Stroboscopic averaging for the nonlinear Schrödinger equation

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

October 28, 2014

Communicated by Arieh Iserles

Abstract

In this paper, we are concerned with an averaging procedure, namely *Stroboscopic averaging*, 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 then illustrated numerically on several examples borrowed from the recent literature.

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

MSC numbers: 35A35, 35J10, 34K33.

1 Introduction

In this article we are concerned with highly-oscillatory evolution equations posed in a Banach space X

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

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

[†]Corresponding author. 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

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 where it is assumed that the above problem is well-posed on a ε-dependent time interval $[0, T/\varepsilon]$. This is a high-frequency system, with one frequency, posed in an infinite-dimensional setting. The highly-oscillatory character of the equation stems from the length of the interval (scaled by $1/\varepsilon$), which makes the dynamics non-trivial: over such intervals, the change in the solution is loosely speaking of size $\mathcal{O}(1)$. Under more stringent hypotheses on the problem, the error bounds derived in this paper hold on even longer intervals of size $\mathcal{O}(1/\varepsilon^{1+\alpha})$ with $0 \leq \alpha < 1$.

The question we address is that of *high-order* averaging. In other words, we look for an approximation $v^\varepsilon(t)$ of $u^\varepsilon(t)$ such that $v^\varepsilon(t)$ is close to $u^\varepsilon(t)$ over $[0, T/\varepsilon]$ to within a small remainder term, *and* such that $v^\varepsilon(t)$ satisfies an *autonomous* non-oscillatory evolution equation. The typical example we have in mind is that of nonlinear PDEs in a high-frequency regime, where only one frequency is present. 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.

Ideally, the purpose of high-order averaging is to find a *periodic*, *near-identity* and smooth change of variable Φ^ε_θ, together with the flow map Ψ^ε_t of an *autonomous* differential equation on X, such that the solution of the original equation $(\mathcal{Q}^\varepsilon)$ takes the *composed form*

$$u^\varepsilon(t) = \Phi^\varepsilon_t \circ \Psi^\varepsilon_t(u_0).$$

Such a form completely[1] separates the dependence of $u^\varepsilon(t)$ upon the fast, periodic variable θ and the dependence of $u^\varepsilon(t)$ upon the slow variable t. Various choices are possible for the change of variables Φ^ε_θ. For instance, it is somehow common to impose that

$$\frac{1}{P}\int_0^P \Phi^\varepsilon_\theta \, d\theta = \mathrm{Id},$$

and it was shown in [CMSS13b] in the context of ordinary differential equations in $\mathbb{R}^d$ that all choices of Φ^ε_θ actually relate to each other through a change of variable independent of θ. However, the only geometric choice is *stroboscopic averaging* for which $\Phi^\varepsilon_{\theta=0} = \mathrm{Id}$. In this case, the original vector field and the averaged vector field lie in the same Lie-algebra.[2] Since the proof of this last result relies strongly upon the use of B-series for ODEs (see [CMSS10, CMSS12] for the single-frequency case and [CMSS13b, CMSS13a] for the quasi-periodic case), it is necessary in the context of infinite dimensional PDEs to use alternative arguments based only on the decomposition $\Phi^\varepsilon_t \circ \Psi^\varepsilon_t$.

The factorized form $u^\varepsilon(t) = \Phi^\varepsilon_t \circ \Psi^\varepsilon_t(u_0)$ 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 $u^\varepsilon(t) = U^\varepsilon(\theta, t)\big|_{\theta=t}$ as is more usual in the context of high-frequency PDEs. Averaging as developed in this paper also constitutes in our view an alternative to other modern techniques, amongst which stand most prominently Birkhoff's forms (see e.g. Bambusi [Bam03, Bam05, BG06, Bam08], Bourgain [Bou96, Bou07], Colliander [CKS+10] and Grébert [GVB11, GT12], to mention just a few authors) and more recently Modulated Fourier

[1] The result we establish in that direction is that such a decomposition may be achieved with error terms of size $\mathcal{O}(e^{-c/\varepsilon})$ for some $c > 0$.

[2] This is for instance not the case whenever $\frac{1}{P}\int_0^P \Phi^\varepsilon_\theta \, d\theta = \mathrm{Id}$.

Expansions (see Hairer and Lubich [HL00] for ODEs and [CHL08, GL10] for Hamiltonian PDEs). Averaging of Hamiltonian PDEs with a single fast frequency has also been considered by Matthies and Scheel [MS03b] in a slightly different context, where the PDE can not be rewritten as an evolution equation in a Banach space. Typically, analytic initial conditions are then necessary to compensate for the possible loss of space derivatives and a weaker, though more general, estimate of the remainder terms is obtained. The decomposition we propose here is also reminiscent of references in fluid dynamics [MS03a, AM78].

Note that most of the papers quoted above deal with the case of non-resonant frequencies and that traditionally, the flow Ψ_t^ε and the change of variables Φ_θ^ε are sought by performing power expansions in ε (see e.g. Meunier [LM88] or Sanders-Verhulst [SVM07]). In the one-frequency context, this procedure goes back to the work of Perko in the finite-dimensional case [Per69]. In contrast, we identify here a new equation (here called *transport equation*) on the change of variables Φ_t^ε, which can be solved using a fixed point procedure in an analytic framework[3]. This transport equation plays a role that is somehow similar to the so-called singular equation of geometric optics, in the context of high-frequency hyperbolic PDEs (see [JMR95]).

Our main theorem (Theorem 2.7) is stated in Subsection 2.2, once preliminary assumptions have been settled, and ideas sustaining stroboscopic averaging outlined in Subsection 2.1. The remaining of Section 2 is devoted to technical proofs and intermediate theorems, with the exception of Subsection 2.4 which is concerned with the linear case, where expansions are shown to converge (in contrast with the general nonlinear 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 instanciation for NLS of Theorems 2.7, 3.5 and 3.7. In the NLS context, Theorem 4.2 can be considered as the main theoretical output of this article. Finally, Section 5 is devoted to some numerical illustrations of stroboscopic averaging for three NLS equations in dimensions 1 and 2.

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

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

where the initial datum u_0 is given and where the function $(\theta, u) \in \mathbb{T} \times X \mapsto g_\theta(u) \in X$ is smooth with respect to $u \in X$ and smooth and periodic[4] in $\theta \in \mathbb{T}$. We actually require $u \mapsto g_\theta(u)$ to be real analytic in a sense we define later (see Assumption 2.3). 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.

[3]Once the change of variables Φ_t^ε is obtained, we show that the flow Ψ_t^ε is easily reconstructed.

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

Assumption 2.1 *The Cauchy problem* (2.1) *is uniformly well-posed in the following sense. There exist $T > 0$, $\varepsilon^* > 0$, 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 $t \leq T/\varepsilon$.*

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[5]

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

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

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

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

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

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

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

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

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

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 $\mathbb{T} \times K_{2R}$.*

The above assumption defines once for all the quantities R and C_K.

[5]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

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

2.1 The formal equations of stroboscopic averaging

Ideally, the purpose of averaging is to find a periodic, near-identity change of variable

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

together with the flow map Ψ_t^ε of an autonomous differential equation with vector field G^ε on X

$$\frac{d}{dt}\Psi_t^\varepsilon(u_0) = \varepsilon\, G^\varepsilon\left(\Psi_t^\varepsilon(u_0)\right), \tag{2.2}$$

such that the solution of the original equation (2.1) takes the *composed form*[6]

$$u^\varepsilon(t) = \Phi_t^\varepsilon \circ \Psi_t^\varepsilon(u_0). \tag{2.3}$$

In the framework of stroboscopic averaging, we further impose that the mapping Φ_θ satisfies $\Phi_{\theta=0}^\varepsilon = \mathrm{Id}$.

Let us now derive equations for Φ_θ^ε and Ψ_t^ε. By differentiating both sides of (2.3) w.r.t. t and using (2.2) we readily get

$$\frac{\partial \Phi_t^\varepsilon}{\partial t}\left(\Psi_t^\varepsilon(u_0)\right) + \varepsilon \frac{\partial \Phi_t^\varepsilon}{\partial u}\left(\Psi_t^\varepsilon(u_0)\right) G^\varepsilon\left(\Psi_t^\varepsilon(u_0)\right) = \varepsilon\, g_t\left(\Phi_t^\varepsilon \circ \Psi_t^\varepsilon(u_0)\right), \tag{2.4}$$

so that, upon replacing $u_0 = \Psi_{-t}(u)$ and then t by $\theta \in \mathbb{T}$, we obtain

$$\frac{\partial \Phi_\theta^\varepsilon}{\partial \theta}(u) + \varepsilon \frac{\partial \Phi_\theta^\varepsilon}{\partial u}(u)\, G^\varepsilon(u) = \varepsilon\, g_\theta\left(\Phi_\theta^\varepsilon(u)\right). \tag{2.5}$$

Firstly, averaging in θ both sides of (2.5) eliminates the term $\partial_\theta \Phi_\theta^\varepsilon$ owing to periodicity, so that

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

where we have used the standard notation

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

for the average w.r.t. θ of a function $(\theta, u) \in \mathbb{T} \times X \mapsto f_\theta(u) \in X$. Assuming for the time being that the linear operator $v \mapsto \dfrac{\partial \langle \Phi^\varepsilon \rangle}{\partial u}(u)\, v$ is invertible for any u, we get

$$G^\varepsilon(u) := \left(\frac{\partial \langle \Phi^\varepsilon \rangle}{\partial u}(u)\right)^{-1} \langle g \circ \Phi^\varepsilon \rangle (u). \tag{2.6}$$

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

[6] *Stricto sensu*, we have frozen the initial datum u_0 up to now, and relation (2.3) 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.

Secondly, inserting previous relation in equation (2.5), we are led to the relation

$$\frac{\partial \Phi^\varepsilon_\theta}{\partial \theta}(u) + \varepsilon \frac{\partial \Phi^\varepsilon_\theta}{\partial u}(u) \left(\frac{\partial \langle \Phi^\varepsilon \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \Phi^\varepsilon \rangle (u) = \varepsilon\, g_\theta \circ \Phi^\varepsilon_\theta(u), \tag{2.7}$$

i.e., in integral form,

$$\Phi^\varepsilon_\theta(u) = u + \varepsilon \int_0^\theta \left(g_\xi \circ \Phi^\varepsilon_\xi(u) - \frac{\partial \Phi^\varepsilon_\xi}{\partial u}(u) \left(\frac{\partial \langle \Phi^\varepsilon \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \Phi^\varepsilon \rangle (u) \right) d\xi. \tag{2.8}$$

This is a closed equation on Φ^ε_θ, though nonlinear and nonlocal. In our perspective, once equation (2.8) is solved for Φ^ε_θ, the vector field G^ε and the associated flow Ψ^ε_t are immediately deduced along formula (2.6). It thus remains to turn the above formal computations into a rigorous analytic procedure. However, it is known [Nei84, CMSS12] that in general, equation (2.7) can only be solved *up to an error* of size $\mathcal{O}(\exp(-c/\varepsilon))$ for some $c > 0$.

2.2 Main result: averaging to within exponentially small remainder terms

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

$$\Lambda(\varphi)_\theta(u) = g_\theta \circ \varphi_\theta(u) - \frac{\partial \varphi_\theta}{\partial u}(u) \left(\frac{\partial \langle \varphi \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \varphi \rangle (u), \tag{2.9}$$

$$\Gamma^\varepsilon(\varphi)_\theta(u) = u + \varepsilon \int_0^\theta \Lambda(\varphi)_\xi(u)\, d\xi. \tag{2.10}$$

Remark 2.4 *Note that if φ_θ is periodic, then $\langle \Lambda(\varphi) \rangle \equiv 0$ so that $\Gamma^\varepsilon(\varphi)_\theta$ is also periodic.*

The basic equation of averaging (2.8) in its integral form,

$$\Phi^\varepsilon_\theta(u) = \Gamma^\varepsilon(\Phi^\varepsilon)_\theta(u), \tag{2.11}$$

or equivalently, in its differential form,

$$\partial_\theta \Phi^\varepsilon_\theta(u) = \varepsilon \Lambda(\Phi^\varepsilon)_\theta(u), \quad \Phi^\varepsilon_{\theta=0}(u) = u, \tag{2.12}$$

can be considered as a fixed point equation. We thus consider the sequence of functions

$$\Phi^{[0]}_\theta = \mathrm{Id}, \quad \Phi^{[k+1]}_\theta = \Gamma^\varepsilon(\Phi^{[k]})_\theta, \quad k = 0, 1, 2, \ldots, n, \tag{2.13}$$

together with the sequence of vector fields

$$G^{[k]}(u) := \left(\frac{\partial \langle \Phi^{[k]} \rangle}{\partial u}(u) \right)^{-1} \langle g \circ \Phi^{[k]} \rangle (u), \quad k = 0, 1, 2, \ldots, n. \tag{2.14}$$

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

$$G_{k+1}(u) = \frac{1}{k!} \left. \frac{d^k G^{[k]}}{d\varepsilon^k} \right|_{\varepsilon=0} (u), \tag{2.15}$$

and define $\widetilde{G}^{[n]}(u) = \sum_{k=0}^{n} \varepsilon^k G_{k+1}(u)$ for $n \geq 0$.

Remark 2.5 *By construction (see Theorem* 2.12*), for all* $n, k \geq 0$*, one has*

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

We now set $\varepsilon_0 := \frac{R}{8C_K P}$, $\varepsilon_2 = \min\left(\frac{\varepsilon_0}{80}, \frac{\varepsilon_0^2 P}{2T}\right)$, $r_n := \frac{R}{n+1}$ and $R_k = 2R - k r_n$ for $k = 0, \ldots, n+1$.

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:*

$$\forall \varepsilon \in \mathbb{C},\ |\varepsilon| < \varepsilon_0/(n+1), \qquad \|\Phi^{[k]} - \mathrm{Id}\|_{R_k} \leq \frac{r_n}{2}. \tag{2.16}$$

Theorem 2.7 *For* $0 < |\varepsilon| < \min(\varepsilon^*, \varepsilon_2)$*, consider* $\widetilde{\Phi}_\theta^\varepsilon = \Phi_\theta^{[n_\varepsilon]}$ *and* $\widetilde{G}^\varepsilon = \widetilde{G}^{[n_\varepsilon]}$ *for* $(n_\varepsilon + 1) = \lfloor \varepsilon_0/(4|\varepsilon|) \rfloor$*. Then the following assertions hold for the solution* $u^\varepsilon(t)$ *of* (2.1)*:*

(i) There exists a function $(\theta, u) \mapsto \mathcal{R}_\theta^\varepsilon(u)$ *periodic and smooth in* θ*, analytic on* K_R*, and bounded*

$$\|\mathcal{R}^\varepsilon\|_R \leq 24\, C_K \exp\left(-\frac{\varepsilon_0}{8\,|\varepsilon|}\right), \tag{2.17}$$

such that

$$\forall t \in [0, T/\varepsilon], \qquad u^\varepsilon(t) = \widetilde{\Phi}_t^\varepsilon(U(t)), \tag{2.18}$$

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

$$\frac{dU}{dt} = \varepsilon\, \widetilde{G}^\varepsilon(U) + \varepsilon\, \mathcal{R}_t^\varepsilon(U). \tag{2.19}$$

(ii) If $\widetilde{\Psi}_t^\varepsilon$ *denotes the* t*-flow of the* autonomous *differential equation*

$$\frac{dU}{dt} = \varepsilon \widetilde{G}^\varepsilon(U), \tag{2.20}$$

then $u^\varepsilon(t)$ *is exponentially close to* $\widetilde{\Phi}_t^\varepsilon \circ \widetilde{\Psi}_t^\varepsilon(u_0)$

$$\forall t \in [0, T/\varepsilon], \quad \left\| u^\varepsilon(t) - \widetilde{\Phi}_t^\varepsilon \circ \widetilde{\Psi}_t^\varepsilon(u_0) \right\|_{X_\mathbb{C}} \leq 36\, R \exp\left(-\frac{\varepsilon_0}{16\,|\varepsilon|}\right). \tag{2.21}$$

(iii) If furthermore $T = +\infty$ *in Assumption* 2.1*, then (2.18) holds for all* $t > 0$*, and* (2.21) *holds on any interval of the form* $[0, \tilde{T}/\varepsilon^{1+\alpha}]$ *with* $\tilde{T} > 0$ *and* $0 < \alpha < 1$*, provided*

$$0 < \varepsilon < \min\left(\varepsilon^*, \frac{\varepsilon_0}{80}, \left(\frac{\varepsilon_0^2 P}{2\tilde{T}}\right)^{\frac{1}{1-\alpha}}\right).$$

Remark 2.8 *Easy though tedious computations using* (2.13) *and* (2.14) *lead to the following expressions of the first three terms of the averaged equation:*

$$G_1(u) = \frac{1}{P}\int_0^P g_\tau(u)d\tau, \quad G_2(u) = \frac{-1}{2P}\int_0^P\int_0^\tau [g_s(u), g_\tau(u)]dsd\tau \qquad \text{and}$$
$$G_3(u) = \frac{1}{4P}\int_0^P\int_0^\tau\int_0^s [[g_r(u), g_s(u)], g_\tau(u)]drdsd\tau + \frac{1}{12P}\int_0^P\int_0^\tau\int_0^\tau [g_r(u), [g_s(u), g_\tau(u)]]drdsd\tau,$$

where $[f^1, f^2] = (\partial_u f^1)f^2 - (\partial_u f^2)f^1$ *denotes the usual Lie-bracket of smooth functions. Further terms can be formally obtained by using a non-linear Magnus expansion. Each of these is a linear combination of iterated integrals of iterated brackets of* g*, which implies that truncation at any order of the averaged vector field* G^ε *lies in the same Lie-algebra as the vector field of the parent system. Note that this also includes the case of volume preserving flows.*

2.3 Technical proofs and intermediate Theorems

In this Subsection, we give the proof of Theorem 2.7 while stating a few intermediate results (fixed-order truncation, non-expanded version of the vector field).

2.3.1 Two basic lemmas

Next lemma gives sufficient conditions for the quantities $\Lambda(\varphi)$ and $\Gamma^\varepsilon(\varphi)$ to be well-defined.

Lemma 2.9 *Let* $0 < \delta < \rho \le 2R$*. 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*

$$\|\varphi - \mathrm{Id}\|_\rho \le \frac{\delta}{2}.$$

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*

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

(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* $|\varepsilon| \ge 0$,

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

Proof. From the assumption $\|\varphi - \mathrm{Id}\|_\rho \le \delta/2$ and a Cauchy estimate applied to $\varphi_\theta(u) - u$, we get

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

Consequently, using Neumann series, we obtain

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

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

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

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

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

Integration in θ next provides $\|\Gamma^\varepsilon(\varphi) - \mathrm{Id}\|_{\rho-\delta} \leq 4\, C_K\, P\, |\varepsilon|$. 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. □

Lemma 2.9 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. We now establish the following contraction property.

Lemma 2.10 *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*

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

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

$$\|\Lambda(\varphi) - \Lambda(\hat{\varphi})\|_{\rho-\delta} \leq \frac{16\, C_K}{\delta} \|\varphi - \hat{\varphi}\|_\rho \quad \text{and} \quad \|\Gamma^\varepsilon(\varphi) - \Gamma^\varepsilon(\hat{\varphi})\|_{\rho-\delta} \leq \frac{16\, C_K\, P\, |\varepsilon|}{\delta} \|\varphi - \hat{\varphi}\|_\rho.$$

Proof. 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

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

Proceeding as in Lemma 2.9, 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 $\|\varphi - \mathrm{Id}\|_\rho \leq \delta/2$ 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 finally 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. □

2.3.2 Existence and uniqueness of quasi-solutions to (2.7) with polynomial remainder

In this section we exhibit (quasi-) solutions to the equations of averaging, i.e. (2.8) for the mapping Φ_θ^ε and (2.6) for the autonomous vector field G^ε. These are quasi-solutions in the sense that equation (2.8) is solved to within an 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. Obviously, function $\Phi_\theta^{[0]} = \mathrm{Id}$ is analytic in the sense of Definition 2.2 on K_{R_0}, analytic for all ε, smooth in θ, and satisfies (2.16). Assume now that, for an integer $k \leq n$, the function $\Phi_\theta^{[k]}$ obeys the conditions of Lemma 2.6. Lemma 2.9 with $\rho = R_k$ and $\delta = \frac{R}{n+1} = r_n$ then shows that the function $\Phi_\theta^{[k+1]}$ defined by (2.13) is analytic on $K_{R_{k+1}}$ and satisfies

$$\forall 0 \leq |\varepsilon| < \varepsilon_0/(n+1), \qquad \|\Phi^{[k+1]} - \mathrm{Id}\|_{R_{k+1}} \leq 4C_K\, P\, |\varepsilon| < \frac{4\, C_K\, P\, \varepsilon_0}{n+1} = \frac{R}{2(n+1)} = \frac{r_n}{2}.$$

It is again analytic in ε as composition of analytic functions, and clearly smooth in θ. This finishes the induction for $\Phi^{[k]}$. From definitions (2.14) and (2.15), $G^{[k]}$ and G_{k+1} are then also analytic in K_{R_k}. □

We are now in position to establish the existence of quasi-solutions to (2.7).

Theorem 2.11 [Existence of quasi-solutions to (2.7)] *For $n \in \mathbb{N}$, consider the sequence of functions*

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

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

$$\varepsilon\delta_\theta^{[k]}(u) := \frac{\partial \Phi_\theta^{[k]}}{\partial\theta}(u) - \varepsilon\Lambda(\Phi^{[k]})_\theta(u), \quad k = 0, \ldots, n. \tag{2.22}$$

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 $|\varepsilon| < \varepsilon_0/(n+1)$

$$\|\Phi^{[n]} - \mathrm{Id}\|_{R+r_n} \leq \frac{r_n}{2}, \qquad \|\delta^{[n]}\|_R \leq 2\, C_K \left((n+1)\frac{2|\varepsilon|}{\varepsilon_0} \right)^n. \tag{2.23}$$

(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 $\|(\Phi^{[n]})^{-1}\|_{R-r_n} \leq R$.

Proof of Theorem 2.11. Statement (i) and first estimate of (2.23) are obvious consequences of Lemma 2.6, up to a possible singularity at $\varepsilon = 0$ which is ruled out by second estimate (2.23), 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, Lemma 2.10 with $\delta = r_n$, $\rho = R + \delta = R_n$ yields (denoting $a_n = \frac{16 C_K P|\varepsilon|}{r_n}$)

$$|\varepsilon| \|\delta^{[n]}\|_R \le \frac{16\, C_K\, |\varepsilon|}{r_n} \|\Phi^{[n]} - \Phi^{[n-1]}\|_{R_n} = \frac{a_n}{P} \left\| \Gamma^\varepsilon \left(\Phi^{[n-1]}\right) - \Gamma^\varepsilon \left(\Phi^{[n-2]}\right) \right\|_{R_n}$$
$$\le \frac{a_n^2}{P} \left\| \Gamma^\varepsilon \left(\Phi^{[n-2]}\right) - \Gamma^\varepsilon \left(\Phi^{[n-3]}\right) \right\|_{R_{n-1}} \le \dots \le \frac{a_n^n}{P} \|\Phi^{[1]} - \Phi^{[0]}\|_{R_1}.$$

On the other hand, by definition of $\Phi_\theta^{[1]}$, we have $\|\Phi^{[1]} - \Phi^{[0]}\|_{R_1} \le 2|\varepsilon|\, P\, C_K$. This proves the second part of (2.23).

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

$$\|u_1 - u_2\|_{X_\mathbb{C}} \le \|\partial_u \Phi^{[n]} - \mathrm{Id}\|_R \, \|u_1 - u_2\|_{X_\mathbb{C}} \le \frac{1}{r_n} \|\Phi^{[n]} - \mathrm{Id}\|_{R+r_n} \, \|u_1 - u_2\|_{X_\mathbb{C}} \le \frac{1}{2} \|u_1 - u_2\|_{X_\mathbb{C}}$$

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

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

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 consequence of the *Inverse Function Theorem.* □

Next theorem establishes that the first n terms in the expansions of $\Phi_\theta^{[n]}$ and $G^{[n]}$ in powers of ε, are *independent* of the construction used. This shows that equations (2.7) and (2.6) are "canonical".

Theorem 2.12 [Uniqueness of quasi-solutions to (2.7)] *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*

$$\widehat{\Phi}_0 = \mathrm{Id}, \qquad \|\widehat{\Phi} - \mathrm{Id}\|_{R+r_n} \le \frac{r_n}{2}. \tag{2.24}$$

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

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

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

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

where $\Phi_\theta^{[n]}$ is the function defined in Theorem 2.11 and $C_3(n)$ is a positive constant depending on n.

Proof of Theorem 2.12. The result stems from successive applications of the contraction Lemma 2.10. First, by Lemma 2.9 and starting from (2.24) and (2.16), we have for any $k \leq n$,

$$\| (\Gamma^\varepsilon)^k (\widehat{\Phi}) - \mathrm{Id}\|_{R+r_n-kr_n} \leq \frac{r_n}{2}, \qquad \| (\Gamma^\varepsilon)^k (\Phi^{[n]}) - \mathrm{Id}\|_{R+r_n-kr_n} \leq \frac{r_n}{2}, \tag{2.25}$$

where we used the fact that $4\, C_K\, P\, |\varepsilon| \leq \frac{r_n}{2}$. These estimates allow for the application of Lemma 2.10 (still denoting $a_n = \frac{16 C_K P |\varepsilon|}{r_n}$) and we get, for all for $k \leq n$,

$$\| (\Gamma^\varepsilon)^{k+1} (\widehat{\Phi}) - (\Gamma^\varepsilon)^k (\widehat{\Phi})\|_{R-kr_n} \leq a_n^k \|\Gamma^\varepsilon(\widehat{\Phi}) - \widehat{\Phi}\|_R \leq \frac{1}{2^k} \|\Gamma^\varepsilon(\widehat{\Phi}) - \widehat{\Phi}\|_R$$

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

$$\| (\Gamma^\varepsilon)^{n+1} (\widehat{\Phi}) - \widehat{\Phi}\|_{r_n} \leq \sum_{k=0}^{n} \| (\Gamma^\varepsilon)^{k+1} (\widehat{\Phi}) - (\Gamma^\varepsilon)^k (\widehat{\Phi})\|_{r_n} \leq 2\, \|\Gamma^\varepsilon(\widehat{\Phi}) - \widehat{\Phi}\|_R, \tag{2.26}$$

and similarly for $\Phi^{[n]}$. On the other hand, using (2.25) and applying again $n+1$ times the contraction Lemma 2.10, we obtain

$$\| (\Gamma^\varepsilon)^{n+1} (\widehat{\Phi}) - (\Gamma^\varepsilon)^{n+1} (\Phi^{[n]})\|_{r_n} \leq a_n^{n+1} \|\widehat{\Phi} - \Phi^{[n]}\|_{R+r_n} \leq a_n^{n+1}\, r_n, \tag{2.27}$$

where the last inequality uses (2.24) and the similar estimate for $\Phi^{[n]}$. Finally, from estimate (2.26) on $\widehat{\Phi}$ and the similar bound on $\Phi^{[n]}$, and from (2.27), we deduce

$$\begin{aligned}
&\|\widehat{\Phi} - \Phi^{[n]}\|_{r_n} \\
&\leq \|\widehat{\Phi} - (\Gamma^\varepsilon)^{n+1} (\widehat{\Phi})\|_{r_n} + \|\Phi^{[n]} - (\Gamma^\varepsilon)^{n+1} (\Phi^{[n]})\|_{r_n} + \| (\Gamma^\varepsilon)^{n+1} (\widehat{\Phi}) - (\Gamma^\varepsilon)^{n+1} (\Phi^{[n]})\|_{r_n} \\
&\leq 2\|\widehat{\Phi} - \Gamma^\varepsilon(\widehat{\Phi})\|_R + 2\|\Phi^{[n]} - \Gamma^\varepsilon(\Phi^{[n]})\|_R + 16 C_K P \left(\frac{16 C_K\, P}{R}\,(n+1)\right)^n |\varepsilon|^{n+1} \qquad (2.28) \\
&\leq 2\widehat{C}\, P\, |\varepsilon|^{n+1} + 4 C_K P \Big(2(n+1)|\varepsilon/\varepsilon_0|\Big)^n |\varepsilon| + 16 C_K P \left(\frac{16 C_K\, P}{R}\,(n+1)\right)^n |\varepsilon|^{n+1},
\end{aligned}$$

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 (2.23) on $\delta^{[n]}$ and the assumption on $\widehat{\delta}$. Gathering the various constants gives the result. □

2.3.3 Proof of Theorem 2.7

By optimizing the choice of the parameter n in $(n+1)^n\, |\varepsilon|^{n+1}$, 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 $(n_\varepsilon + 1)|\varepsilon| \le \frac{\varepsilon_0}{4} < \varepsilon_0$, Theorem 2.11 applies with $n = n_\varepsilon$ and we have that

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

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

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

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

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

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 2.11 and the choice of n^ε ensure that the defect $\delta^{[n_\varepsilon]}_\theta$ satisfies

$$\|\delta^{[n_\varepsilon]}\|_R \le 2\, C_K \left(2(n_\varepsilon + 1) \frac{|\varepsilon|}{\varepsilon_0} \right)^{n_\varepsilon} \le 2\, C_K \left(\frac{1}{2}\right)^{n_\varepsilon},$$

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

$$\begin{aligned} \|G^{[n_\varepsilon]}(u) - \widetilde{G}^\varepsilon(u)\|_{X_\mathbb{C}} &= \left\| \sum_{k \ge n_\varepsilon + 1} \frac{\varepsilon^k}{k!} \left.\frac{d^k G^{[n_\varepsilon]}}{d\varepsilon^k}\right|_{\varepsilon=0} (u) \right\|_{X_\mathbb{C}} \\ &\le \sum_{k \ge n_\varepsilon + 1} \frac{\varepsilon^k}{k!} \frac{k!}{\delta^k} \sup_{|\varepsilon| < \delta} \|G^{[n_\varepsilon]}(u)\|_{X_\mathbb{C}} \le \frac{(\varepsilon/\delta)^{n_\varepsilon+1}}{1 - (\varepsilon/\delta)} \sup_{|\varepsilon|<\delta} \|G^{[n_\varepsilon]}\|_R \le 4 C_K \left(\frac{1}{2}\right)^{n_\varepsilon + 1}, \end{aligned}$$

where we have used $\left\| \left(\partial_u \langle \Phi^{[n_\varepsilon]} \rangle\right)^{-1} \right\|_R \le 2$ to bound[7] $G^{[n_\varepsilon]}$ on K_R by $2\, C_K$ and $|\varepsilon|/\delta \le \frac{1}{2}$. To conclude, it remains to write, using $n_\varepsilon \ge (\varepsilon_0/(4|\varepsilon|)) - 2$, that

$$\|\mathcal{R}^\varepsilon\|_R \;\le\; \|G^{[n_\varepsilon]} - \widetilde{G}^\varepsilon\|_R + \left\| \left(\partial_u \Phi^{[n_\varepsilon]}\right)^{-1} \right\|_R \left\|\delta^{[n_\varepsilon]}\right\|_R \le 24\, C_K \exp\left(-\frac{\varepsilon_0}{8|\varepsilon|} \right).$$

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

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

[7] This stems from Lemma 2.9-(i) together with the known estimate $\|\Phi^{[n_\varepsilon]} - \mathrm{Id}\|_{R+r_{n_\varepsilon}} \le r_{n_\varepsilon}/2$ (where $r_{n_\varepsilon} = R/(n_\varepsilon + 1)$), as established in Theorem 2.11.

There exists $T_1 > 0$ such that $\widetilde{\Psi}_t^\varepsilon(u_0)$ is well-defined for all $0 \leq t \leq T_1/\varepsilon$, given that $\widetilde{G}^\varepsilon$ 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 $\tilde{u}^\varepsilon(t) = \widetilde{\Phi}_t^\varepsilon \circ \widetilde{\Psi}_t^\varepsilon(u_0)$

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

as long as $\tilde{u}^\varepsilon$ and $(\widetilde{\Phi}_t^\varepsilon)^{-1}(\tilde{u}^\varepsilon(t))$ 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

$$\|u^\varepsilon(t) - \tilde{u}^\varepsilon(t)\|_{X_\mathbb{C}} \leq \frac{3}{2}\|\mathcal{R}^\varepsilon\|_R \frac{e^{|\varepsilon|Lt} - 1}{L} \leq 36\, R\, e^{|\varepsilon|Lt - \frac{\varepsilon_0}{8|\varepsilon|}} := M(t,\varepsilon)$$

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

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 $|\varepsilon| < \varepsilon_2$ with ε_2 small enough so that $M(T/|\varepsilon|^{1+\alpha}, \varepsilon) \leq R - r_{n_\varepsilon}$, ensures that $\tilde{u}^\varepsilon$ and $(\widetilde{\Phi}_t^\varepsilon)^{-1}(\tilde{u}^\varepsilon(t))$ remain in K_R (see Theorem 2.11 (iii)) whenever $t \leq T/|\varepsilon|^{1+\alpha}$ (hence $T_1 \geq T/|\varepsilon|^\alpha$). Now we claim that, with the choice

$$0 < \varepsilon < \min\left(\varepsilon^*, \left(\frac{\varepsilon_0^2 P}{2T}\right)^{\frac{1}{1-\alpha}}, \frac{\varepsilon_0}{80}\right), \tag{2.29}$$

(which means $0 < \varepsilon < \min(\varepsilon^*, \varepsilon_2)$ in the case $\alpha = 0$), we obtain estimate (2.21).

Proof of the claim. For $|\varepsilon|^{1-\alpha} \leq \varepsilon_0^2 \frac{P}{2T}$, we have $e^{|\varepsilon|Lt - \frac{\varepsilon_0}{8|\varepsilon|}} \leq e^{-\frac{\varepsilon_0}{16|\varepsilon|}}$ on $[0, T/\varepsilon^{1+\alpha}]$, so that for all $0 \leq t \leq T/|\varepsilon|^{1+\alpha}$

$$M(t,\varepsilon) \leq 36\, R\, e^{-\frac{\varepsilon_0}{16|\varepsilon|}}.$$

The quantity $M(T/|\varepsilon|^{1+\alpha}, \varepsilon)$ 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

$$|\varepsilon| < \frac{\varepsilon_0}{16} \frac{1}{\log 72}.$$

A combined bound on ε is given by (2.29), a condition under which

$$\forall t \in [0, T/\varepsilon^{1+\alpha}], \qquad \|u^\varepsilon(t) - \tilde{u}^\varepsilon(t)\|_{X_\mathbb{C}} \leq 36\, R\, e^{-\frac{\varepsilon_0}{16\,|\varepsilon|}}.$$

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

2.4 The linear case

In this section, we consider the case $g_\theta(u) \equiv A_\theta u$, where A_θ is a bounded linear operator on X. The initial value problem then reads

$$\begin{aligned} \frac{d}{dt} u^\varepsilon(t) &= \varepsilon A_t\, u^\varepsilon(t), \qquad & u^\varepsilon(t) \in X, \\ u^\varepsilon(0) &= u_0, & u_0 \in X. \end{aligned} \tag{2.30}$$

In this situation, it turns out the change of variable solution of (2.7) can be exactly constructed, due to the fact that our iterative procedure actually converges. Naturally, Assumption 2.3 has to be replaced here by the following:

Assumption 2.13 *The map $\theta \mapsto A_\theta \in \mathcal{L}(X)$ is continuous and P-periodic.*

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 2.14 [Exact averaging in the linear case] *Consider $u^\varepsilon(t)$ the solution of* (2.30) *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 $0 \le \varepsilon < \varepsilon_l$, there exists a map $\theta \mapsto \Phi^\varepsilon_\theta \in \mathcal{L}(X)$, such that*

(i) The function Φ^ε_θ is P-periodic and C^1 in θ, and satisfies $\Phi_{\theta=0} = \mathrm{Id}$.

(ii) For all $\theta \in \mathbb{T}$, the operator Φ^ε_θ is invertible in $\mathcal{L}(X)$.

(iii) For all $u_0 \in X$, the solution of (2.30) *admits the factorized form*

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

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

$$G^\varepsilon = \langle \Phi^\varepsilon \rangle^{-1} \langle A \Phi^\varepsilon \rangle. \tag{2.31}$$

Sketch of Proof. In the linear framework, equation (2.7) becomes

$$\frac{d\Phi^\varepsilon_\theta}{d\theta} + \varepsilon \Phi^\varepsilon_\theta \langle \Phi^\varepsilon \rangle^{-1} \langle A \Phi^\varepsilon \rangle = \varepsilon A \Phi^\varepsilon_\theta,$$

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

$$\Gamma^\varepsilon(\Phi)_\theta = \mathrm{Id} + \varepsilon \int_0^\theta \big(A_\xi \Phi_\xi - \Phi_\xi \langle \Phi \rangle^{-1} \langle A \Phi \rangle \big)\, d\xi. \tag{2.32}$$

We now prove by induction that the sequence $\Phi^{[0]}_\theta = \mathrm{Id}$, with $\Phi^{[n+1]}_\theta = \Gamma^\varepsilon(\Phi^{[n]})_\theta$, $n \in \mathbb{N}$, satisfies the estimate[8]

$$d_n := \|\Phi^{[n]} - \mathrm{Id}\| < d_* := \frac{1}{2} - \varepsilon\alpha - \sqrt{\varepsilon^2\alpha^2 - 3\varepsilon\alpha + \frac{1}{4}}. \tag{2.33}$$

[8]Note that the term $\varepsilon^2\alpha^2 - 3\varepsilon\alpha + \frac{1}{4}$ is positive due to the assumption $\varepsilon < \varepsilon_l$.

Assume that (2.33) holds for some n. From $0 < \varepsilon < \varepsilon_l$ and the induction assumption, we first deduce $d_n < d_* < 1$, so that $\Phi_\theta^{[n]}$ is invertible for all θ and $(\Phi_\theta^{[n]})^{-1} \in C(\mathbb{T}, \mathcal{L}(X))$. In particular, $\Phi_\theta^{[n+1]} = \Gamma^\varepsilon(\Phi^{[n]})_\theta$ is well-defined. Moreover, from (2.32), we estimate

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

and as a consequence, $d_{n+1} \leq 2\varepsilon\alpha \, \dfrac{1 + d_n}{1 - d_n}$. Studying $f(d) = 2\varepsilon\alpha \frac{1+d}{1-d}$, it is easy to prove that $d_{n+1} \leq f(d_n) < d_*$ and (2.33) follows. It remains to to prove that the Lipschitz constant of Γ^ε is less than one on the domain $\{\Phi \text{ s.t. } \|\Phi - \mathrm{Id}\| \leq d^*\}$. Considering $\Phi \in C(\mathbb{T}, \mathcal{L}(X))$ and $\widehat{\Phi} \in C(\mathbb{T}, \mathcal{L}(X))$ with $\|\Phi - \mathrm{Id}\| < d_*$ and $\|\widehat{\Phi} - \mathrm{Id}\| < d_*$, both Φ_θ and $\widehat{\Phi}_\theta$ are invertible on X, with $\|\Phi^{-1}\| \leq 1/(1 - d^*)$ and $\|\widehat{\Phi}^{-1}\| \leq 1/(1 - d^*)$. Hence, as in Lemma 2.10, we write

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

Denoting d^* the largest root of $d(1-d) - 2\varepsilon\alpha(1+d)$, we have $d^* d_* = 2\varepsilon\alpha$ and $d^* + d_* = 1 - 2\varepsilon\alpha$, so that $\frac{4\varepsilon\alpha}{(1-d_*)^2} < \frac{4\varepsilon\alpha}{(1-d_*)(1-d^*)} = 1$ and the convergence of $\Phi^{[n]}$ in $C(\mathbb{T}, \mathcal{L}(X))$ follows. $\square$

3 Geometric aspects

One of the advantages of stroboscopic averaging is that it preserves the geometric properties of the initial equation: both the Hamiltonian structure and the invariants of the original equation (if any) are inherited.

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 $J : X \to X$, 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*

$$\forall (\theta, u, v) \in \mathbb{T} \times K \times X, \qquad (\partial_u H_\theta)(u)\, v = (J g_\theta(u), v)_Z\,. \tag{3.1}$$

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

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

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*

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

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

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

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*

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

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 2.11 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* 2.11 *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 $0 \le k \le n$, i.e. that for all $(\theta, u) \in \mathbb{T} \times K$ and $v, w \in X$*

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

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

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

with Hamiltonian

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

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

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

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

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

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

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

These results eventually provide the relation

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

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 2.11. This provides, by picking up $v \in X$ and taking the scalar product with $J\partial_u \Phi_\theta(u)\, v$, the relation

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

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

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

This is where assumption $(\mathcal{S}_k)$ is used. It provides, using $k+1 \leq n+1$,

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

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 2.12) gives

$$\partial_u H^{[n]}(u)\, v = \big(JG^{[n]}(u)\,,\; v\big)_Z + \mathcal{O}(\varepsilon^{k+1}\|v\|_X).$$

The proof of the lemma is complete. □

Lemma 3.4 allows to establish the

Theorem 3.5 [Stroboscopic averaging preserves the Hamiltonian structure]
Under the assumptions of Theorem 2.11 *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)*),*

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

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

Remark 3.6 *Consider the truncated Hamiltonian*

$$\widetilde{H}^{[n]}(u) = \sum_{k=0}^{n} \frac{\varepsilon^k}{k!}\, \frac{d^k H^{[n]}}{d\varepsilon^k}\bigg|_{\varepsilon=0}(u). \tag{3.7}$$

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 is exactly *Hamiltonian, namely*

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

(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:

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

We claim that

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

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

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

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

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

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 2.11 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

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

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. □

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 $Q : \mathbb{T}\times X \to \mathbb{R}$, 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

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

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 2.11*, 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* $u \in K$*,* $\theta \in \mathbb{T}$*,*

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

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* $G^{[n+1]} - \widetilde{G}^{[n]} = \mathcal{O}(\varepsilon^{n+1})$ *and remarking that* $\widetilde{G}^{[n]}$ *is a polynomial of degree* n *in* ε*, one deduces from* (3.11) *that we have*

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

so that Q_0 *is* exactly *preserved by the autonomous equation* (2.20)*.*

Proof of Theorem 3.7. Relation (2.22), 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

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

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

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

In particular, taking averages on both sides yields

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

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 $\Phi^{[n]} = \mathrm{Id} + \mathcal{O}(\varepsilon)$ 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. □

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

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

Here f is a real-analytic function. We set $Z = L^2(\Omega)$ where $\Omega \subset \mathbb{R}^d$ is open, and $A : D(A) \subset Z \to Z$ 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 $V(x) \to +\infty$ as $|x| \to \infty$ 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

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

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 $G: \mathbb{C} \to \mathbb{C}$ satisfying $G(0) = 0$, there exists a nondecreasing C^1 function $C_G : \mathbb{R}_+ \to \mathbb{R}_+$ such that, for all $u \in X$, we have

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

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

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

with periodic boundary conditions. In this case we set $A = -\Delta_x$ with domain

$$D(A) = \left\{u \in H^2(\mathbb{T}_a^d) \text{ s.t. } u|_{x_j=0} = u|_{x_j=a} \text{ for } j = 1, \ldots, d\right\},$$

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 $A : D(A) \to Z = L^2(\mathbb{T}_a^d; \mathbb{C})$ is self-adjoint non-negative with compact resolvent, the embedding $D(A) \subset Z$ is dense, and the spectrum of A is

$$\sigma(A) = \left\{(2\pi/a)^2|k|^2 = (2\pi/a)^2\left(k_1^2 + \cdots + k_d^2\right); \, k \in \mathbb{Z}^d\right\} \subset (2\pi/a)^2\mathbb{N}.$$

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

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

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

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

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

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

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

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

In this case we set $A = -\Delta + \omega^2|x|^2 - d\omega$, with domain

$$D(A) = \left\{u \in L^2(\mathbb{R}^d) \text{ s.t. } Au \in L^2(\mathbb{R}^d)\right\} = \left\{u \in H^2(\mathbb{R}^d) : |x|^2 u \in L^2(\mathbb{R}^d)\right\}.$$

The operator $A : D(A) \to Z = L^2(\mathbb{R}^d;\mathbb{C})$ 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

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

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

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

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

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

The same argument shows that X is an algebra. □

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 $T^\varepsilon_{\max} \in]0,+\infty]$ and a unique maximal solution $\psi^\varepsilon \in C([0,T^\varepsilon_{\max}[,X)$ to* (4.1). *This solution is maximal in the sense that*

$$\textit{if } T^\varepsilon_{\max} < +\infty, \ \textit{then} \ \limsup_{t\to T^\varepsilon_{\max}} \|\psi^\varepsilon(t)\|_{L^\infty} = +\infty.$$

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

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

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

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

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

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

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

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

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

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

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

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

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 $t \le T^\varepsilon_\kappa/\varepsilon$ we recover

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

and the Gronwall lemma asserts, whenever $0 \le t \le T^\varepsilon_\kappa/\varepsilon$, the estimate

$$\|\psi^\varepsilon(t)\|_X \le \|\psi_0\|_X\, e^{T^\varepsilon_\kappa C_{\tilde f}(c_0\kappa\|\psi_0\|_X)}.$$

This in turn ensures

$$T^\varepsilon_\kappa \ge T_\kappa := \frac{\log\kappa}{C_{\tilde f}\left(c_0\kappa\|\psi_0\|_X\right)}. \tag{4.13}$$

□

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

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

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 $Ay^\varepsilon \equiv \begin{pmatrix} Ap^\varepsilon \\ Aq^\varepsilon \end{pmatrix}$, 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 $\theta \to e^{\theta J^{-1} A}$ is periodic (with period $2\pi/\lambda$). Consider now the function

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

Inserting (4.15) in (4.14) immediately yields

$$\partial_t u^\varepsilon(t) = \varepsilon g_t\left(u^\varepsilon(t)\right), \qquad u^\varepsilon(0) = u_0, \tag{4.16}$$

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

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

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 $\kappa - 1 > 0$)

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

□

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 $\widetilde{f}(z) = f(z^2)z$. The function $\widetilde{f}$ 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

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

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

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

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

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

and the assumed fact that X is an algebra, together with the fact that $\widetilde{f}$ 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 $u \mapsto \widetilde{f}(u)$ is analytic on K_{2R} in the sense of Definition 2.2. Moreover, we remark that for all $\theta \in \mathbb{T}$ and $u \in K_{2R}$,

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

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 $\theta \in \mathbb{T}$, $u \in K_{2R}$,

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

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 $F(u) = \displaystyle\int_0^u f(v)dv$, it is clear that J is skew-symmetric with respect to the scalar product on $Z \times Z$, and one can check the identity

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

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 $\Phi^\varepsilon_\theta(u)$, analytic in $u \in X$, continuously differentiable and periodic in $\theta \in \mathbb{T}$ with $\Phi^\varepsilon_0 = \mathrm{id}$, 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*

$$\sup_{0 \leq t \leq \frac{T}{\varepsilon}} \left\| \psi^\varepsilon(t) - e^{-itA} \Phi^\varepsilon_t \left(\widetilde{\psi}^\varepsilon(t) \right) \right\|_X \leq C \exp\left(-\frac{C^{-1}}{\varepsilon} \right), \tag{4.21}$$

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

$$\frac{\partial \widetilde{\psi}^\varepsilon}{\partial t} = \varepsilon G^\varepsilon(\widetilde{\psi}^\varepsilon), \qquad \widetilde{\psi}^\varepsilon(0) = \psi_0. \tag{4.22}$$

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

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

and the almost conservation of energy

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

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 $G^\varepsilon = \widetilde{G}^{[n_\varepsilon]}$ and $\Phi^\varepsilon_\theta = \Phi^{[n_\varepsilon]}_\theta$ (and n_ε is as in Theorem 2.7). In that context, the Hamiltonian H^ε is the function $\widetilde{H}^{[n_\varepsilon]}$ given according to Remark 3.6. Lastly, the original equation (4.16) preserves the mass $\|u\|^2_{L^2}$ as well as the filtered energy

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

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. □

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 writen 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*

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

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

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

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 Numerical experiments

In this section, we present numerical simulations based on stroboscopic averaging for some specific solutions to NLS found in the literature. Interesting physical informations are illustrated on these examples.

5.1 A numerical counterpart of stroboscopic averaging

A numerical method can not rely on the analytical computation of such terms as those involved in G^ε and this rules out the direct approximation of (2.2). In order to approach $G^\varepsilon(u)$ at a given point $u \in X$, we first use the group property of Ψ_t^ε to assert that

$$\varepsilon G^\varepsilon(u) = \frac{d}{dt}\Psi_t^\varepsilon(u)\bigg|_{t=0}$$

and then interpolate the derivative of $\Psi^\varepsilon(t,u)$ (say at order 2 for ease of presentation)

$$G^\varepsilon(u) \approx \frac{1}{2P\varepsilon}\left(\Psi^\varepsilon_P(u) - \Psi^\varepsilon_{-P}(u)\right) = G_1(u) + \varepsilon G_2(u) + \mathcal{O}(\varepsilon^2). \tag{5.1}$$

To complete the procedure, it remains to use the fact that

$$\Phi^\varepsilon_P(\Psi^\varepsilon_P(u)) = \Psi^\varepsilon_P(u) \quad \text{and} \quad \Phi^\varepsilon_{-P}(\Psi^\varepsilon_{-P}(u)) = \Psi^\varepsilon_{-P}(u),$$

a consequence of the stroboscopic property. We finally approximate $\Phi^\varepsilon_P(\Psi^\varepsilon_P(u))$ and $\Phi^\varepsilon_{-P}(\Psi^\varepsilon_{-P}(u))$ by solving the equations

$$\dot{U}^\varepsilon = \varepsilon g_t(U^\varepsilon),\ t \in [0,P],\ U^\varepsilon(0) = u \ \text{and}\ \dot{U}^\varepsilon = \varepsilon g_t(U^\varepsilon),\ t \in [-P,0],\ U^\varepsilon(0) = u, \tag{5.2}$$

by a standard one-step method S^ε_h (Strang splitting here) where the step size h used is small enough to resolve one oscillation, i.e. $h = P/n$ with $n \in \mathbb{N}$. The outcome of this procedure is a *micro-macro* algorithm (called SAM for Stroboscopic Averaging Method) which computes a sequence of approximations to the averaged solution of (2.2).

5.2 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 5.1 *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]). The first-order averaged[9] Hamiltonian H_1 can be computed as

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

[9]Other terms can be obtained by a formal Magnus expansion (see for instance the preprint version of this paper).

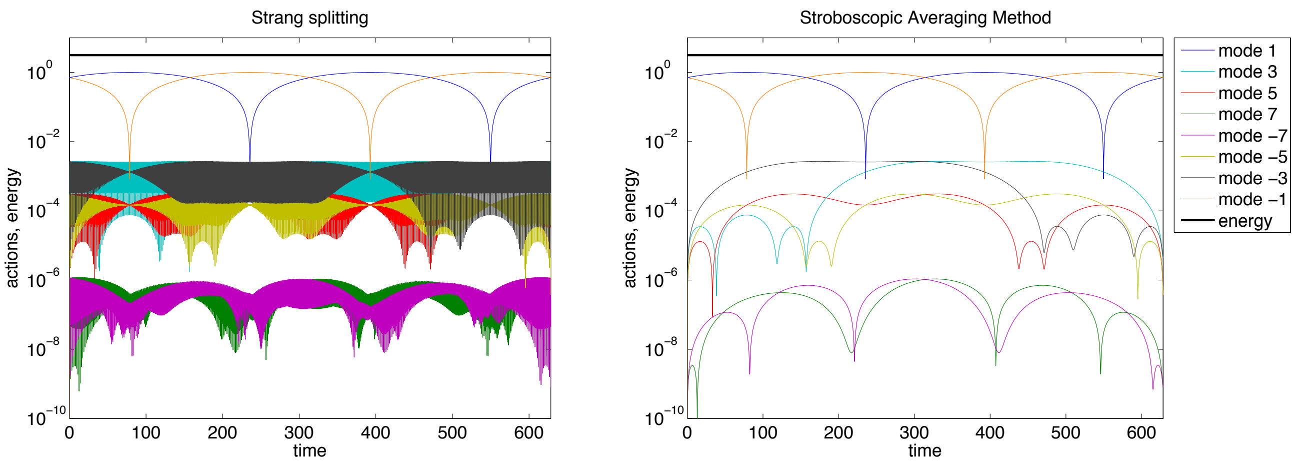

Figure 1: Example 1 with the Strang splitting method and the Stroboscopic Averaging Method.

The proof of Theorem 5.1 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:

- The original system integrated by Strang splitting with tiny stepsizes (see left of Fig. 1).
- The original system integrated by SAM (see right of Fig. 1).
- 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 5.1, 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.

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 (2.20)

$$\frac{dU}{dt} = \varepsilon \widetilde{G}^{\varepsilon}(U) = \varepsilon G_1(U) + \varepsilon^2 G_2(U) + \varepsilon^3 G_3(U) + \cdots .$$

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

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

We observe that Fig. 2 left and Fig. 2 right coincide up to remainder 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 remainder 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).

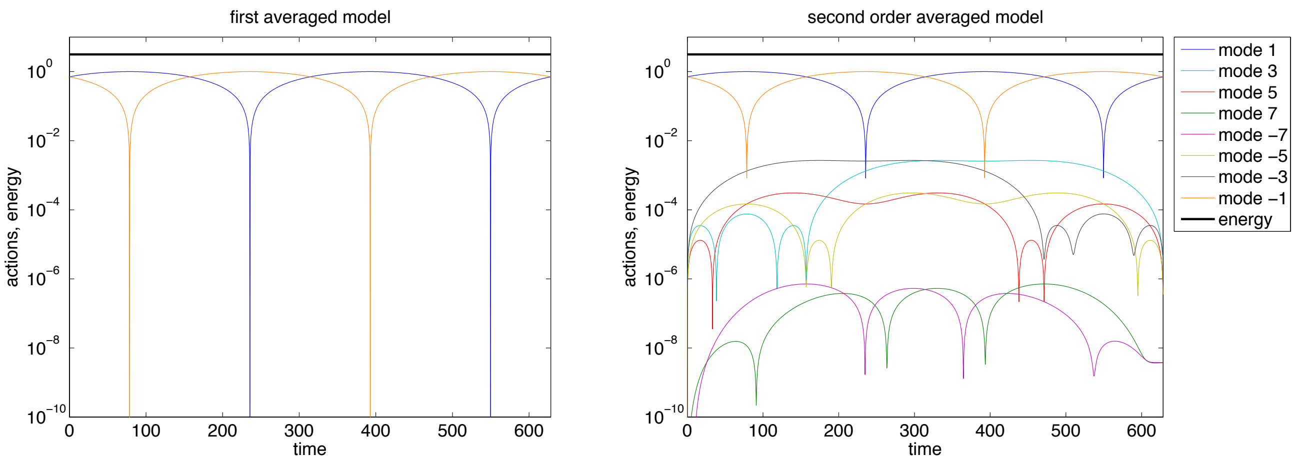

Figure 2: Example 1, first and second order averaged models.

5.3 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 $|\psi|^4\psi$:

$$i\partial_t \psi^\varepsilon = -\Delta\psi^\varepsilon + \varepsilon|\psi^\varepsilon|^4\psi^\varepsilon, \quad t \geq 0, \quad \psi^\varepsilon(t,\cdot) \in H^1(\mathbb{T}_{2\pi}) \tag{5.3}$$

$$\psi(0,x) = \sqrt{\kappa_0}\, e^{ix} + \sqrt{\frac{\kappa_0}{2}}\, e^{-ix} + \sqrt{\kappa_0}\, e^{5ix} + \sqrt{\frac{1-\kappa_0}{2}}\, e^{7ix} \tag{5.4}$$

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 5.2 *There exist $\widetilde{P} > 0$, $\widetilde{\varepsilon}_0 > 0$ and a $2\widetilde{P}$-periodic function $\kappa : \mathbb{R} \to]0;1[$ such that if $0 < \varepsilon < \varepsilon_0$, the solution to* (5.3) *satisfies for all $0 \leq t \leq \varepsilon^{-3/2}$*

$$\psi^\varepsilon(t,x) = \sum_{k=-1,1,5,7} \xi_j(t)e^{ikx} + \varepsilon^{1/4}q_1(t,x) + \varepsilon^{3/2}tq_2(t,x)$$

with

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

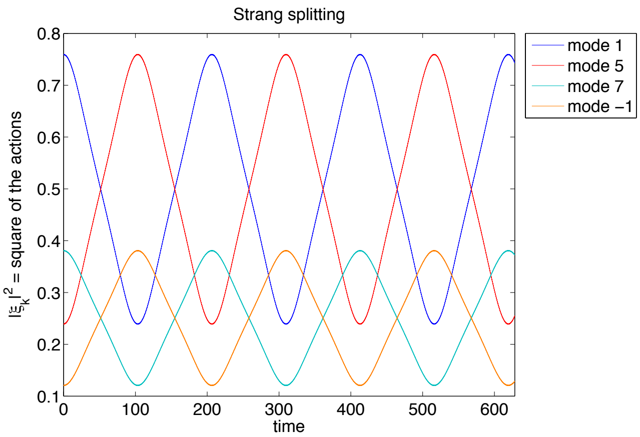

Figure 3: Example 2, beating effect between modes 1, 5 ,7 and -1.

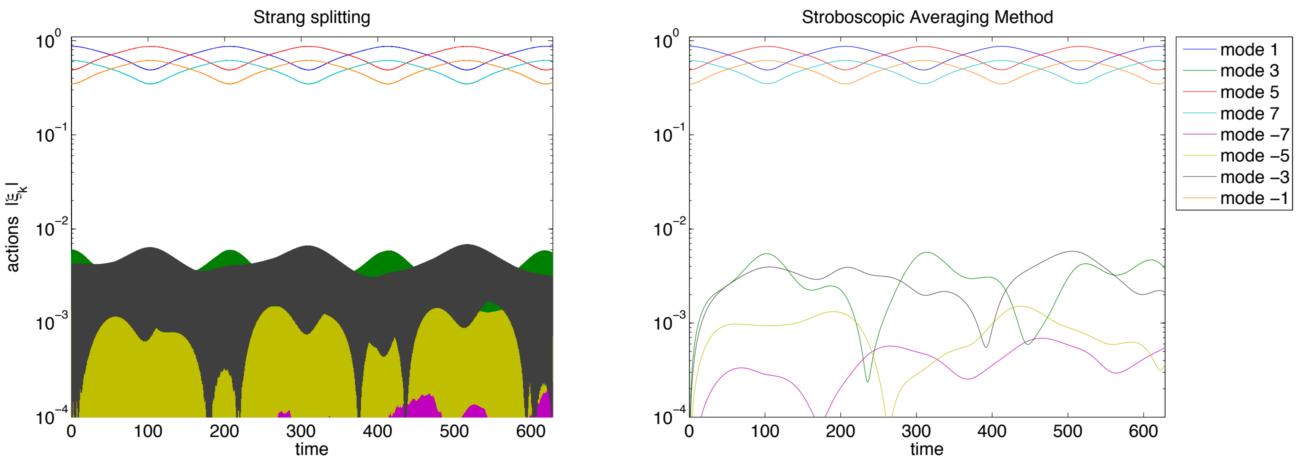

Figure 4: Example 2 with the Strang splitting method and the Stroboscopic Averaging Method.

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 5.2.

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 by formulated. As a matter of facts, 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 is here 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 $\mathcal{O}(\varepsilon)$.

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

The problem considered by Carles and Faou in [CF12] (Example 3) involves a cubic nonlinearity $|\psi|^2\psi$ and is posed in the two-dimensional torus $\mathbb{T}^2_{2\pi} = [0, 2\pi] \times [0, 2\pi]$:

$$\begin{aligned} i\partial_t \psi^\varepsilon &= -\Delta\psi^\varepsilon \pm \varepsilon|\psi^\varepsilon|^2\psi^\varepsilon, \quad t \geq 0, \quad \psi^\varepsilon(t, \cdot) \in H^1(\mathbb{T}^2_{2\pi}) && (5.5) \\ \psi^\varepsilon(0, x) &= 1 + 2\cos(x_1) + 2\cos(x_2). && (5.6) \end{aligned}$$

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.

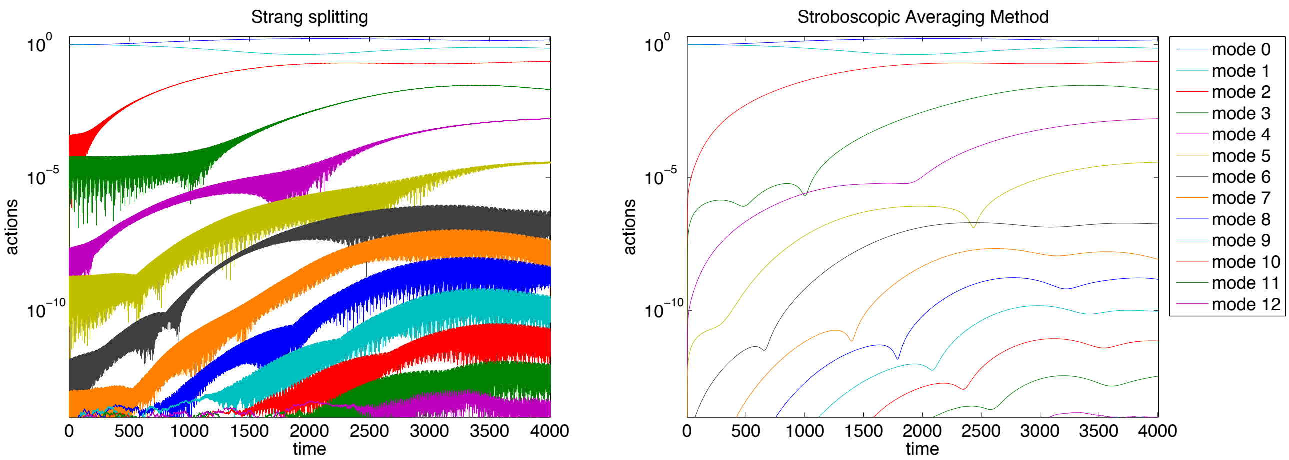

Figure 5: Example 3 with the Strang splitting method and the Stroboscopic Averaging Method.

Acknowledgments

F. Castella, P. Chartier and F. Méhats were partially supported by the French National Research Agency - ANR under project Lodiquas. Ander Murua was partially supported by Ministerio de Ciencia e Innovacin (Spain) under project MTM2010-18246-C03-03 and by the Basque Government under the Consolidated Research Group IT649-13.

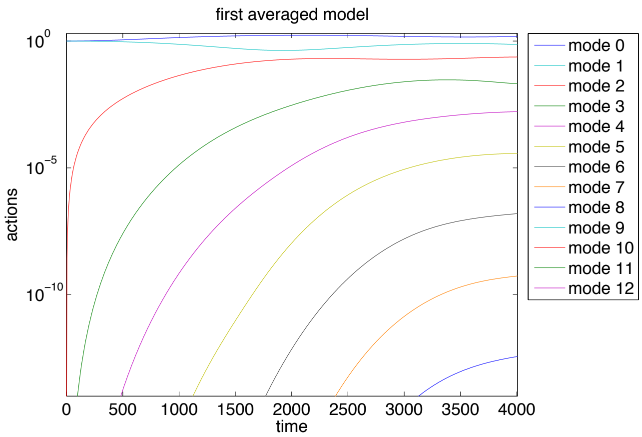

Figure 6: Example 3, first order averaged model.

References

[AM78] D. G. Andrews and M. E. McIntyre. On wave-action and its relatives. *J. Fluid Mech.*, 89(4):647–664, 1978.

[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.

[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.

[CHL08] 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.

[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.

[CMSS13a] P. Chartier, A. Murua, and J.M. Sanz-Serna. Higher-order averaging, formal series and numerical integration iii: Error bounds. *Foundations of Computational Mathematics*, pages 1–22, 2013.

[CMSS13b] 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, 2013.

[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.

[HLW06] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

[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.

[MS03a] Jerrold E. Marsden and Steve Shkoller. The anisotropic Lagrangian averaged Euler and Navier-Stokes equations. *Arch. Ration. Mech. Anal.*, 166(1):27–46, 2003.

[MS03b] Karsten Matthies and Arnd Scheel. Exponential averaging for Hamiltonian evolution equations. *Trans. Amer. Math. Soc.*, 355(2):747–773 (electronic), 2003.

[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.

[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.