 M^2.$$

This immediately implies that the solution $\psi^\varepsilon(t) \in X$ exists for all times.

5 Computing explicitly the averaged vector field

5.1 A formal Magnus expansion for non-linear time-dependent vector fields

In this subsection, we very briefly expose the form of Magnus expansion and a way to compute it. We refer the further interested reader to [BCOR10] and to the Appendix section of this paper for an elementary proof. Prior to stating the result, we introduce the following notation: Given two time-dependent vector fields $(t, u) \mapsto f_t(u)$ and $(t, u) \mapsto g_t(u)$ we write $h = f \triangleleft g$ for the vector field defined by

$$h_t(u) = \int_0^t [f_\tau(u), g_t(u)] d\tau = \int_0^t \left(\frac{\partial f_\tau(u)}{\partial u} g_t(u) - \frac{\partial g_t(u)}{\partial u} f_\tau(u) \right) d\tau. \quad (5.1)$$

Theorem 5.1 (Magnus) *Given the differential equation (2.1), there exists a formal vector field*

$$\varepsilon G^\varepsilon(u) = \varepsilon G_1(u) + \varepsilon^2 G_2(u) + \varepsilon^3 G_3(u) + \dots \quad (5.2)$$

such that, for arbitrary initial value u_0 , the solution $u^\varepsilon(t)$ of (2.1) and the (formal) solution $U(t)$ of the initial value problem

$$\frac{d}{dt}U = \varepsilon G^\varepsilon(U), \quad U(0) = u_0, \quad (5.3)$$

(formally) coincide at $t = P$. The terms G_i in (5.2) are given by

$$G_i(u) = \frac{1}{P} \int_0^P R_t^{[i]}(u) dt, \quad (5.4)$$

where the time-dependent vector fields $R_t^{[i]}(u)$ are uniquely determined by requiring that

$$\varepsilon R_t^\varepsilon(u) = \varepsilon R_t^{[1]}(u) + \varepsilon^2 R_t^{[2]}(u) + \varepsilon^3 R_t^{[3]}(u) + \dots \quad (5.5)$$

satisfies the equation

$$\varepsilon g = \varepsilon R^\varepsilon + \frac{\varepsilon^2}{2} R^\varepsilon \triangleleft R^\varepsilon + \frac{\varepsilon^3}{3!} R^\varepsilon \triangleleft (R^\varepsilon \triangleleft R^\varepsilon) + \frac{\varepsilon^4}{4!} R^\varepsilon \triangleleft (R^\varepsilon \triangleleft (R^\varepsilon \triangleleft R^\varepsilon)) + \dots \quad (5.6)$$

It is worth noticing that the use of the same notations for G^ε (as compared to Subsection 2.2) and for the corresponding terms of its expansion, G_1 , G_2 and so on, is no coincidence. As a matter of fact, denoting $\chi_{0,t}^\varepsilon$ the solution operator of (2.1) (i.e. writing as $\chi_{0,t}^\varepsilon(u_0)$ the solution $u^\varepsilon(t)$ of (2.1) with initial condition $u^\varepsilon(0) = u_0$), it is possible to define the change of variables Φ_t^ε of Subsection 2.2 as follows:

$$\Phi_t^\varepsilon(U) = (\chi_{0,t}^\varepsilon \circ \Psi_{-t}^\varepsilon)(U).$$

Now, as soon as $(t, u) \mapsto g_t(u)$ is P -periodic in t , one has for all u_0 and all $n \in \mathbb{Z}$ the relation $\Psi_{nP}^\varepsilon(u_0) = \chi_{0,nP}^\varepsilon(u_0)$, and this implies that $(t, U) \mapsto \Phi_t^\varepsilon(U)$ is P -periodic in t . In conclusion, this shows that the vector field G^ε obtained by Magnus expansion coincides (at least formally) with the vector field obtained in Subsection 2.2 and further, in a form which is more amenable to geometric considerations ⁴.

To conclude this subsection, we now give explicitly the first terms of G^ε :

$$\begin{aligned} R^{[1]} &= g, & G_1 &= \langle g \rangle, \\ R^{[2]} &= -\frac{1}{2} g \triangleleft g, & G_2 &= -\frac{1}{2} \left\langle \int_0^t [g_\tau(u), g_t(u)] d\tau \right\rangle, \\ R^{[3]} &= \frac{1}{12} g \triangleleft (g \triangleleft g) + \frac{1}{4} (g \triangleleft g) \triangleleft g, & \dots & \end{aligned}$$

5.2 Expressions of the averaged vector fields in terms of Fourier coefficients

We first note that whenever the system is Hamiltonian, i.e. whenever $g_\theta(u) = J^{-1} \nabla_u H_\theta(u)$, then, in accordance with Theorem 3.5, G_1 and G_2 are also Hamiltonian with Hamiltonians H_1 and H_2 given by

$$\begin{aligned} H_1(u) &= \frac{1}{P} \int_0^P H_s(u) ds, \\ H_2(u) &= -\frac{1}{2P} \int_0^P \int_0^{s_2} \{H_{s_1}(u), H_{s_2}(u)\} ds_1 ds_2. \end{aligned} \quad (5.7)$$

⁴Though as already mentioned in introduction, stucked to the one-frequency situation.

where $\{\cdot, \cdot\}$ denotes the Lie-Poisson bracket on X . Later in this section, we consider the nonlinear Schrödinger equation (4.1) with $f(z) = z^p$, $A = -\Delta_x$ (See Section 4) and the additional multiplicative term $\gamma(x)$

$$i\partial_t \psi^\varepsilon(t, x) = -\Delta_x \psi^\varepsilon(t, x) + \varepsilon \gamma(x) |\psi^\varepsilon(t, x)|^{2p} \psi^\varepsilon(t, x), \quad x \in [0, a]^d, \quad t \geq 0,$$

and write its solution in terms of the Fourier coefficients $\xi_k(t)$

$$\psi^\varepsilon(t, x) = \sum_{k \in \mathbb{Z}^d} e^{i\omega k \cdot x} \xi_k(t) \text{ where } \omega = \frac{2\pi}{a}.$$

In order to conveniently and concisely write down various terms occurring in the sequel, we first introduce, for $\mu \in \mathbb{N}^*$, *super* multi-indices $\mathbf{K} = (\mathbf{K}_1, \dots, \mathbf{K}_\mu) \in (\mathbb{Z}^d)^\mu$ and denote

- $\sigma_1(\mathbf{K}) = \sum_{j=1}^\mu (-1)^{j+1} \mathbf{K}_j \in \mathbb{Z}^d$,
- $\sigma_2(\mathbf{K}) = \sum_{j=1}^\mu (-1)^{j+1} |\mathbf{K}_j|^2 \in \mathbb{Z}$,
- $(uv)_{\mathbf{K}} = u_{K_1} v_{K_2} u_{K_3} \cdots v_{K_\mu}$ for even μ and $(uv)_{\mathbf{K}} = u_{K_1} v_{K_2} u_{K_3} \cdots u_{K_\mu}$ for odd μ ,
- $(vu)_{\mathbf{K}} = v_{K_1} u_{K_2} v_{K_3} \cdots u_{K_\mu}$ for even μ and $(vu)_{\mathbf{K}} = v_{K_1} u_{K_2} v_{K_3} \cdots v_{K_\mu}$ for odd μ .

Equation (4.1) is then formally equivalent to the infinite dimensional system

$$\forall n \in \mathbb{Z}^d, \quad i\dot{\xi}_n(t) = \omega^2 |n|^2 \xi_n(t) + \varepsilon \sum_{\mathbf{K} \in (\mathbb{Z}^d)^{2p+1}} \gamma_{(\sigma_1(\mathbf{K})-n)} (\xi \bar{\xi})_{\mathbf{K}}(t)$$

where we have defined

$$\forall k \in \mathbb{Z}^d, \gamma_k = \frac{1}{a^d} \int_{\mathbb{T}_a^d} \gamma(x) e^{i\omega k \cdot x} dx.$$

This system which can be put in Hamiltonian form by introducing the conjugate variables $\eta_n(t) = \bar{\xi}_n(t)$:

$$\forall n \in \mathbb{Z}^d, \quad \begin{cases} \dot{\xi}_n(t) &= -i \frac{\partial H}{\partial \eta_n}(\xi, \eta) \\ \dot{\eta}_n(t) &= i \frac{\partial H}{\partial \xi_n}(\xi, \eta) \end{cases},$$

with

$$H(\xi, \eta) = \omega^2 \sum_{k \in \mathbb{Z}^d} |k|^2 \xi_k \eta_k + \frac{\varepsilon}{p+1} \sum_{\mathbf{K} \in (\mathbb{Z}^d)^{2p+2}} \gamma_{\sigma_1(\mathbf{K})} (\xi \eta)_{\mathbf{K}}.$$

Now, filtering out the main oscillations by changing the variables $\xi_n(t) = e^{-i\omega^2 |n|^2 t} u_n(t)$ and $\eta_n(t) = e^{i\omega^2 |n|^2 t} v_n(t)$, we obtain the non-autonomous Hamiltonian system

$$\forall n \in \mathbb{Z}^d, \quad \begin{cases} \dot{u}_n(t) &= -i\varepsilon \frac{\partial H_t}{\partial v_n}(u, v) \\ \dot{v}_n(t) &= i\varepsilon \frac{\partial H_t}{\partial u_n}(u, v) \end{cases}$$

with

$$H_\theta(u, v) = \frac{1}{p+1} \sum_{\mathbf{K} \in (\mathbb{Z}^d)^{2p+2}} \gamma_{\sigma_1(\mathbf{K})} e^{-i\omega^2 \sigma_2(\mathbf{K})\theta} (uv)_{\mathbf{K}}$$

In this setting, the Poisson bracket of two Hamiltonian functions F_1 and F_2 reads

$$\{F_1, F_2\} = -i \sum_{p \in \mathbb{Z}^d} \left(\frac{\partial F_1}{\partial u_p} \frac{\partial F_2}{\partial v_p} - \frac{\partial F_2}{\partial u_p} \frac{\partial F_1}{\partial v_p} \right),$$

so that one has

$$\frac{d}{dt} F(\xi(t), \eta(t)) = \{F, H\}(\xi(t), \eta(t)).$$

The term H_1 is obtained easily and is of the form

$$H_1(u, v) = \frac{1}{(p+1)} \sum_{\mathbf{K} \in (\mathbb{Z}^d)^{2p+2}, \sigma_2(\mathbf{K})=0} \gamma_{\sigma_1(\mathbf{K})} (uv)_{\mathbf{K}}. \quad (5.8)$$

Now, the Poisson bracket of H_s and H_t may be computed according to the following steps

$$\begin{aligned} \frac{\partial H_s}{\partial u_k} &= \sum_{\mathbf{K} \in (\mathbb{Z}^d)^{2p+1}} \gamma_{(\sigma_1(\mathbf{K})-k)} e^{i\omega^2(\sigma_2(\mathbf{K})-|k|^2)s} (vu)_{\mathbf{K}}, \\ \frac{\partial H_t}{\partial v_k} &= \sum_{\mathbf{L} \in (\mathbb{Z}^d)^{2p+1}} \gamma_{(\sigma_1(\mathbf{L})-k)} e^{-i\omega^2(\sigma_2(\mathbf{L})-|k|^2)t} (uv)_{\mathbf{L}}, \\ \{H_s, H_t\} &= 2 \sum_{\substack{\mathbf{K} \in (\mathbb{Z}^d)^{2p+1}, \\ \mathbf{L} \in (\mathbb{Z}^d)^{2p+1}, \\ k \in \mathbb{Z}^d}} \gamma_{(\sigma_1(\mathbf{K})-k)} \gamma_{(\sigma_1(\mathbf{L})-k)} \sin \left(\omega^2(\sigma_2(\mathbf{K})s - \sigma_2(\mathbf{L})t + |k|^2(t-s)) \right) (vu)_{\mathbf{K}} (uv)_{\mathbf{L}} \end{aligned}$$

We thus have to evaluate expressions of the form

$$I(\alpha, \beta) = \frac{\omega^2}{2\pi} \int_0^{\frac{2\pi}{\omega^2}} \int_0^s (\sin(\omega^2(\alpha s - \beta t))) dt ds$$

for $(\alpha, \beta) = (\sigma_2(\mathbf{K}) - |\mathbf{K}|^2, \sigma_2(\mathbf{L}) - |\mathbf{K}|^2) \in \mathbb{Z}^2$. This gives

$$\begin{aligned} I(0, 0) &= 0, \quad I(\alpha, 0) = \frac{1}{\omega^2 \alpha} \text{ for all } \alpha \neq 0, \quad I(0, \beta) = \frac{1}{\omega^2 \beta} \text{ for all } \beta \neq 0, \\ I(\alpha, \alpha) &= \frac{-1}{\omega^2 \alpha} \text{ for all } \alpha \neq 0, \quad I(\alpha, \beta) = 0 \text{ for all } \alpha \neq 0, \beta \neq 0, \alpha - \beta \neq 0. \end{aligned}$$

Hence, $H_2 = A + B + C$ with

$$\begin{aligned} A &= -\omega^{-2} \sum_{\substack{(\mathbf{K}, \mathbf{L}) \in (\mathbb{Z}^d)^{2p+1} \times (\mathbb{Z}^d)^{2p+1}, \\ k \in \mathbb{Z}^d, \sigma_2(\mathbf{K}) = |k|^2, \sigma_2(\mathbf{L}) \neq |k|^2}} \gamma_{(\sigma_1(\mathbf{K})-k)} \gamma_{(\sigma_1(\mathbf{L})-k)} \frac{(vu)_{\mathbf{K}} (uv)_{\mathbf{L}}}{\sigma_2(\mathbf{L}) - |k|^2}, \\ B &= -\omega^{-2} \sum_{\substack{(\mathbf{K}, \mathbf{L}) \in (\mathbb{Z}^d)^{2p+1} \times (\mathbb{Z}^d)^{2p+1}, \\ k \in \mathbb{Z}^d, \sigma_2(\mathbf{K}) \neq |k|^2, \sigma_2(\mathbf{L}) = |k|^2}} \gamma_{(\sigma_1(\mathbf{K})-k)} \gamma_{(\sigma_1(\mathbf{L})-k)} \frac{(vu)_{\mathbf{K}} (uv)_{\mathbf{L}}}{\sigma_2(\mathbf{K}) - |k|^2}, \\ C &= \omega^{-2} \sum_{\substack{(\mathbf{K}, \mathbf{L}) \in (\mathbb{Z}^d)^{2p+1} \times (\mathbb{Z}^d)^{2p+1}, \\ k \in \mathbb{Z}^d, \sigma_2(\mathbf{K}) = \sigma_2(\mathbf{L}) \neq |k|^2}} \gamma_{(\sigma_1(\mathbf{K})-k)} \gamma_{(\sigma_1(\mathbf{L})-k)} \frac{(vu)_{\mathbf{K}} (uv)_{\mathbf{L}}}{\sigma_2(\mathbf{L}) - |k|^2}. \end{aligned}$$

6 Numerical experiments

In this section, we present numerical simulations of some specific solutions to NLS found in the recent literature in order to illustrate these examples in the light of stroboscopic averaging. Stroboscopic averaging gives access not only to the first order averaged model, which is often the only model described explicitly in the literature, but also to higher order approximations, which may carry some interesting physical information, as it will be seen on these examples.

6.1 A problem of Grébert and Villegas-Blas [GVB11]

In what follows, we briefly derive the first averaged model and simulate a problem considered by B. Grébert and C. Villegas-Blas in [GVB11], which involves a cubic nonlinearity $|\psi|^2\psi$ multiplied by an inciting term of the form $2 \cos(2x)$. More precisely we consider the following Cauchy problem (Example 1)

$$\begin{aligned} i\partial_t \psi^\varepsilon &= -\Delta \psi^\varepsilon + 2\varepsilon \cos(2x)|\psi^\varepsilon|^2 \psi^\varepsilon, \quad t \geq 0, \quad \psi^\varepsilon(t, \cdot) \in H^s(\mathbb{T}_{2\pi}) \\ \psi(0, x) &= \cos x + \sin x. \end{aligned}$$

Classical arguments based on the conservation of the energy imply that this problem has a unique global solution in all Sobolev spaces $H^s(\mathbb{T}_{2\pi})$ for $s \geq 0$ (see [GVB11] for details and references therein). Writing the solution in Fourier $\psi^\varepsilon(t, x) = \sum_{k \in \mathbb{Z}} \xi_k(t) e^{ikx}$, Grébert and Villegas-Blas prove the following result.

Theorem 6.1 *For ε small enough, one has for all $|t| \leq \varepsilon^{-9/8}$ the following estimates:*

$$\begin{aligned} |\xi_1(t)|^2 &= \frac{1 + \sin(2\varepsilon t)}{2} + \mathcal{O}(\varepsilon^{1/8}), \\ |\xi_{-1}(t)|^2 &= \frac{1 - \sin(2\varepsilon t)}{2} + \mathcal{O}(\varepsilon^{1/8}). \end{aligned}$$

The estimate for $|\xi_1(t)|^2 + |\xi_{-1}(t)|^2$ implies that all the energy remains essentially on the two Fourier modes $+1$ and -1 , while estimates for $|\xi_1(t)|^2$ and $|\xi_{-1}(t)|^2$ account for the energy exchange between these two modes (named “beating effect” in [GVB11]). We have here $a = 2\pi$, $p = 1$, $d = 1$ and $\omega = 1$, while the coefficients γ_k , $k \in \mathbb{Z}$, read

$$\gamma_k = \frac{1}{2\pi} \int_0^{2\pi} 2 \cos(2x) e^{ikx} dx = \delta_{k-2} + \delta_{k+2}$$

with $\delta_n = 1$ if $n = 0$ and $\delta_n = 0$ otherwise. The Hamiltonian H_1 given by (5.8) is thus of the form

$$H_1(u, v) = \frac{1}{2} \sum_{\substack{(k, l, m, n) \in \mathbb{Z}^4, \\ k - l + m - n = \pm 2, \\ k^2 - l^2 + m^2 - n^2 = 0}} u_k v_l u_m v_n.$$

The proof of Theorem 6.1 in [GVB11] relies on a careful analysis of interactions between modes in the system corresponding to H_1 and we do not reproduce it here. Instead, we illustrate it by simulating:

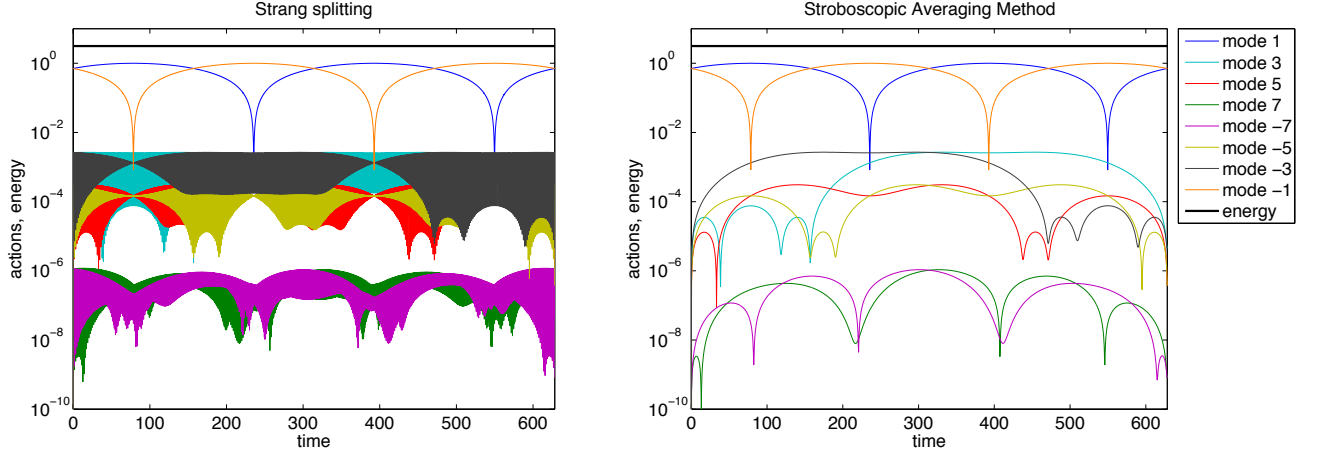


Figure 1: Example 1 with the Strang splitting method and the Stroboscopic Averaging Method.

- The original system integrated by Strang splitting with tiny stepsizes (see left of Fig. 1).
- The original system integrated by the Stroboscopic Averaging Method (SAM) which is presented in [CMSS10, CCMSS11] (see right of Fig. 1). Without getting into intricate details of this method, let us point out that it simulates the autonomous equation (2.22).
- The first order averaged system corresponding to εH_1 (see left of Fig. 2).
- The second order averaged system corresponding to $\varepsilon H_1 + \varepsilon^2 H_2$, (see right of Fig. 2).

In all figures, we represent (in logarithmic scale) so-called actions $|\xi_k|$, for modes $k = \pm 1, \pm 3, \pm 5, \pm 7$. Modes ± 1 are of order $\mathcal{O}(1)$ and one can observe on all pictures the beating effect between these two modes: all four pictures corroborate estimates of Theorem 6.1, which are based on the first order averaged system. The nonlinear energy is also represented and is well conserved by the four numerical methods. Meanwhile, higher-order averaged systems exhibit other interesting “beating” effects not yet investigated, which can be observed thanks to the Stroboscopic Averaging Method (SAM) or through their direct simulation.

Indeed, one can observe of the left of Fig. 1 that modes $\pm 3, \pm 5$ are of order $\mathcal{O}(\varepsilon)$, modes ± 7 of order $\mathcal{O}(\varepsilon^2)$, and that all these modes are highly oscillating, thus difficult to analyse. In contrast, high oscillations have been filtered out by the SAM (Fig. 1 right): the simulated equation on the right is the non-stiff equation (2.22)

$$\frac{dU}{dt} = \varepsilon \tilde{G}^\varepsilon(U) = \varepsilon G_1(U) + \varepsilon^2 G_2(U) + \varepsilon^3 G_3(U) + \dots,$$

amenable to much more efficient numerical method (see [CCMSS11]). Note that both solutions (left and right) on Figure 1 coincide at stroboscopic times $2\pi n$, $n \in \mathbb{N}$.

The equations simulated on Figure 2, left and right, are truncated versions of this averaged equation, respectively

$$\frac{dU}{dt} = \varepsilon G_1(U) \quad \text{and} \quad \frac{dU}{dt} = \varepsilon G_1(U) + \varepsilon^2 G_2(U).$$

We observe that Fig. 2 left and Fig. 2 right coincide up to terms of order $\mathcal{O}(\varepsilon)$ (i.e. only modes ± 1 are correctly simulated by the first order averaged model), while Fig. 2 right and Fig. 1 right coincide up to terms of order $\mathcal{O}(\varepsilon^2)$ (i.e. only modes $\pm 1, \pm 3, \pm 5$ are correctly simulated by the second order averaged model).

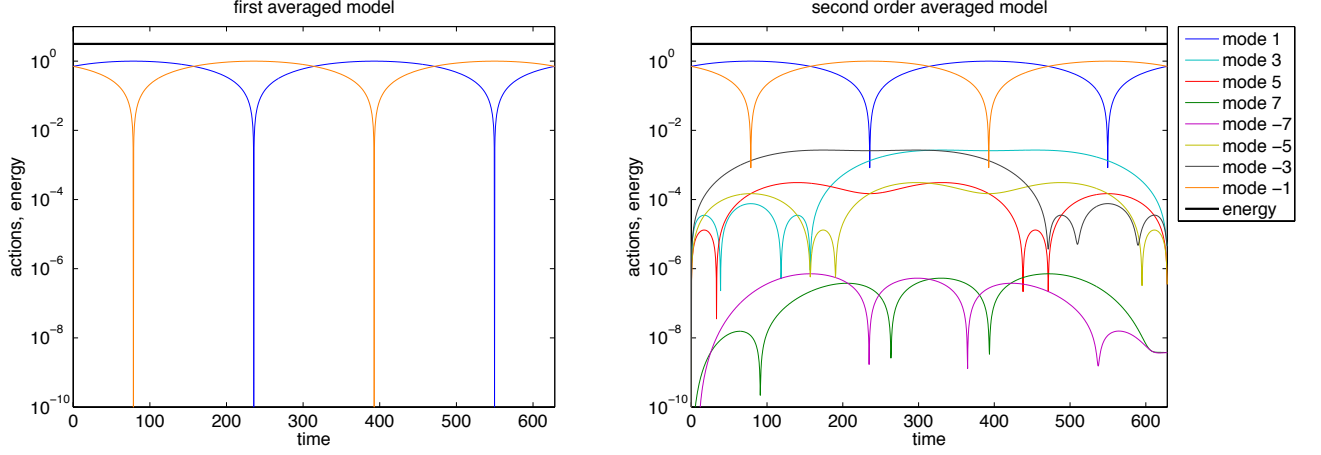


Figure 2: Example 1, first and second order averaged models.

6.2 A problem of Grébert and Thomann with quintic nonlinearity [GT12].

Our second example (Example 2) is the following problem considered by Grébert and Thomann in [GT12], involving a quintic nonlinearity $|\psi|^4\psi$:

$$i\partial_t \psi^\varepsilon = -\Delta \psi^\varepsilon + \varepsilon |\psi^\varepsilon|^4 \psi^\varepsilon, \quad t \geq 0, \quad \psi^\varepsilon(t, \cdot) \in H^1(\mathbb{T}_{2\pi}) \quad (6.1)$$

$$\psi(0, x) = \sqrt{\kappa_0} e^{ix} + \sqrt{\frac{\kappa_0}{2}} e^{-ix} + \sqrt{\kappa_0} e^{5ix} + \sqrt{\frac{1-\kappa_0}{2}} e^{7ix} \quad (6.2)$$

with $\kappa_0 = 0.24$. Thanks to energy conservation, they show that there exists a global solution $\psi^\varepsilon(t, \cdot) \in H^1(\mathbb{T}_{2\pi})$ provided the initial value lies itself in $H^1(\mathbb{T}_{2\pi})$. Moreover, if this initial value is chosen in such a way that its non-zero Fourier modes belong to a specific resonant set, then the solution exhibits periodic exchanges of energy. We state this result below in the specific case of a resonant set made of modes $-1, 1, 5$, and 7 :

Theorem 6.2 *There exist $\tilde{P} > 0$, $\tilde{\varepsilon}_0 > 0$ and a $2\tilde{P}$ -periodic function $\kappa : \mathbb{R} \rightarrow]0; 1[$ such that if $0 < \varepsilon < \varepsilon_0$, the solution to (6.1) satisfies for all $0 \leq t \leq \varepsilon^{-3/2}$*

$$\psi^\varepsilon(t, x) = \sum_{k=-1,1,5,7} \xi_j(t) e^{ikx} + \varepsilon^{1/4} q_1(t, x) + \varepsilon^{3/2} t q_2(t, x)$$

with

$$|\xi_5(t)|^2 = 2|\xi_{-1}(t)|^2 = \kappa(\varepsilon t), \quad |\xi_1(t)|^2 = 2|\xi_7(t)|^2 = 1 - \kappa(\varepsilon t),$$

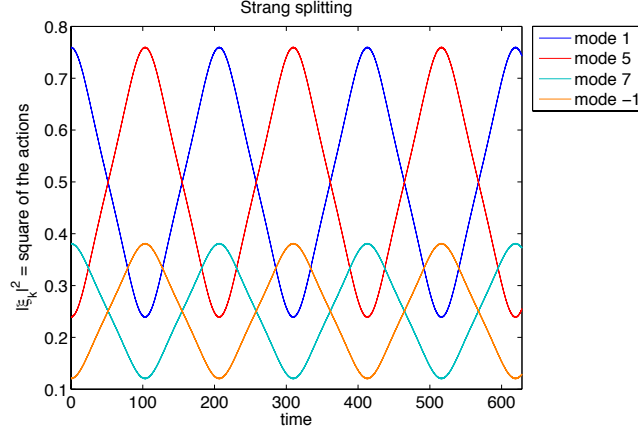


Figure 3: Example 2, beating effect between modes 1, 5, 7 and -1.

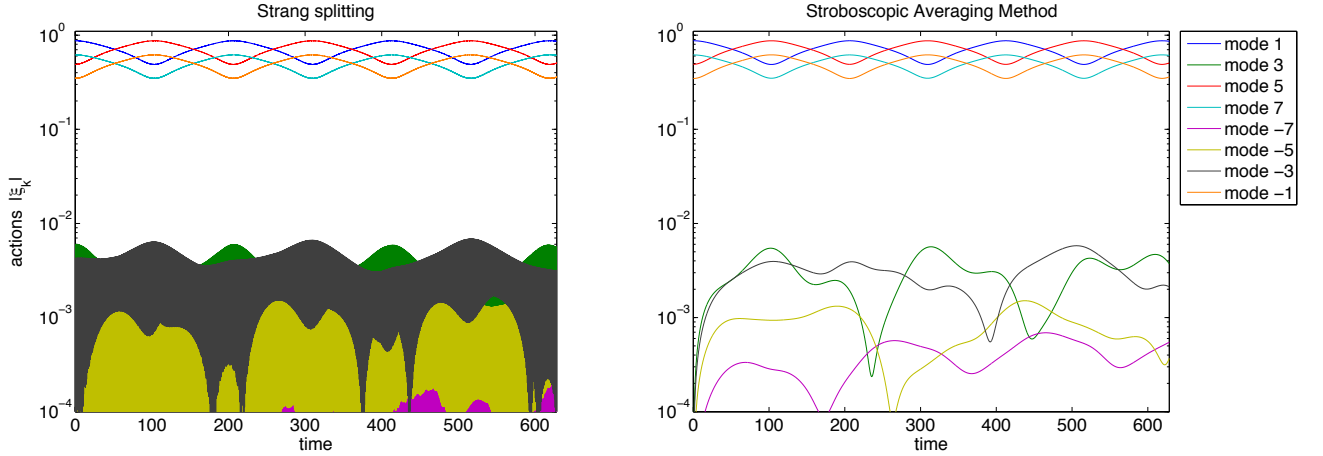


Figure 4: Example 2 with the Strang splitting method and the Stroboscopic Averaging Method.

and where for all $s \in \mathbb{R}$, $\|q_1(t, \cdot)\|_{H^s(\mathbb{T}_{2\pi})} \leq C_s$ for all $t \geq 0$, and $\|q_2(t, \cdot)\|_{H^s(\mathbb{T}_{2\pi})} \leq C_s$ for all $0 \leq t \leq \varepsilon^{-3/2}$.

Figure 3 presents (in linear scale) the evolution of the square of the actions $|\xi_1|^2$, $|\xi_{-1}|^2$, $|\xi_5|^2$, $|\xi_7|^2$ obtained by the Strang splitting method: the predicted energy exchange between modes 1, -1, 5 and 7 is clearly observed, in full agreement with Theorem 6.2.

On Figure 4, we represent (in logarithmic scale) the evolution of actions $|\xi_k|$ for $k \in \{\pm 1, \pm 3, \pm 5, \pm 7\}$ obtained by Strang's method (left) and by the SAM (right). The same observations as for Example 1 can be formulated. As a matter of fact, modes 3, -3, -5 and -7 are of order $\mathcal{O}(\varepsilon)$ and entail high oscillations. Those are filtered (see the right picture) by the high-order averaging process developed in this paper (actually, the SAM used here is exact up to $\mathcal{O}(\varepsilon^6)$ only, a sufficient accuracy here) and this sheds new light on an interesting smooth

macroscopic behavior at order $\mathcal{O}(\varepsilon)$.

6.3 A problem of Carles and Faou in two dimensions [CF12].

Our Example 3 is the problem considered by Carles and Faou in [CF12] (see also [Fao12]). It is the following nonlinear Schrödinger equation involving a cubic nonlinearity $|\psi|^2\psi$ and posed in the two-dimensional torus $\mathbb{T}_{2\pi}^2 = [0, 2\pi] \times [0, 2\pi]$:

$$i\partial_t\psi^\varepsilon = -\Delta\psi^\varepsilon \pm \varepsilon|\psi^\varepsilon|^2\psi^\varepsilon, \quad t \geq 0, \quad \psi^\varepsilon(t, \cdot) \in H^1(\mathbb{T}_{2\pi}^2) \quad (6.3)$$

$$\psi^\varepsilon(0, x) = 1 + 2\cos(x_1) + 2\cos(x_2). \quad (6.4)$$

Carles and Faou describe energy exchanges between the actions, as a cascade: high modes become significantly large at a time that depends on the mode and which is increasing with the size of the mode. Their analysis relies again on the careful study of the first order averaged system. Note that a related result is stated by Colliander et al. in [CKS⁺10].

In [CF12] Figure 1, this energy cascade is numerically illustrated by a direct simulation of the system using Strang splitting. We reproduce this experiment on Figure 5 left. As observed above, higher-order modes $\{|\xi_k|, k \geq 2\}$ are highly oscillatory and it is not obvious to distinguish on this graph first-order effects from higher-order (in ε) effects. To complete the picture, we show, on the right of Figure 5, the results obtained by the SAM: a smooth macroscopic behavior appears distinctly, though more complex than what the first averaged model suggests on Figure 6.

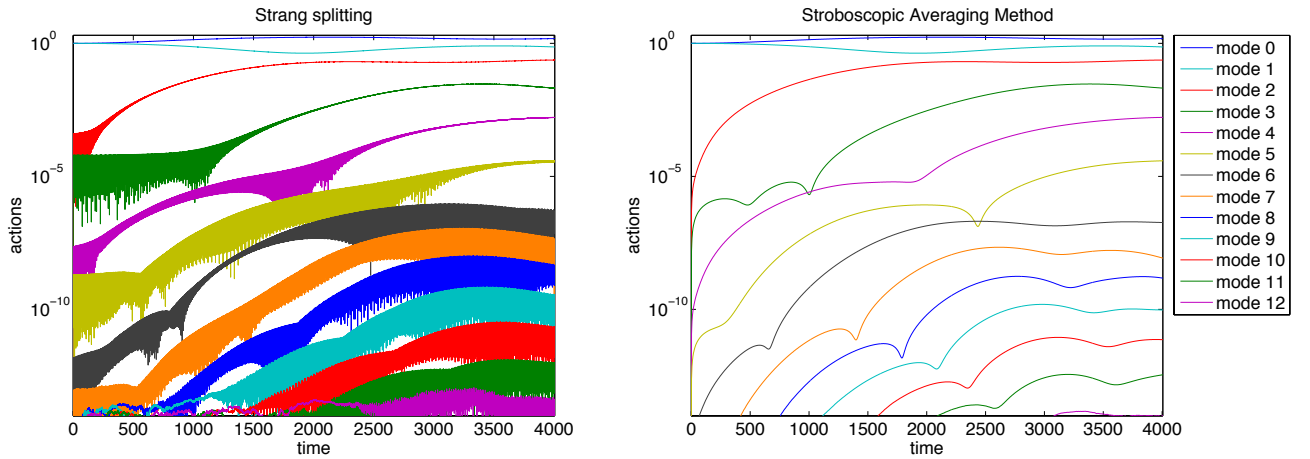


Figure 5: Example 3 with the Strang splitting method and the Stroboscopic Averaging Method.

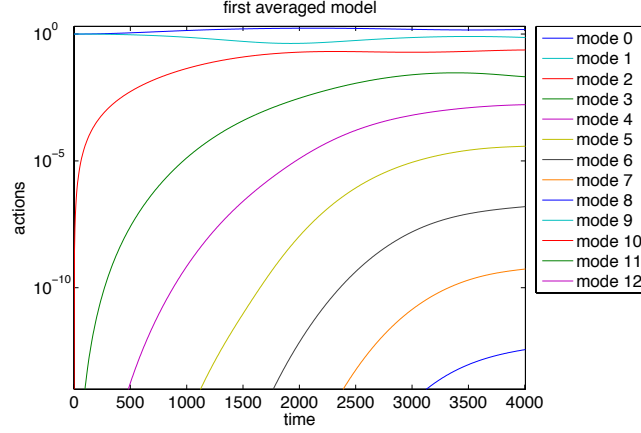


Figure 6: Example 3, first order averaged model.

Appendix

In this section, we give an elementary proof of Magnus expansion which does not make use of Lie-operators. The proof is indeed a direct consequence of the following

Proposition 6.3 *Given a vector field $(t, u) \mapsto g_t(u)$, there exist unique (formal) series*

$$V_{s,t}^\varepsilon(u) = \sum_{i \geq 1} \varepsilon^{i-1} V_{s,t}^{[i]}(u), \quad W_t^\varepsilon(u) = \sum_{i \geq 1} \varepsilon^{i-1} W_t^{[i]}(u), \quad (6.5)$$

such that

$$V_{0,t}^\varepsilon(u) = 0, \quad W_0^\varepsilon(u) = 0 \text{ and } V_{1,t}^\varepsilon(u) = g_t(u), \quad (6.6)$$

and satisfying the following requirement: for all u_0 , there exists $z(s, t) = u_0 + \sum_{i \geq 1} \varepsilon^i z_i$ such that

$$\begin{cases} \frac{\partial}{\partial t} z(s, t) = \varepsilon V_{s,t}^\varepsilon(z(s, t)), & z(s, 0) = u_0, \\ \frac{\partial}{\partial s} z(s, t) = \varepsilon W_t^\varepsilon(z(s, t)), & z(0, t) = u_0. \end{cases} \quad (6.7)$$

The $W^{[i]}$'s are given by

$$W_t^{[i]}(u) = \int_0^t R_\tau^{[i]}(u) d\tau, \quad (6.8)$$

where the $R^{[i]}$'s and the $V^{[i]}$'s are determined by

$$\varepsilon V_{s,\cdot}^\varepsilon = (\varepsilon s) R^\varepsilon + \frac{(\varepsilon s)^2}{2} R^\varepsilon \triangleleft R^\varepsilon + \frac{(\varepsilon s)^3}{3!} R^\varepsilon \triangleleft (R^\varepsilon \triangleleft R^\varepsilon) + \dots, \quad (6.9)$$

and where $R_t^\varepsilon(u) = \sum_{i \geq 1} \varepsilon^{i-1} R_t^{[i]}(u)$.

Recall that given two vector fields $f_t(u)$ and $g_t(u)$ the notation $h = f \triangleleft g$ is defined by (5.1). One can then readily check for arbitrary u_0 that, $z(1, t) = u^\varepsilon(t)$ is the unique solution of (2.1), that $U(t) = z(t/P, P)$ is the solution of the i.v.p.

$$\frac{d}{dt}U(t) = \frac{\varepsilon}{P} W_P^\varepsilon(U(t)), \quad U(0) = u_0,$$

and finally that $G^\varepsilon(U) := \frac{1}{P} W_P^\varepsilon(U)$ owing to $u^\varepsilon(P) = z(1, t) = U(P)$.

Proof of Proposition 6.3. Assume for the moment that the series in (6.5) satisfy (6.6) and (6.7). Deriving respectively the first and second equation in (6.7) w.r.t. respectively s and t , we arrive at

$$\partial_s V_{s,t}^\varepsilon(u) - \partial_t W_t^\varepsilon(u) = \varepsilon \left(\partial_u W_t^\varepsilon(u) \right) V_{s,t}^\varepsilon(u) - \varepsilon \left(\partial_u V_{s,t}^\varepsilon(u) \right) W_t^\varepsilon(u) \quad (6.10)$$

evaluated at $(s, t, u) = (s, t, z(s, t))$, where $z(s, t)$ is the unique solution of (6.7) corresponding to an arbitrary initial value u_0 . This implies that (6.10) must actually hold for arbitrary (s, t, u) . Now, by considering

$$R_t^\varepsilon(u) = \partial_t W_t^\varepsilon(u), \quad (6.11)$$

condition (6.10) can be rewritten as

$$\partial_s V_{s,t}^\varepsilon = R^\varepsilon + \varepsilon R^\varepsilon \triangleleft V_{s,t}^\varepsilon. \quad (6.12)$$

which leads to (6.9), taking into account that $V_{0,t}^\varepsilon(u) \equiv 0$. We finally have that (6.5), (6.6) and (6.7), imply (6.9) with (6.8).

Assume now conversely that a pair $(V^\varepsilon, W^\varepsilon)$ as in (6.5) satisfies (6.6), (6.8) and (6.9). The $R^{[i]}$'s are then clearly uniquely determined by equation (6.9) with $s = 1$, so that the $W^{[i]}$'s are also uniquely determined by (6.8). It remains to show that the pair $(V^\varepsilon, W^\varepsilon)$ satisfies requirement (6.7). Indeed, let $z(s, t) = u_0 + \sum_{i \geq 1} \varepsilon^i z_i(s, t)$ be the unique solution of the first equation of (6.7). Since $\partial_t z(0, t) = \varepsilon V_{0,t}^\varepsilon(z(0, t)) = 0$, we get that $z(0, t) = u_0$, and we only need to prove by induction on i , that for all $i \geq 1$,

$$\partial_s z(s, t) = \varepsilon W_t^\varepsilon(z(s, t)) + \mathcal{O}(\varepsilon^i).$$

Clearly, this holds true for $i = 1$. Assume that it holds true for a given $i \geq 1$. By derivation with respect to s of the first equation of (6.7) and use of the induction hypothesis, we get

$$\begin{aligned} \partial_{t,s}^2 z(s, t) &= \varepsilon (\partial_s V_{s,t}^\varepsilon)(z(s, t)) + \varepsilon \left(\partial_u V_{s,t}^\varepsilon(z(s, t)) \right) \partial_s z(s, t) \\ &= \varepsilon (\partial_s V_{s,t}^\varepsilon)(z(s, t)) + \varepsilon^2 \left(\partial_u V_{s,t}^\varepsilon(z(s, t)) \right) W_t^\varepsilon(z(s, t)) + \mathcal{O}(\varepsilon^{i+1}), \\ &= \varepsilon (\partial_t W_t^\varepsilon)(z(s, t)) + \varepsilon^2 \left(\partial_u W_t^\varepsilon(z(s, t)) \right) V_{s,t}^\varepsilon(z(s, t)) + \mathcal{O}(\varepsilon^{i+1}), \end{aligned}$$

where we used (6.10) (which can indeed be deduced from (6.9) through (6.12)). Finally, since $\partial_s z(s, 0) \equiv 0$, we get

$$\begin{aligned} \partial_s z(s, t) &= \varepsilon \int_0^t \left((\partial_\tau W_\tau^\varepsilon)(z(s, \tau)) + \varepsilon \partial_u W_\tau^\varepsilon(z(s, \tau)) V_{s,\tau}^\varepsilon(s, \tau, z(s, \tau)) \right) d\tau + \mathcal{O}(\varepsilon^{i+1}) \\ &= \varepsilon W_t^\varepsilon(z(s, t)) - \varepsilon W_0^\varepsilon(z(s, 0)) + \mathcal{O}(\varepsilon^{i+1}) \\ &= \varepsilon W_t^\varepsilon(z(s, t)) + \mathcal{O}(\varepsilon^{i+1}). \end{aligned}$$

□

References

- [BACM08] N. Ben Abdallah, F. Castella, and F. Méhats. Time averaging for the strongly confined nonlinear Schrödinger equation, using almost-periodicity. *J. Differ. Equations*, 245(1):154–200, 2008.
- [Bam03] D. Bambusi. Birkhoff normal form for some nonlinear PDEs. *Comm. Math. Phys.*, 234(2):253–285, 2003.
- [Bam05] D. Bambusi. Birkhoff normal form for some quasilinear Hamiltonian PDEs. In *XIVth International Congress on Mathematical Physics*, pages 273–280. World Sci. Publ., Hackensack, NJ, 2005.
- [Bam08] D. Bambusi. A Birkhoff normal form theorem for some semilinear PDEs. In *Hamiltonian dynamical systems and applications*, NATO Sci. Peace Secur. Ser. B Phys. Biophys., pages 213–247. Springer, Dordrecht, 2008.
- [BCOR10] S. Blanes, F. Casas, J.A. Oteo, and J. Ros. A pedagogical approach to the Magnus expansion. *Eur. J. Phys.*, 31(4):907–918, 2010.
- [BG06] D. Bambusi and B. Grébert. Birkhoff normal form for partial differential equations with tame modulus. *Duke Math. J.*, 135(3):507–567, 2006.
- [Bou96] J. Bourgain. Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations. *Geom. Funct. Anal.*, 6(2):201–230, 1996.
- [Bou07] J. Bourgain. Normal forms and the nonlinear Schrödinger equation. In *Perspectives in nonlinear partial differential equations*, volume 446 of *Contemp. Math.*, pages 153–157. Amer. Math. Soc., Providence, RI, 2007.
- [Bri26] L. Brillouin. La mécanique ondulatoire de Schrödinger; une méthode générale de résolution par approximations successives. *C. B.*, 183:24–26, 1926.
- [Car08] R. Carles. *Semi-classical analysis for nonlinear Schrödinger equations*. Hackensack, NJ: World Scientific. xi, 243 p. , 2008.
- [Caz03] T. Cazenave. *Semilinear Schrödinger equations*. Courant Lecture Notes in Mathematics 10. Providence, RI: American Mathematical Society (AMS); New York, NY: Courant Institute of Mathematical Sciences. xiii, 323 p. , 2003.
- [CCMSS11] M.P. Calvo, Ph. Chartier, A. Murua, and J.-M. Sanz-Serna. Numerical stroboscopic averaging for ODEs and DAEs. *Appl. Numer. Math.*, 61(10):1077–1095, 2011.
- [CF12] R. Carles and E. Faou. Energy cascades for NLS on the torus. *Discrete Contin. Dyn. Syst.*, 32(6):2063–2077, 2012.

- [CH98] T. Cazenave and A. Haraux. *An introduction to semilinear evolution equations. Transl. by Yvan Martel. Revised ed.* Oxford Lecture Series in Mathematics and its Applications. 13. Oxford: Clarendon Press. xiv, 186 p. , 1998.
- [CHL08a] D. Cohen, E. Hairer, and C. Lubich. Conservation of energy, momentum and actions in numerical discretizations of non-linear wave equations. *Numer. Math.*, 110(2):113–143, 2008.
- [CHL08b] D. Cohen, E. Hairer, and C. Lubich. Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions. *Arch. Ration. Mech. Anal.*, 187(2):341–368, 2008.
- [CKS⁺10] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation. *Invent. Math.*, 181(1):39–113, 2010.
- [CMSS] Ph. Chartier, A. Murua, and J.-M. Sanz-Serna. Higher-order averaging, formal series and numerical integration ii: The quasi-periodic case. *Foundations of Computational Mathematics*, pages 1–38. 10.1007/s10208-012-9118-8.
- [CMSS10] Ph. Chartier, A. Murua, and J.M. Sanz-Serna. Higher-order averaging, formal series and numerical integration. I: B-series. *Found. Comput. Math.*, 10(6):695–727, 2010.
- [CMSS12] Ph. Chartier, A. Murua, and J.-M. Sanz-Serna. A formal series approach to averaging: exponentially small error estimates. *Discrete Contin. Dyn. Syst.*, 32(9):3009–3027, 2012.
- [Fao12] E. Faou. *Geometric numerical integration and Schrödinger equations.* Zurich Lectures in Advanced Mathematics. Zürich: European Mathematical Society (EMS). viii, 138 p. , 2012.
- [GL10] L. Gauckler and C. Lubich. Nonlinear Schrödinger equations and their spectral semi-discretizations over long times. *Found. Comput. Math.*, 10(2):141–169, 2010.
- [GT12] B. Grébert and L. Thomann. Resonant dynamics for the quintic nonlinear Schrödinger equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 29(3):455–477, 2012.
- [GVB11] B. Grébert and C. Villegas-Blas. On the energy exchange between resonant modes in nonlinear Schrödinger equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(1):127–134, 2011.
- [HL00] E. Hairer and C. Lubich. Long-time energy conservation of numerical methods for oscillatory differential equations. *SIAM J. Numer. Anal.*, 38(2):414–441 (electronic), 2000.
- [JMR95] J.-L. Joly, G. Métivier, and J. Rauch. Coherent and focusing multidimensional nonlinear geometric optics. *Ann. Sci. École Norm. Sup. (4)*, 28(1):51–113, 1995.

- [Kra26] H. A. Kramers. Wellenmechanik und halbzahlige Quantisierung. *Z. f. Physik*, 39:828–840, 1926.
- [LM88] P. Lochak and C. Meunier. *Multiphase averaging for classical systems. With applications to adiabatic theorems. Transl. from the French by H. S. Dumas*. Applied Mathematical Sciences, 72, New York etc.: Springer-Verlag. xi, 360 p. DM 78.00 , 1988.
- [Nei84] A.I. Neishtadt. The separation of motions in systems with rapidly rotating phase. *J. Appl.Math. Mech.*, 48:133–139, 1984.
- [Per69] L.M. Perko. Higher order averaging and related methods for perturbed periodic and quasiperiodic systems. *SIAM J. Appl. Math.*, 17:698–724, 1969.
- [PT87] J. Pöschel and E. Trubowitz. *Inverse spectral theory*. Pure and Applied Mathematics, Vol. 130. Boston etc.: Academic Press, Inc., Harcourt Brace Jovanovich, Publishers. X, 192 p.; \$ 29.95 , 1987.
- [Sim94] C. Simó. Averaging under fast quasiperiodic forcing. In *Hamiltonian mechanics (Toruń, 1993)*, volume 331 of *NATO Adv. Sci. Inst. Ser. B Phys.*, pages 13–34. Plenum, New York, 1994.
- [SVM07] J.A. Sanders, F. Verhulst, and J. Murdock. *Averaging methods in nonlinear dynamical systems. 2nd ed.* Applied Mathematical Sciences 59. New York, NY: Springer. xxi, 431 p. , 2007.
- [Wen26] G. Wentzel. Eine Verallgemeinerung der Quantenbedingungen für die Zwecke der Wellenmechanik. *Z. f. Physik*, 38:518–529, 1926.