An averaging theory for nonlinear partial differential equations

Guan Huang

ÉCOLE POLYTECHNIQUE

CENTRE DE MATHÉMATHIQUES LAURENT SCHWARTZ

École Doctorale de l'École Polytechnique

Thèse de doctorat

Discipline : Mathémathiques

présentée par

HUANG Guan

Une théorie de la moyenne pour les équations aux dérivées partielles nonlinéaires

---

dirigée par Sergei KUKSIN

Soutenue le 4 Juin 2014 devant le jury composé de :

M. Jacques Féjoz	Université Paris IX	
M. Raphaël Krikorian	Université Paris VI	
M. Sergei Kuksin	Université Paris VII	Directeur
M. Pierre Lochak	Université Paris VI	
M. Laurent Niederman	Université Paris XI	
M. Thierry Paul	École Polytechnique	
M. Nikolay Tzvetkov	Université de Cergy-Pontoise	Rapporteur

Rapporteur non présent à la soutenance :
M. Dario Bambusi   Università degli studi di Milano

Thèse présentée pour obtenir
la garde de docteur
de l'École Polytechnique

CMLS - UMR 7640
École Polytechnique
91 128 Palaiseau cedex
France

To see a world in a grain of sand
And a heaven in a wild flower
Hold infinity in the palm of your hand
And eternity in an hour
——William Blake

OK! GO!

Remerciements

Il y a beaucoup de personnes à qui je suis reconnaissant...

Tout d'abord, je tiens à exprimer toute ma gratitude à mon directeur de thèse, Monsieur Sergei Kuksin, pour tout ce qu'il a apporté dans ma vie dès le premier jour où il m'a connu. Il m'a proposé des sujets de recherche avec ses expériences. Il a été toujours volontaire et disponible à répondre toutes mes questions. Il a partagé avec générosité ses connaissances et idées mathématiques et son savoir culturel. Il a montré une très grande patience pendant de nombreuses heures de discussions, qui ont été essentielles pour des étapes de cette thèse. Il a consacré beaucoup de temps à lire et relire, à noter et commenter un grand nombre de mes textes mathématiques et non mathématiques. Il a montré une grande tolérance quand j'ai fait des erreurs. Il m'a encouragé beaucoup quand j'étais déçu. Je lui suis dont très sincèrement reconnaissant, et par toutes les experiences que j'ai eues durant les quatre dernières années, je me sens heureux que j'ai eu la grande chance et l'honneur de pouvoir faire mes etudes et mes recherches en mathématiques avec lui.

Je tiens à remercier mon directeur de master, Monsieur Cheng Chongqing, qui m'a conduit au domaine des systèmes dynamiques et m'a recommandé à Sergei. Sa gentillesse et son enthousiasme pour les mathématiques m'ont beaucoup influencé.

Je suis très privilégié que Dario Bambusi et Nikolay Tzvetkov aient accepté d'être rapporteurs de cette thèse. Je les remercie profondément pour leur travails précieux.

Je remercie chaleureusement Jacques Féjoz, Pierre Lochak, Raphaël Krikorian, Laurent Niederman et Thierry Paul de m'avoir fait l'honneur de faire partie du jury.

Merci à Benoît Grébert pour m'avoir invité à faire exposé dans leur séminaire. Merci à Anatoly Neishtadt pour m'avoir invité à le rencontrer et pour nos plusieurs discussions.

C'est grace au programme de Chinese Scholar Council que j'ai eu l'occasion de poursuivre mes etudes en France. Je suis heureux de pouvoir exprimer ici ma gratitude à tous ceux qui y on participe.

Cette thèse est effectuée au sein du Centre Mathématiques Laurent Schwartz, École Polytechnique. J'en remercie tous les membres pour y ont créé des excellents conditions scientifiques, culturelles et matérielles, tout particulièrement, nos secrétaires travailleuses et gentilles, Michèle Lavalette, Pascale Fuseau et Marine Aimer. Pendant les deux dernières années, j'ai aussi travaillé très souvent dans le bâtiment Sophie Germain de Paris 7. J'en remercie pour son hospitalité.

Merci aux mes amis et co-organisateurs de notre group de travail : Kai Jiang, Qiaoling Wei et Lei Zhao, pour les nombreuses lectures que nous avons organisés et le temps que nous avons passé ensemble. Merci à Alexandre pour m'avoir donné des conseils divers. Merci à Xin Nie, Shanwen Wang et Shaoshi Chen pour m'avoir aidé quand je commençais mon séjour en France. Merci à Marcel et Sasha, pour les plusieurs heures de discussion qu'on a passé ensemble.

J'adresse également des remerciements à mes amis pour partager des bonheurs et des peines de la vie : Jiatu Cai, Xin Fang, Haibiao Huang, Benben Liao, Jyunao Lin, Chunhui Liu, Zhengfang Wang, ling Wang, Junyi Xie, Kang Xu, Yao Xu, Jinxin Xue, Huafeng Zhang, Zhiyan Zhao, Qi Zhou, etc.

Tout Particulièrement, j'exprime ma gratitude du fond du coeur à toutes mes familles, surtout mes parents, pour me soutenir toujours. Je ne pourrai jamais exprimer tout ce que je leur dois.

Enfin, merci à Jiani.

RÉSUMÉ

Résumé

Cette thèse se consacre aux études des comportements de longtemps des solutions pour les EDPs nonlinéaires qui sont proches d'une EDP linéaire ou intégrable hamiltonienne. Une théorie de la moyenne pour les EDPs nonlinéaires est presenté. Les modèles d'équations sont les équations Korteweg-de Vries (KdV) perturbées et quelques équations aux dérivées partielles nonlinéaires faiblement.

Considère une équation KdV perturbée sur le cercle :

$$u_t + u_{xxx} - 6uu_x = \epsilon f(u)(x), \quad x \in \mathbb{T} = \mathbb{R}/\mathbb{Z}, \int_{\mathbb{T}} u(x,t)dx = 0, \tag{$*$}$$

où la perturbation nonlinéaire définit les opérateurs analytiques $u(\cdot) \mapsto f(u(\cdot))$ dans les espaces de Sobolev suffisamment lisses. Soit $I(u) = (I_1(u), I_2(u), \cdots) \in \mathbb{R}_+^\infty$ le vecteur formé par les intégrales de KdV, calculé pour le potentiel $u(x)$. Supposons que l'équation $(*)$ satisfait des hypothèses modérées supplémentaires et possède une mesure μ qui est ϵ-quasi-invariante. Soit $u^\epsilon(t)$ est une solution. Il est ici obtenue que sur des intervalles de temps de l'ordre ϵ^{-1}, ses actions $I(u^\epsilon(t,\cdot))$ peuvent être estimés par des solutions d'une certaine équation en moyenne bien posée, à condition que la donnée initiale $u^\epsilon(0)$ est μ-typique et que le ϵ est assez petit.

Considère une EDP nonlinéaire faiblement sur le tore :

$$\frac{d}{dt}u + i(-\triangle u + V(x)u) = \epsilon \mathcal{P}(\triangle u, \nabla u, u, x), \quad x \in \mathbb{T}^d. \tag{$**$}$$

Soient $\{\zeta_1(x), \zeta_2(x), \ldots\}$ les L_2-bases formées par les fonctions propres de l'opérateur $-\triangle + V(x)$. Pour une fonction complexe $u(x)$, on l'écrit comme $u(x) = \sum_{k\geqslant 1} v_k \zeta_k(x)$ et définit $I(u) = (I_k(u), k \geqslant 1)$, où $I_k(u) = \frac{1}{2}|v_k|^2$. Alors, pour toutes les solutions $u(t,x)$ de l'équation linéaire $(**)_{\epsilon=0}$, on a $I(u(t,\cdot)) = const$. Dans cette thèse, il est prouvé que si $(**)$ est bien posée sur des intervalles de temps $t \lesssim \epsilon^{-1}$ et satisfait-il des hypothèses a-priori bénins, alors pour tout ses solutions $u^\epsilon(t,x)$, le comportement limité de la courbe $I(u^\epsilon(t,\cdot))$ sur des intervalles de temps de l'ordre ϵ^{-1}, comme $\epsilon \to 0$, peut être caractérisée uniquement par des solutions d'une certaine équation efficace bien posée.

Mots-clefs

KdV, NLS, EDPs nonlinéaires faiblement, L'équation en moyenne, L'équation efficace.

An averaging theory for nonlinear PDEs

Abstract

This Ph.D thesis focuses on studying the long-time behavior of solutions for non-linear PDEs that are close to a linear or an integrable Hamiltonian PDE. An averaging theory for nonlinear PDEs is presented. The model equations are the perturbed Korteweg-de Vries (KdV) equations and some weakly nonlinear partial differential equations.

Consider a perturbed KdV equation on the circle:

$$u_t + u_{xxx} - 6uu_x = \epsilon f(u)(x), \quad x \in \mathbb{T} = \mathbb{R}/\mathbb{Z}, \ \int_{\mathbb{T}} u(x,t)dx = 0, \tag{$*$}$$

where the nonlinear perturbation defines analytic operators $u(\cdot) \mapsto f(u)(\cdot)$ in sufficiently smooth Sobolev spaces. Let $I(u) = (I_1(u), I_2(u), \cdots) \in \mathbb{R}^{\infty}_{+}$ be the vector, formed by the KdV integrals of motion, calculated for the potential $u(x)$. Assume that the equation $(*)$ has an ϵ-quasi-invariant measure μ and satisfies some additional mild assumptions. Let $u^{\epsilon}(t)$ be a solution. Then it is obtained here that on time intervals of order ϵ^{-1}, its actions $I(u^{\epsilon}(t,\cdot))$ can be approximated by solutions of a certain well-posed averaged equation, provided that the initial datum is μ-typical and that the ϵ is small enough.

Consider a weakly nonlinear PDE on the torus:

$$\frac{d}{dt}u + i(-\triangle u + V(x)u) = \epsilon \mathcal{P}(\triangle u, \nabla u, u, x), \quad x \in \mathbb{T}^d. \tag{$**$}$$

Let $\{\zeta_1(x), \zeta_2(x), \ldots\}$ be the L_2-basis formed by eigenfunctions of the operator $-\triangle + V(x)$. For any complex function $u(x)$, write it as $u(x) = \sum_{k\geqslant 1} v_k \zeta_k(x)$ and set $I_k(u) = \frac{1}{2}|v_k|^2$. Then for any solution $u(t,x)$ of the linear equation $(**)_{\epsilon=0}$ we have $I(u(t,\cdot)) = const$. In this thesis it is proved that if $(**)$ is well posed on time-intervals $t \lesssim \epsilon^{-1}$ and satisfies there some mild a-priori assumptions, then for any its solution $u^{\epsilon}(t,x)$, the limiting behavior of the curve $I(u^{\epsilon}(t,\cdot))$ on time intervals of order ϵ^{-1}, as $\epsilon \to 0$, can be uniquely characterized by solutions of a certain well-posed effective equation.

Keywords

KdV, NLS, Weakly nonlinear PDEs, Averaged equation, Effective equation.

Table des matières

Introduction

In mathematics and physics, the linear or integrable partial differential equations are usually deduced from idealization of certain physical processes. Normally, their solutions possess very clear and nice structures. However it is a fact of life that most of the processes that we encounter in physical realities are neither linear nor integrable but nonlinear and non-integrable. Fortunately, many important processes are described by suitable (nonlinear) perturbations of an integrable or linear equation. To study up to what extends the nice structures of the unperturbed equation can help us understand the behaviours of the perturbed systems has been an important and popular topic in the mathematics community. In this thesis, I will present several averaging-type theorems that describe the long-time behaviours of solutions for some perturbed Korteweg-de Vries (KdV) equations (which are perturbations of an integrable PDE) and nonlinear Schrödinger equations (NLS) with small nonlinearities (which are perturbations of a linear system).

On perturbed KdV. Consider a perturbed KdV equation with zero mean-value periodic boundary condition :

$$u_t + u_{xxx} - 6uu_x = \epsilon f(u)(x),\ x \in \mathbb{T} = \mathbb{R}/\mathbb{Z},\ \int_{\mathbb{T}} u(x,t)dx = 0, \tag{0.0.1}$$

where ϵf is a nonlinear perturbation to be specified below. For any $p \geqslant 0$, denote by H^p the Sobolev space of real valued functions on $\mathbb{T}$ with zero mean-value and by $||\cdot||_p$ the Sobolev norm or some related norms. It is well known that KdV is integrable. It means that the space H^p admits analytic coordinates

$$v = (\mathbf{v}_1, \mathbf{v}_2, \dots) = \Psi_K(u(\cdot)),$$

where $\mathbf{v}_j = (v_j, v_{-j})^t \in \mathbb{R}^2$, such that the quantities $I_j = \frac{1}{2}|\mathbf{v}_j|^2$ and $\varphi_j = \mathrm{Arg}\ \mathbf{v}_j$, $j \geqslant 1$, are action-angle variables for KdV. In the (I,φ)-varibles, KdV takes the integrable form

$$\dot{I} = 0, \quad \dot{\varphi} = W(I),$$

where $W(I) \in \mathbb{R}^\infty$ is the KdV frequency vector (see [38]).

One of the fundamental problems related to the solutions $u(t)$ of the perturbed equation (0.0.1) is the behaviours of the action variables $I(u(t))$ for $t \gg 1$. The KAM theory for PDEs (see [47, 38]) affirms that if (0.0.1) is a Hamiltonian system, then for typical finite dimensional initial data u_0 such that $\#\{j : I_j(u_0) \neq 0\} < +\infty$, we have $\sup_{t\in\mathbb{R}} |I(u(t)) - I(u_0)| \leqslant \epsilon^\sigma$ for some $\sigma \in (0,1)$. However these initial data form a null-set with respect to any reasonable measure in the Sobolev space H^p.

What happens if the initial datum is outside this null-set or if the perturbation is not hamiltonian ? The work here mainly concerns the dynamics of $I(u(t))$ in the time interval of order ϵ^{-1} for general initial data and general perturbations. Let us fix some $\zeta_0 \geqslant 0$, $p \geqslant 3$, $T > 0$, and assume :

Assumption A : *(i) For any $u_0 \in H^p$ and the equation (0.0.1), there exists a unique solution $u(\cdot) \in C([0, \epsilon^{-1}T], H^p)$ with $u(0) = u_0$. It satisfies*

$$||u||_p \leqslant C(T, p, ||u_0||_p), \quad 0 \leqslant t \leqslant T\epsilon^{-1}.$$

(ii) There exists a $p' = p'(p) < p$ such that for $q \in [p', p]$, the perturbation term defines an analytic mapping

$$H^q \to H^{q+\zeta_0}, \quad u(\cdot) \mapsto f(u)(\cdot).$$

Passing to slow time $\tau = \epsilon t$, write the equation in action-angle variables (I, φ),

$$\frac{dI}{d\tau} = F(I, \varphi), \quad \frac{d\varphi}{d\tau} = \epsilon^{-1}W(I) + G(I, \varphi). \tag{0.0.2}$$

Here $I \in \mathbb{R}^\infty$ and $\varphi \in \mathbb{T}^\infty$, where $\mathbb{T}^\infty := \{\theta = (\theta_j)_{j\geqslant 1}, \theta_j \in \mathbb{T}\}$ is the infinite-dimensional torus, endowed with the Tikhonov topology. The two functions $F(I, \varphi)$ and $G(I, \varphi)$ represent the perturbation term f, written in the action-angle variables. Inspired by finite dimensional averaging theory, we consider an averaged equation for the actions $I(\cdot)$:

$$\frac{dJ}{d\tau} = \langle F\rangle(J), \quad \langle F\rangle(J) = \int_{\mathbb{T}^\infty} F(J, \varphi)d\varphi, \tag{0.0.3}$$

where $d\varphi$ is the Haar measure on $\mathbb{T}^\infty$. It turns out that under the Assumption A, the equation (0.0.3) is well-posed, at least locally. The main task is to study the relation between the actions $I(\tau)$ of the solutions for equation (0.0.2) and solutions $J(\tau)$ of equation (0.0.3), for $\tau \in [0, T]$. One of the main obstacles here is that the frequency vector $W(I)$ of KdV is resonant in a dense subset of the space H^p. How to insure that the solutions of the perturbed KdV do not stay "too long" in this dense subset ? Our strategy to handle this is introducing a "new" tool : the *ϵ-quasi-invariant measures*[1] (see Definition 2.1.1).

The following theorem is proved in Chapter 2. Let $u^\epsilon(t)$ stand for solutions of equation (0.0.1) and $v^\epsilon(\tau) = \Psi_K(u^\epsilon(\epsilon^{-1}t))$. By Assumption A, for $\tau \in [0, T]$ we have $||I(v^\epsilon(\tau))||_p \leqslant C_1\Big(T, ||I(v^\epsilon(0))||_p\Big)$. Denote

$$\tilde{T}(I_0) := \min\{\tau \in \mathbb{R}_+ : J(0) = I_0 \;\&\; ||J(\tau)||_p \geqslant C_1(T, ||I_0||_p) + 1\}.$$

Theorem 0.0.1. *Let μ_p be an ϵ-quasi-invariant measure for equation (0.0.1) on H^p. Suppose that Assumption A holds and $\mathcal{B} \subset H^p$ is a bounded Borel set. Then*

1. Let S^t_ϵ be the flow map of the perturbed KdV (0.0.1) on H^p. Then the ϵ-quasi-invariant μ for equation (0.0.1) on H^p satisfies $e^{-\epsilon t C}\mu(A) \leqslant \mu(S^t_\epsilon A) \leqslant e^{\epsilon t C}\mu(A)$, for every bounded Borel subset $A \subset H^p$ and $t \in [0, \epsilon^{-1}T]$, where the constant C depends only on the bound of A and T.

(i) For any $\rho > 0$ and $q < p + \min\{1, \zeta_0/2\}$, there exists $\epsilon_{\rho,q} > 0$ and a Borel subset $\Gamma^{\epsilon}_{\rho,q} \subset \mathcal{B}$ such that

$$\lim_{\epsilon \to 0} \mu_p(\mathcal{B} \setminus \Gamma^{\epsilon}_{\rho,q}) = 0,$$

and for $\epsilon \leqslant \epsilon_\rho$, we have that if $u^\epsilon(0) \in \Gamma^{\epsilon}_{\rho,q}$, then

$$||I(v^\epsilon(\tau)) - J(\tau)||_q \leqslant \rho, \quad \text{for} \quad 0 \leqslant \tau \leqslant \min\{T, \tilde{T}(I_0^\epsilon)\}.$$

Here $I_0^\epsilon = I(v^\epsilon(0))$ and $J(\cdot)$ is a unique solution of the averaged equation (0.0.3) with initial data I_0^ϵ.

(ii) (ii) Let $\lambda_\epsilon^{v_0}$ be the probability measure on $\mathbb{T}^\infty$ defined by the relation

$$\int_{\mathbb{T}^\infty} f(\varphi)\, d\lambda_\epsilon^{v_0}(d\varphi) = \frac{1}{T}\int_0^T f(\varphi(v^\epsilon(\tau))d\tau, \quad \forall f \in C(\mathbb{T}^\infty),$$

where $v_0 = v_0(u_0) := \Psi_K(u^\epsilon(0))$. Then the averaged measure

$$\lambda_\epsilon := \frac{1}{\mu(\mathcal{B})}\int_{\mathcal{B}} \lambda_\epsilon^{v_0(u_0)} d\mu(u_0)$$

converges weakly, as $\epsilon \to 0$, to the Haar measure $d\varphi$ on $\mathbb{T}^\infty$

On the existence of the ϵ-quasi-invariant measures. We will provide two sufficient (not necessary) conditions for the existence of the ϵ-quasi-invariant measures for the perturbed KdV (0.0.1) in Section 2.5 of Chapter 2. Let

$$\mathcal{P}_K(v) = d\Psi_K(u)\big(f(u)\big), \quad v = \Psi_K(u)$$

Theorem 0.0.2. *If Assumption A holds and the map $v \mapsto \mathcal{P}_K(v)$ is ζ_0'-smoothing with $\zeta_0' > 1$, then there exist ϵ-quasi-invariant measures for the perturbed KdV (0.0.1).*

However, due to the complexity of the nonlinear Fourier transform Ψ_K, the additional smoothing condition in this theorem is not easy to verify. So it would be convient to have sufficient conditions directly on the map $f(u)$ in the Sobolev space H^p.

As is known, for solutions of KdV, there are countably many conservation laws $\mathcal{J}_n(u)$, $n \geqslant 0$, where $\mathcal{J}_0(u) = \frac{1}{2}||u||_0^2$ and

$$\mathcal{J}_n(u) = \int_{\mathbb{T}} \Big\{\frac{1}{2}(\partial_x^n u)^2 + c_n u(\partial_x^{n-1}u)^2 + \mathcal{Q}_n(u, \dots, \partial_x^{n-2}u)\Big\}dx, \tag{0.0.4}$$

for $n \geqslant 1$, where c_n are real constants, and $\mathcal{Q}_n$ are polynomial in their arguments (see e.g. [38]). Let μ_n be the Gibbs measures on the space H^n, generated by the conservation law $\mathcal{J}_n(u)$, which formally may be written as

$$d\mu_n = e^{-\mathcal{J}_{n+1} + \frac{1}{2}||u||_{p+1}^2} d\eta_n,$$

where η_n is the Gaussian measure on the space H^n with correlation operator ∂_x^{-2}. They are invariant for KdV ([75]). We have the following :

Theorem 0.0.3. *Let $p \in \mathbf{N}$. Then if Assumption A holds with $\zeta_0 \geqslant 2$, then the Gibbs measure μ_p is ϵ-quasi-invariant for the perturbed KdV (0.0.1).*

Particularly, this theorem and Theorem 0.0.1 apply to the equation :

$$u_t + u_{xxx} - 6uu_x = \epsilon f(x),$$

where $f(x)$ is a smooth function on the circle with zero mean value.

On weakly nonlinear equations. Consider a weakly nonlinear equation

$$\frac{d}{dt}u + i(-\triangle u + V(x)u) = \epsilon \mathcal{P}(\triangle u, \nabla u, u, x), \quad x \in \mathbb{T}^d, \tag{0.0.5}$$

where $\mathcal{P} : \mathbb{C}^{d+2} \times \mathbb{T}^d \to \mathbb{C}$ is a smooth function, $1 \leqslant V(x) \in C^n(\mathbb{T}^d)$ is a potential (we will assume that n is sufficiently large). *We fix some $p \geqslant d/2 + 4$ and suppose that the item (i) of the Assumption A holds for equation (0.0.5).*

Denote by A_V the Schrödinger operator

$$A_V u := -\triangle u + V(x)u.$$

Let $\{\zeta_k\}_{k\geqslant 1}$ and $\{\lambda_k\}_{k\geqslant 1}$ be its real eigenfunctions and eigenvalues, ordered in such a way that

$$1 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \cdots .$$

The potential $V(x)$ is called *non-resonant* if $\sum_{k=1}^{\infty} \lambda_k s_k \neq 0$, for every finite non-zero integer vector $(s_1, s_2, \cdots)$. For any complex-valued function $u(x) \in H^p$, we denote by

$$\Psi_S(u) := v = (v_1, v_2, \cdots), \quad v_j \in \mathbb{C},$$

the vector of its Fourier coefficients with respect to the basis $\{\zeta_k\}$, i.e. $u(x) = \sum_{k=1}^{\infty} v_k \zeta_k$. Denote $I_k = \frac{1}{2}|v_k|^2$, $k \geqslant 1$. We are mainly concerned of the behaviour of the quantity $I(t) = (I_1(t), \dots)$ in the time interval of order ϵ^{-1}. Using the mapping Ψ_S, we can rewrite equation (0.0.5) in the v-variables and in slow time $\tau = \epsilon t$ as,

$$\frac{dv}{d\tau} = \epsilon^{-1} d\Psi_S(u)(-iA_V(u)) + P(v). \tag{0.0.6}$$

Here $P(v)$ is the perturbation term $\mathcal{P}$, written in v-variables. This equation is singular when $\epsilon \to 0$. Following the work of S. Kuksin in [49], we introduce the *effective equation* for (0.0.6) as

$$\frac{dv}{d\tau} = \int_{\mathbb{T}^\infty} \Phi_{-\theta} P(\Phi_\theta v) d\theta, \tag{0.0.7}$$

where Φ_θ is the linear operator in the space of complex sequences $(v_1, v_2, \cdots)$, which multiplies each component v_j with $e^{i\theta_j}$. *We assume that the effective equation (0.0.7) is locally well posed in the space H^p.*

The following result is presented in Chapter 3. Let $v^\epsilon(\tau)$ be the Fourier transform of a solution $u^\epsilon(t, x)$ for the problem (0.0.5) with initial data in H^p, written in the slow time $\tau = \epsilon t$:

$$v^\epsilon(\tau) = \Psi_S\Big(u^\epsilon(\epsilon^{-1}\tau)\Big), \quad \tau \in [0, T].$$

Assume also that the potential $V(x)$ is non-resonant.

Theorem 0.0.4. *The curves $I(v^\epsilon(\tau))$, $\tau \in [0,T]$, converge to a curve $I^0(\tau)$, $\tau \in [0,T]$, as $\epsilon \to 0$, uniformly in $\tau \in [0,T]$. Moreover $I^0(\tau) = I(v(\tau))$, where $v(\cdot)$ is the unique solution of the effective equation (0.0.7), equal to $\Psi_S(u_0)$ at $\tau = 0$.*

Particularly, the theorem applies to a complex Ginzburg-Landau equation :

$$\frac{du}{d\tau} + \epsilon^{-1} i(-\triangle u + V(x)u) = \triangle u - \gamma_R f_p(|u|^2)u - i\gamma_I f_q(|u|^2)u, \quad x \in \mathbb{T}^d,$$

where the constants γ_R, $\gamma_I > 0$, the functions $f_p(r)$ and $f_q(r)$ are the monomials $|r|^p$ and $|r|^q$, smoothed out near zero, and

$$0 \leqslant p, q < \infty \quad \text{if} \quad d = 1, 2 \quad \text{and} \quad 0 \leqslant p, q < \min\{\frac{d}{2}, \frac{2}{d-2}\} \quad \text{if} \quad d \geqslant 3.$$

In the completely resonant case where in the equation (0.0.5) the potential $V(x) \equiv 0$, the assertion of Theorem 0.0.4 also holds true if the nonlinearity $\mathcal{P}$ is a polynomial of the unknown functions u and $\bar{u}$. In this situation, the corresponding effective equation is constructed through a certain *resonant averaging* process. See Chapter 4.

Organization of the thesis. In Chapter 1, we would cover some background knowledge on the finite dimensional integrable systems, classic averaging principle in finite dimensional space, Gaussian measures in Hilbert space and the integrability of the KdV equation. The averaging theory for perturbed KdV equations (Theorems 0.0.1-0.0.3) would be showed in Chapter 2. The Chapters 3 and 4 would discuss averaging theorems for weakly nonlinear equations. Except Chapter 1, each chapter here is self contained. Every chapter can be read independently.

Chapitre 1

Background

This chapter contains some background knowledge on finite dimensional integrable systems, classic averaging theory in finite dimensional space, Gaussian measure in Hilbert space and the integrality of the KdV equation. The Sections 1.1 and 1.4 are directly taken from the review [33].

1.1 Finite dimensional integrable systems

Classically, integrable systems are particular hamiltonian systems that can be integrated in quadratures. It was observed by Liouville that for a hamiltonian system with n degrees of freedom to be integrable, it has to possess n independent integrals in involution. This assertion can be understood globally (in the vicinity of an invariant torus or an invariant cylinder) and locally (in the vicinity of an equilibrium). Now we recall corresponding finite-dimensional definitions and results.

1.1.1 Liouville-integrable systems

Let $Q \subset \mathbb{R}^{2n}_{(p,q)}$ be a $2n$-dimensional domain. We provide it with the standard symplectic form $\omega_0 = dp \wedge dq$ and the corresponding Poisson bracket

$$\{f, g\} = \nabla_p f \cdot \nabla_q g - \nabla_q f \cdot \nabla_p g,$$

where $g, f \in C^1(Q)$ and “ $\cdot$ ” stands for the Euclidean scalar product in $\mathbb{R}^n$ (see [1]). If $\{f, g\} = 0$, the functions f and g are called *commuting*, or *in involution*. If $h(p, q)$ is a C^1-function on Q, then the hamiltonian system with the Hamiltonian h is

$$\dot{p} = -\nabla_q h, \quad \dot{q} = \nabla_p h. \tag{1.1.1}$$

Definition 1.1.1. *(Liouville-integrability). The hamiltonian system (1.1.1) is called integrable in the sense of Liouville if its Hamiltonian h admits n independent integrals in involution $h_1, \ldots, h_n$. That is, $\{h, h_i\} = 0$ for $1 \leqslant i \leqslant n$; $\{h_i, h_j\} = 0$ for $1 \leqslant i,\ j \leqslant n$, and $dh_1 \wedge \cdots \wedge dh_n \neq 0$.*

A nice structure of an Liouville-integrable system is assured by the celebrated Liouville-Arnold-Jost theorem (see [1, 63]) It claims that if an integrable systems is such that the level sets $T_c = \{(p,q) \in Q : h_1(p,q) = c_1, \dots, h_n(p,q) = c_n\}$, $c = (c_1, \dots, c_n) \in \mathbb{R}^n$ are compact, then each non-empty set T_c is an embedded n-dimensional torus. Moreover for a suitable neighborhood O_{T_c} of T_c in Q there exists a symplectomorphism

$$\Theta : O_{T_c} \to O \times \mathbb{T}^n = \{(I, \varphi)\}, \quad O \subset \mathbb{R}^n,$$

where the symplectic structure in $O \times \mathbb{T}^n$ is given by the 2-form $dI \wedge d\varphi$. Finally, there exists a function $\bar{h}(I)$ such that $h(p,q) = \bar{h}(\Theta(p,q))$. This result is true both in the smooth and analytic categories.

The coordinates (I, φ) are called the *action-angle variables* for (1.1.1). Using them the hamiltonian system may be written as

$$\dot{I} = 0, \quad \dot{\varphi} = \nabla_I \bar{h}(I). \tag{1.1.2}$$

Accordingly, in the original coordinates (p,q) solutions of the system are

$$(p,q)(t) = \Theta^{-1}(I_0, \varphi_0 + \nabla_I \bar{h}(I_0)t).$$

On $O \times \mathbb{T}^n$, consider the 1-form $Id\varphi = \sum_{j=1}^n I_j d\varphi_j$, then $d(Id\varphi) = dI \wedge d\varphi$. For any vector $I \in O$, and for $j = 1, \dots, n$, denote by $C_j(I)$ the cycle

$$\{(I, \varphi) \in O \times \mathbb{T}^n : \varphi_j \in [0, 2\pi] \text{ and } \varphi_i = const, \text{ if } i \neq j\}.$$

Then

$$\frac{1}{2\pi} \int_{C_j} I d\varphi = \frac{1}{2\pi} \int_{C_j} I_j d\varphi_j = I_j.$$

Consider a disc $D_j \subset O \times \mathbb{T}^n$ such that $\partial D_j = C_j$. For any 1-form ω_1, satisfying $d\omega_1 = dI \wedge d\varphi$, we have

$$\frac{1}{2\pi} \int_{C_j} (Id\varphi - \omega_1) = \frac{1}{2\pi} \int_{D_j} d(Id\varphi - \omega_1) = 0.$$

So

$$I_j = \frac{1}{2\pi} \int_{C_j(I)} \omega_1, \quad \text{if} \quad d\omega_1 = dI \wedge d\varphi. \tag{1.1.3}$$

This is the *Arnold formula for actions.*

1.1.2 Birkhoff Integrable systems

We denote by $\mathbb{J}$ the standard symplectic matrix $\mathbb{J} = \operatorname{diag}\left\{ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\}$, operating in any $\mathbb{R}^{2n}$ (e.g. in $\mathbb{R}^2$). Assume that the origin is an elliptic critical point of a smooth Hamiltonian h, i.e. $\nabla h(0) = 0$ and that the matrix $\mathbb{J}\nabla^2 h(0)$ has only pure imaginary eigenvalues. Then there exists a linear symplectic change of coordinates which puts h to the form

$$h = \sum_{i=1}^n \lambda_i (p_i^2 + q_i^2) + h.o.t, \qquad \lambda_j \in \mathbb{R} \quad \forall j.$$

If the frequencies $(\lambda_1, \dots, \lambda_n)$ satisfy some non-resonance conditions, then this normalization process can be carried out to higher order terms. The result of this normalization is known as the *Birkhoff normal form for the Hamiltonian h.*

Definition 1.1.2. *The frequencies* $\lambda_1, \dots, \lambda_n$ *are non-resonant up to order* $m \geqslant 1$ *if* $\sum_{i=1}^n k_i\lambda_i \neq 0$ *for each* $k \in \mathbb{Z}^n$ *such that* $1 \leqslant \sum_{i=1}^n |k_i| \leqslant m$. *They are called non-resonant if* $k_1\lambda_1 + \dots + k_n\lambda_n = 0$ *with integers* $k_1, \dots, k_n$ *only when all* k_j*'s vanish.*

Theorem 1.1.3. *(Birkhoff normal form, see [62, 63]) Let* $H = N_2 + \cdots$ *be a real analytic Hamiltonian in the vicinity of the origin in* $(\mathbb{R}^{2n}_{(p,q)}, dp \wedge dq)$ *with the quadratic part* $N_2 = \sum_{i=1}^n \lambda_i(q_i^2 + p_i^2)$. *If the (real) frequencies* $\lambda_1, \dots, \lambda_n$ *are non-resonant up to order* $m \geqslant 3$, *then there exists a real analytic symplectic trasformation* $\Psi_m = Id + \cdots$, *such that*

$$H \circ \Psi_m = N_2 + N_4 + \cdots + N_m + h.o.t.$$

Here N_i *are homogeneous polynomials of order* i, *which are actually smooth functions of variables* $p_1^2 + q_1^2, \dots, p_n^2 + q_n^2$. *If the frequencies are non-resonant, then there exists a formal symplectic transformation* $\Psi = Id + \cdots$, *represented by a formal power series, such that* $H \circ \Psi = N_2 + N_4 + \cdots$ *(this equality holds in the sense of formal series).*

If the transformation, converting H to the Birkhoff normal form, was convergent, then the resulting Hamiltonian would be integrable in a neighborhood of the origin with the integrals $p_1^2 + q_1^2, \dots, p_n^2 + q_n^2$. These functions are not independent when $p_i = q_i = 0$ for some i. So the system is not integrable in the sense of Liouville. But it is integrable in a weaker sense :

Definition 1.1.4. *Functions* $f_1, \dots, f_k$ *are functionally independent if their differentials* $df_1, \dots, df_k$ *are linearly independent on a dense open set. A* $2n$*-dimensional Hamiltonian is called Birkhoff integrable near an equilibrium* $m \in \mathbb{R}^{2n}$, *if it admits* n *functionally independent integrals in involution in the vicinity of* m.

Birkhoff normal form provides a powerful tool to study the dynamics of hamiltonian PDEs, e.g. see [54, 8] and references in [8] .

1.1.3 Vey theorem

The results of this subsection hold both in the C^∞-smooth and analytic categories.

Definition 1.1.5. *Consider a Birkhoff integrable system, defined near an equilibrium* $m \in \mathbb{R}^{2n}$, *with independent commuting integrals* $F = (F_1, \dots, F_n)$. *Its Poisson algebra is the linear space* $\mathcal{A}(F) = \Big\{G : \{G, F_i\} = 0,\ i = 1, \dots, n\Big\}$.

Note that although the integrals of an integrable system are not defined in a unique way, the corresponding algebra $\mathcal{A}(F)$ is.

Definition 1.1.6. *A Poisson algebra* $\mathcal{A}(F)$ *is said to be non-resonant at a point* $m \in \mathbb{R}^{2n}$, *if it contains a Hamiltonian with a non-resonant elliptic critical point at* m *(i.e., around* m *one can introduce symplectic coordinates* (p, q) *such that the quadratic part of that Hamiltonian at* m *is* $\sum \lambda_j(p_j^2 + q_j^2)$, *where the real numbers* λ_j *are non-resonant).*

It is easy to verify that if some $F_1 \in \mathcal{A}(\mathcal{F})$ is elliptic and non-resonant at the equilibrium m, then all other functions in $\mathcal{A}(\mathcal{F})$ are elliptic at m as well.

Theorem 1.1.7. *(Vey's theorem). Let $F = (F_1, \dots, F_n)$ be n functionally independent functions in involution in a neighbourhood of a point $m \in \mathbb{R}^{2n}$. If the Poisson algebra $\mathcal{A}(F)$ is non-resonant at m, then one can introduce around m symplectic coordinates (p, q) so that $\mathcal{A}(F)$ consists of all functions, which are actually functions of $p_1^2 + q_1^2, \dots, p_n^2 + q_n^2$.*

Example. Let $F = (f_1, \dots, f_n)$ be a system of smooth commuting Hamiltonians, defined in the vicinity of their joint equilibrium $m \in \mathbb{R}^{2n}$, such that the hessians $\nabla^2 f_i(m)$, $1 \leqslant i \leqslant n$, are linear independent. Then the theorem above applies to the Poisson algebra $\mathcal{A}(F)$.

In [72] Vey proved the theorem in the analytic case with an additional non-degeneracy condition, which was later removed by Ito in [34]. The results in [72, 34] also apply to non-elliptic cases. The smooth version of Theorem 1.1.7 is due to Eliasson [23]. There exists an infinite dimensional extension of the theorem, see [52].

1.2 The averaging principle

If a small perturbation is imposed upon an integrable conservative system, then the quantities that were integrals in the unperturbed system begins to slowly evolve. We assume the perturbed system can be written as

$$\dot{I} = \epsilon f(I, \varphi, \epsilon), \quad \dot{\varphi} = W(I) + \epsilon G(I, \varphi, \epsilon), \tag{1.2.1}$$

where $I \in D \subset \mathbb{R}^n$, $\varphi \in \mathbb{T}^m$, $f(I, \varphi, \epsilon)$, $W(I)$, $G(I, \varphi, \epsilon)$ are smooth functions of their parameters and ϵ is small. In the system (1.2.1) the variables I are called the slow variables and the phase φ are called the fast variables. Over times of order 1, the slow variables change only a bit, but over times of order ϵ^{-1}, their evolution may be considerable (of order 1). In many applications one is usually mainly interested in the behaviours of the slow variables. The *averaging principle* consists in using the *averaged system* :

$$\dot{J} = \epsilon \langle f \rangle (J), \quad \langle f \rangle (J) = \int_{\mathbb{T}^m} f(J, \varphi, 0) d\varphi, \tag{1.2.2}$$

for the approximate description of the evolution of the slow variables on the time interval of order ϵ^{-1}. This method has a very long history which dates back to the epoch of Lagrange and Laplace, who applied it to the problems of celestial mechanics, without proper justifications. Only in the last fifty years rigorous mathematical justification of the principle has been obtained, see in [66, 2, 57].

Let μ be the Lebesgue measure on $D \times \mathbb{T}^m \subset \mathbb{R}^{n+m}$. Assume that $I(t) \in D$ for $t \in [0, \epsilon^{-1}]$. Suppose the function $W(I)$ satisfies some non-degenerate conditions, mainly, the vector $W(I)$ is non-resonant for a.a $I \in D$. Then the following averaging principle is well established.

Averaging principle : *Let $(I(t), \varphi(t))$ and $J(t)$ satisfy equations (1.2.1) and (1.2.2) respectively. If $I(0) = J(0) = I_0$, then for each $\rho > 0$, there exists $\epsilon_\rho > 0$ and a subset $D_\epsilon \subset D \times \mathbb{T}^m$ such that*

1) $\lim_{\epsilon\to 0}\mu(D\times\mathbb{T}^m\setminus D_\epsilon)=0$.
2) For $\epsilon\leqslant\epsilon_\rho$, $(I_0,\varphi_0)\in D_\epsilon$, *we have* $\sup_{t\in[0,\epsilon^{-1}]}|I(t)-J(t)|\leqslant\rho$.

1.3 The Gaussian measure on Hilbert space

In this subsection we will recall the definition and some basis properties of the Gaussian measure on Hilbert space. Proofs and details may be found in [11].

Let H be a separable real Hilbert space with an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$. We denote the inner product on H by $(\cdot,\cdot)_H$ and the corresponding norm by $|\cdot|_H$. Let Y be a linear bounded self-adjoint operator acting on H and its eigen elements coincide with the basis $\{e_n\}_{n\in\mathbb{N}}$. Assume

$$\mathrm{Y}e_n=\lambda_n e_n,\quad n=1,2,\dots,$$

where $\lambda_n>0$ for all n. We recall that Y is an operator *of trace class* if $\sum_{n=1}^\infty\lambda_n<\infty$. We also introduce the operator $\mathrm{Y}^{1/2}$ defined by

$$\mathrm{Y}^{1/2}e_n=\lambda_n^{1/2}e_n,\quad n=1,2,\dots.$$

Definition 1.3.1. *We call a set* $\mathcal{M}\subset H$ *cylindrical if there exists an integer* $n>0$ *and a Borel set* $\mathcal{F}\subset\mathbb{R}^n$ *such that*

$$\mathcal{M}=\{x\in H:\ [(x,e_1)_H,\dots,(x,e_n)_H]\in\mathcal{F}\}.\tag{1.3.1}$$

We denote by $\mathcal{A}$ the collection of all cylindrical subsets of H. Clearly, it is an algebra.

Proposition 1.3.2. *The minimal* σ*-algebra containing the algebra* $\mathcal{A}$ *is the Borel* σ*-algebra of the Hilbert space* H*.*

Definition 1.3.3. *We call the additive (may not countably additive) measure* μ *defined on the algebra* $\mathcal{A}$ *by the rule : for* $\mathcal{M}\in\mathcal{A}$ *be as in (1.3.1),*

$$\mu(\mathcal{M})=(2\pi)^{-n/2}\prod_{j=1}^n\lambda_j^{-1/2}\int_{\mathcal{F}}e^{-\frac12\sum_{j=1}^n\lambda_j^{-1}x_j^2}dx_1\dots dx_n.$$

the (centered) Gaussian measure in H *with correlation operator Y.*

Proposition 1.3.4. *The Gaussian measure defined in Definition 1.3.3 is countably additive if and only if the correlation operator Y is of trace class.*

Therefore a Gaussian measure in the Hilbert space is a well defined probability measure if and only if its correlation operator is of trace class.

Theorem 1.3.5. *Let* μ *be a well defined Gaussian measure in the Hilbert space* H*, then for any* $x\in H$ *and* $r>0$*, we have* $\mu(\{x_1\in H:|x_1-x|_H\leqslant r)\})>0$.

For any $x_0 \in H$, consider the map $T : H \mapsto H$, $Tx = x + x_0$. Let μ_T be the push forward of the well defined Gaussian measure μ with correlation operator Y :

$$\mu_T(\mathcal{M}) = \mu(T^{-1}(\mathcal{M})), \quad \text{for every Borel set} \quad \mathcal{M} \subset H.$$

We have the following famous result describing the relation between the measures μ and μ_T :

Theorem 1.3.6. *(Cameron-Martin formula) 1) The Gaussian measures μ and μ_T are equivalent if and only if $x_0 \in Y^{1/2}(H)$ and the Radon-Nikodym derivative is given by*

$$\frac{d\mu_T}{d\mu}(x) = \exp[-|\,Y^{-1/2}x_0|_H^2 + (\,Y^{-1/2}x_0,\, Y^{-1/2}x)_H].$$

2) The measure μ and μ_T are singular if and only if $x_0 \notin Y^{1/2}(H)$.

1.4 Preliminary of KdV

The famous Korteweg-de Vries (KdV) equation

$$u_t = -u_{xxx} + 6uu_x, \quad x \in \mathbb{R},$$

was first proposed by Joseph Boussinesq [16] as a model for shallow water wave propagation. It became famous later when two Dutch mathematicians, Diederik Korteweg and Gustav De Vries [43], used it to explain the existence of a soliton water wave, previously observed by John Russel in physical experiments. Their work was so successful that this equation is now named after them. Since the mid-sixties of 20th century the KdV equation received a lot of attention from mathematical and physical communities after the numerical results of Kruskal and Zabusky [45] led to the discovery that its solitary wave solutions interact in an integrable way. It turns out that in some suitable setting, the KdV equation can be viewed as an integrable infinite dimensional hamiltonian system.

1.4.1 KdV under periodic boundary conditions as a hamiltonian system

Consider the KdV equation under zero mean value periodic boundary condition :

$$u_t + u_{xxx} - 6uu_x = 0, \quad x \in \mathbb{T} = \mathbb{R}/\mathbb{Z}, \quad \int_{\mathbb{T}} udx = 0. \tag{1.4.1}$$

(Note that the mean-value $\int_{\mathbb{T}} udx$ of a space-periodic solution u is a time-independent quantity, to simplify presentation we choose it to be zero.) To fix the setup, for any integer $p \geqslant 0$, we introduce the Sobolev space of real valued functions on $\mathbb{T}$ with zero mean-value :

$$H^p = \Big\{u \in L^2(\mathbb{T}, \mathbb{R}) : \ ||u||_p < +\infty, \ \int_{\mathbb{T}} u = 0\Big\}, \quad ||u||_p^2 = \sum_{k\in\mathbb{N}} |2\pi k|^{2p}(|\hat{u}_k|^2 + |\hat{u}_{-k}|^2).$$

Here $\hat{u}_k$, $\hat{u}_{-k}$, $k \in \mathbb{N}$, are the Fourier coefficients of u with respect to the trigonometric base

$$e_k = \sqrt{2}\cos 2\pi kx, \quad k > 0 \quad \text{and} \quad e_k = \sqrt{2}\sin 2\pi kx, \quad k < 0, \tag{1.4.2}$$

i.e.

$$u = \sum_{k\in\mathbb{N}} \hat{u}_k e_k + \hat{u}_{-k} e_{-k}. \tag{1.4.3}$$

In particular, H^0 is the space of L^2-functions on $\mathbb{T}$ with zero mean-value. By $\langle\cdot,\cdot\rangle$ we denote the scalar product in H^0 (i.e. the L^2-scalar product).

For a C^1-smooth functional F on some space H^p, we denote by ∇F its gradient with respect to $\langle\cdot,\cdot\rangle$, i.e.

$$dF(u)(v) = \langle \nabla F(u), v\rangle,$$

if u and v are sufficiently smooth. So $\nabla F(u) = \frac{\delta F}{\delta u(x)} + const$, where $\frac{\delta F}{\delta u}$ is the variational derivative, and the constant is chosen in such a way that the mean-value of the r.h.s vanishes. See [47, 38] for details. The initial value problem for KdV on the circle $\mathbb{T}$ is well posed on every Sobolev space H^p with $p \geqslant 1$, see [70, 13]. The regularity of KdV in function spaces of lower smoothness was studied intensively, see [19, 41] and references in these works ; also see [19] for some qualitative results concerning the KdV flow in these spaces. We avoid this topic.

It was observed by Gardner [26] that if we introduce the Poisson bracket which assigns to any two functionals $F(u)$ and $G(u)$ the new functional $\{F,G\}$,

$$\Big\{F,G\Big\}(u) = \int_{\mathbb{T}} \frac{d}{dx}\nabla F(u(x))\nabla G(u(x))dx \tag{1.4.4}$$

(we assume that the r.h.s is well defined, see [47, 48, 38] for details), then KdV becomes a hamiltonian PDE. Indeed, this bracket corresponds to a differentiable hamiltonian function F a vector filed $\mathcal{V}_F$, such that

$$\langle \mathcal{V}_F(u), \nabla G(u)\rangle = \{F,G\}(u)$$

for any differentiable functional G. From this relation we see that $\mathcal{V}_F(u) = \frac{\partial}{\partial x}\nabla F(u)$. So the KdV equation takes the hamiltonian form

$$u_t = \frac{\partial}{\partial x}\nabla\mathcal{H}(u), \tag{1.4.5}$$

with the KdV Hamiltonian

$$\mathcal{H}(u) = \int_{\mathbb{T}} (\frac{u_x^2}{2} + u^3)dx. \tag{1.4.6}$$

The Gardner bracket (1.4.4) corresponds to the symplectic structure, defined in H^0 (as well as in any space H^p, $p \geqslant 0$) by the 2-form

$$\omega_2^G(\xi,\eta) = \Big\langle (-\frac{\partial}{\partial x})^{-1}\xi, \eta\Big\rangle \quad \text{for} \quad \xi,\eta \in H^0. \tag{1.4.7}$$

Indeed, since $\omega_2^G(\mathcal{V}_F(u),\xi) \equiv -\langle \nabla F(u), \xi\rangle$, then the 2-form ω_2^G also assigns to a Hamiltonian F the vector field $\mathcal{V}_F$ (see [1, 38, 47, 48]).

We note that the bracket (1.4.4) is well defined on the whole Sobolev spaces $H^p(\mathbb{T}) = H^p \oplus \mathbb{R}$, while the symplectic form ω_2^G is not, and the affine subspaces $\{u \in H^p(\mathbb{T}) : \int_{\mathbb{T}} u dx = const\} \simeq H^p$ are symplectic leaves for this Poisson system. We study the equation only on the leaf $\int_{\mathbb{T}} u dx = 0$, but on other leaves it may be studied similarly.

Writing a function $u(x) \in H^0$ as in (1.4.3) we see that $\omega_2^G = \sum_{k=1}^{\infty} k^{-1} d\hat{u}_k \wedge d\hat{u}_{-k}$ and that $\mathcal{H}(u) = H(\hat{u}) := \Lambda(\hat{u}) + G(\hat{u})$ with

$$\Lambda(\hat{u}) = \sum_{k=1}^{+\infty} (2\pi k)^2 \Big(\frac{1}{2}\hat{u}_k^2 + \frac{1}{2}\hat{u}_{-k}^2\Big), \qquad G(\hat{u}) = \sum_{k,l,m\neq 0, k+l+m=0} \hat{u}_k \hat{u}_l \hat{u}_m.$$

Accordingly, the KdV equation may be written as the infinite chain of hamiltonian equations

$$\frac{d}{dt}\hat{u}_j = -2\pi j \frac{\partial H(\hat{u})}{\partial \hat{u}_{-j}}, \quad j = \pm 1, \pm 2, \ldots.$$

1.4.2 Lax pair

The KdV equation (1.4.1) admits infinitely many integrals in involution, and there are different ways to obtain them, see [26, 61, 64, 55, 74]. Below we present an elegant way to construct a set of Poisson commuting integrals by considering the spectrum of an associated Schrödinger operator, due to Peter Lax [55] (see [56] for a nice presentation of the theory).

Let $u(x)$ be a L^2-function on $\mathbb{T}$. Consider the differential operators L_u and B_u, acting on 2-periodic functions [1]

$$L_u = -\frac{d^2}{dx^2} + u(x), \quad B_u = -4\frac{d^3}{dx^3} + 3u(x)\frac{d}{dx} + 3\frac{d}{dx}u(x),$$

where we view $u(x)$ as a multiplication operator $f \mapsto u(x)f$. The operators B_u and L_u are called the *Lax pair* for KdV. Calculating the commutator $[B_u, L_u] = B_u L_u - L_u B_u$, we see that most of the terms cancel and the only term left is $-u_{xxx} + 6uu_x$. Therefore if $u(t,x)$ is a solution of (1.4.1), then the operators $L(t) = L_{u(t,\cdot)}$ and $B(t) = B_{u(t,\cdot)}$ satisfy the operator equation

$$\frac{d}{dt}L(t) = [B(t), L(t)]. \tag{1.4.8}$$

Note that the operator $B(t)$ are skew-symmetric, $B(t)^* = -B(t)$. Let $U(t)$ be the one-parameter family of unitary operators, defined by the differential equation

$$\frac{d}{dt}U = B(t)U, \quad U(0) = \mathrm{Id}.$$

Then $L(t) = U^{-1}(t)L(0)U(t)$. Therefore, the operator $L(t)$ is unitary conjugated to $L(0)$. Consequently, its spectrum is independent of t. That is, the spectral data of the operator L_u provide a set of conserved quantities for the KdV equation (1.4.1).

1. note the doubling of the period.

Since L_u is the strurm-Liouville operator with a potential $u(x)$, then in the context of this theory functions $u(x)$ are called *potentials*.

It is well known that for any L^2-potential u the spectrum of the Sturm-Liouville operator L_u, regarded as an unbounded operator in $L^2(\mathbb{R}/2\mathbb{Z})$, is a sequence of simple or double eigenvalues $\{\lambda_j : j \geqslant 0\}$, tending to infinity :

$$\text{spec}(u) = \{\lambda_0 < \lambda_1 \leqslant \lambda_2 < \cdots \nearrow \infty\}.$$

Equality or inequality may occur in every place with a "$\leqslant$" sign (see [58, 38]). The segment $[\lambda_{2j-1}, \lambda_{2j}]$ is called the *n-th spectral gap*. The asymptotic behaviour of the periodic eigenvalues is

$$\lambda_{2n-1}(u),\ \lambda_{2n}(u) = n^2\pi^2 + [u] + l^2(n),$$

where $[u]$ is the mean value of u, and $l^2(n)$ is the n-th number of an l^2 sequence. Let $g_n(u) = \lambda_{2n}(u) - \lambda_{2n-1}(u) \geqslant 0$, $n \geqslant 1$. These quantities are conserved under the flow of KdV. We call g_n the *n-th gap-length* of the spectrum. The n-th gap is called *open* if $g_n > 0$, otherwise it is *closed*. However, from the analytic point of view the periodic eigenvalues and the gap-lengths are not satisfactory integrals, since λ_n is not a smooth function of the potential u when $g_n = 0$. Fortunately, the squared gap lengths $g_n^2(u)$, $n \geqslant 1$, are real analytic functions on L^2, which Poisson commute with each other (see [59, 56, 38]). Moreover, together with the mean value, the gap lengths determine uniquely the periodic spectrum of a potential, and their asymptotic behavior characterizes the regularity of a potential in exactly the same way as its Fourier coefficients [58, 27].

This method applies to integrate other hamiltonian systems in finite or infinite dimension. It is remarkably general and is referred to as the *method of Lax pair.*

1.4.3 Action-angle coordinates

We denote by $\text{Iso}(u_0)$ the isospectral set of a potential $u_0 \in H^0$:

$$\text{Iso}(u_0) = \Big\{u \in H^0 : \quad \text{spec}(u) = \text{spec}(u_0)\Big\}.$$

It is invariant under the flow of KdV and may be characterized by the gap lengths

$$\text{Iso}(u_0) = \Big\{u \in H^0 : \quad g_n(u) = g_n(u_0),\ n \geqslant 1\Big\}.$$

Moreover, for any $n \geqslant 1$, $u_0 \in H^n$ if and only if $\text{Iso}(u_0) \subset H^n$.

In [59], McKean and Trobwitz showed that the $\text{Iso}(u_0)$ is homemorphic to a compact torus, whose dimension equals the number of open gaps. So the phase space H^0 is foliated by a collection of KdV-invariant tori of different dimensions, finite or infinite. A potential $u \in H^0$ is called *finite-gap* if only a finite number of its spectral gaps are open. The finite-dimensional KdV-invariant torus $\text{Iso}(u_0)$ is called a *finite-gap torus*. For any $n \in \mathbb{N}$ let us set

$$\mathcal{J}^n = \Big\{u \in H^0 : \ g_j(u) = 0 \ \text{ if } \ j > n\Big\}. \tag{1.4.9}$$

We call the sets $\mathcal{J}^n$, $n \in \mathbb{N}$, the finite-gap manifolds.

Theorem 1.4.1. *For any $n \in \mathbb{N}$, the finite gap manifold $(\mathcal{J}^n, \omega_2^G)$ is a smooth symplectic $2n$-manifold, invariant under the flow of KdV (1.4.1), and*

$$T_0\mathcal{J}^n = \Big\{ u \in H^0 : \ \hat{u}_k = 0 \ \text{ if } |k| \geqslant n+1 \Big\},$$

(see (1.4.3)). Moreover, the square gap lengths $g_k^2(u)$, $k = 1, \dots, n$, form n commuting analytic integrals of motions, non-degenerated everywhere on the dense domain $\mathcal{J}_0^n = \{u \in \mathcal{J}^n : g_1(u), \dots, g_n(u) > 0\}$.

Therefore, the Liouville-Arnold-Jost theorem applies everywhere on $\mathcal{J}_0^n$, $n \in \mathbb{N}$. Furthermore, the union of the finite gap manifolds $\cup_{n\in\mathbb{N}}\mathcal{J}^n$ is dense in each space H^s (see [58]). This hints that on the spaces H^s, $s \geqslant 0$, it may be possible to construct global action-angle coordinates for KdV. In [25], Flaschka and McLaughlin used the Arnold formula (1.1.3) to get an explicit formula for action variables of KdV in terms of the 2-period spectral data of L_u. To explain their construction, denote by $y_1(x, \lambda, u)$ and $y_2(x, \lambda, u)$ the standard fundamental solutions of the equation $-y'' + uy = \lambda y$, defined by the initial conditions

$$\begin{aligned} y_1(0, \lambda, u) &= 1, \quad & y_2(0, \lambda, u) &= 0, \\ y_1'(0, \lambda, u) &= 0, \quad & y_2'(0, \lambda, u) &= 1. \end{aligned}$$

The quantity $\triangle(\lambda, u) = y_1(1, \lambda, u) + y_2'(1, \lambda, u)$ is called the *discriminant*, associated with this pair of solutions. The periodic spectrum of u is precisely the zero set of the entire function $\triangle^2(\lambda, u) - 4$, for which we have the explicit representation (see e.g. [74, 59])

$$\triangle^2(\lambda, u) - 4 = 4(\lambda_0 - \lambda) \prod_{n\geqslant 1} \frac{(\lambda_{2n} - \lambda)(\lambda_{2n-1} - \lambda)}{n^4\pi^4}.$$

This function is a spectral invariant. We also need the spectrum of the differential operator $L_u = -\frac{d^2}{dx^2} + u$ under Dirichlet boundary conditions on the interval $[0, 1]$. It consists of an unbounded sequence of single Dirichlet eigenvalues

$$\mu_1(u) < \mu_2(u) < \cdots \nearrow \infty,$$

which satisfy $\lambda_{2n-1}(u) \leqslant \mu_n(u) \leqslant \lambda_{2n}(u)$, for all $n \in \mathbb{N}$. Thus, the n-th Dirichlet eigenvalue μ_n is always contained in the n-th spectral gap. The Dirichlet spectrum provides coordinates on the isospectral sets (see [59, 58, 38]). For any $z \in \mathbb{T}$, denote by $\{\mu_j(u, z), \ j \geqslant 1\}$ the spectrum of the operator L_u under the shifted Dirichlet boundary conditions $y(z) = y(z+1) = 0$ (so $\mu_j(u, 0) = \mu_j(u)$) ; still $\lambda_{2n-1} \leqslant \mu_n(u, z) \leqslant \lambda_{2n}(u)$. Jointly with the spectrum $\{\lambda_j\}$, it defines the potential $u(x)$ via the remarkable *trace formula* (see [74, 21, 38, 59]) :

$$u(z) = \lambda_0(u) + \sum_{j=1}^{\infty} (\lambda_{2j-1}(u) + \lambda_{2j}(u) - 2\mu_j(u, z)).$$

Define

$$f_n(u) = 2\log(-1)^n y_2'(1, \mu_n(u), u), \quad \forall n \in \mathbb{N}.$$

Flashka and McLaughlin [25] observed that the quantities $\{\mu_n, f_n\}_{n\in\mathbb{N}}$ form canonical coordinates of H^0, i.e.

$$\{\mu_n, \mu_m\} = \{f_n, f_m\} = 0, \quad \{\mu_n, f_m\} = \delta_{n,m}, \quad \forall n, m \in \mathbb{N}.$$

Accordingly, the symplectic form ω_2^G (see (1.4.7)) equals $d\omega_1$, where ω_1 is the 1-form $\sum_{n\in\mathbb{N}} f_n d\mu_n$. Now the KdV action variables are given by the Arnold formula (1.1.3), where C_n is a circle on the invariant torus $\operatorname{Iso}(u)$, corresponding to $\mu_n(u)$. It is shown in [25] that

$$I_n = \frac{2}{\pi}\int_{\lambda_{2n-1}}^{\lambda_{2n}} \lambda \frac{\dot{\triangle}(\lambda)}{\sqrt{\triangle^2(\lambda)-4}} d\lambda, \quad \forall n \in \mathbb{N}.$$

The analytic properties of the functions $u \mapsto I_n$ and of the mapping $u \mapsto I = (I_1, I_2, \dots)$ were studied later by Kappeler and Korotyaev (see references in [38, 42] and below). In particular, it was shown that $I_n(u)$, $n \in \mathbb{N}$, are real analytic functions on H^0 of the form $I_n = g_n^2 +$ $h.o.t$, and $I_n = 0$ if and only if $g_n = 0$, see in [38]. For any vector $I = (I_1, I_2, \dots)$ with non-negative components we will denote

$$T_I = \{u(x) \in H^0 : I_n(u) = I_n \quad \forall\, n\}. \tag{1.4.10}$$

The angle-variables φ^n on the finite-gap manifolds $\mathcal{J}^n$ were found in 1970's by Soviet mathematicians, who constructed them from the Dirichlet spectrum $\{\mu_j(u)\}$ by means of the Abel transform, associated with the Riemann surface of the function $\sqrt{\triangle^2 - 4}$, see [21, 58, 74], and see [35, 20, 44, 9] for the celebrated explicit formulas for angle-variables φ^n and for finite-gap solutions of KdV in terms of the theta-functions.

In [46] and [47], Section 7, the action-angle variables (I^n, φ^n) on a finite-gap manifold $\mathcal{J}^n$ and the explicit formulas for solutions of KdV on manifolds $\mathcal{J}^N$, $N \geq n$, from the works [20, 44, 9] were used to obtain an analytic symplectic coordinate system (I^n, φ^n, y) in the vicinity of $\mathcal{J}^n$ in H^p. The variable y belongs to a ball in a subspace $Y \subset H^p$ of co-dimention $2n$, and in the new coordinates the KdV Hamiltonian (1.4.6) reads

$$\mathcal{H} = \text{const} + h^n(I^n) + \langle A(I^n)y, y\rangle + O(y^3). \tag{1.4.11}$$

The selfadjoint operator $A(I^n)$ is diagonal in some fixed symplectic basis of Y. The nonlinearity $O(y^3)$ defines a hamiltonian operator of order one. That is, the KdV's linear operator, which is an operator of order three, mostly transforms to the linear part of the new hamiltonian operator and "does not spread much" to its nonlinear part. This is the crucial property of (1.4.11). The normal form (1.4.11) is instrumental for the purposes of the KAM-theory, see [47]

McKean and Trubowitz in [59, 60] extended the construction of angles on finite-gap manifolds to the set of all potentials, thus obtaining angle variables $\varphi = (\varphi_1, \varphi_2, \dots)$ on the whole space H^p, $p \geqslant 0$. The angles $(\varphi_k(u), k \geq 1)$ are well defined Gateaux-analytic functions of u outside the locus

$$\eth = \{u(x) : g_j(u) = 0 \text{ for some } j\}, \tag{1.4.12}$$

which is dense in each space H^p. The action-map $u \mapsto I$ was not considered in [59, 60], but it may be shown that outside $\eth$, in a certain weak sense, the variables (I, φ) are KdV's action-angles (see the next section for a stronger statement). This result is nice and elegant, but it is insufficient to study perturbations of KdV since the transformation to the variables (I, φ) is singular at the dense locus $\eth$.

1.4.4 Birkhoff coordinates and nonlinear Fourier transform

In a number of publications (see in [38]), Kappeler with collaborators proved that the Birkhoff coordinates $v = \{v_n,\ n = \pm 1, \pm 2, \dots\}$, associated with the action-angles variables (I, φ),

$$v_n = \sqrt{2I_n}\cos(\varphi_n), \quad v_{-n} = \sqrt{2I_n}\sin(\varphi_n), \quad \forall n \in \mathbb{N}, \tag{1.4.13}$$

are analytic on the whole of H^0 and define there a global coordinate system, in which the KdV Hamiltonian (1.4.6) is a function of the actions only. This remarkable result significantly specifies the normal form (1.4.11). To state it exactly, for any $p \in \mathbb{R}$, we introduce the Hilbert space h^p,

$$h^p := \Big\{ v = (\mathbf{v}_1, \mathbf{v}_2, \cdots) : |v|_p^2 = \sum_{j=1}^{+\infty} (2\pi j)^{2p+1} |\mathbf{v}_j|^2 < \infty,\ \mathbf{v}_j = (v_j, v_{-j})^t \in \mathbb{R}^2,\ j \in \mathbb{N} \Big\},$$

and the weighted l^1-space h_I^p,

$$h_I^p := \Big\{ I = (I_1, \dots) \in \mathbb{R}^\infty : |I|_p^\sim = 2\sum_{j=1}^{+\infty} (2\pi j)^{2p+1} |I_j| < +\infty \Big\}.$$

Define the mappings

$$\pi_I :\ h^p \to h_I^p, \quad v \mapsto I = (I_1, I_2, \dots), \ \text{ where }\ I_k = \frac{1}{2}|\mathbf{v}_k|^2 \quad \forall\, k,$$

$$\pi_\varphi :\ h^p \to \mathbb{T}^\infty, \quad v \mapsto \varphi = (\varphi_1, \varphi_2, \dots), \ \text{ where }\ \varphi_k = \arctan\big(\frac{v_{-k}}{v_k}\big)$$
$$\text{if }\ \mathbf{v}_k \neq 0, \ \text{ and }\ \varphi_k = 0\ \text{ if }\ \mathbf{v}_k = 0.$$

Since $|\pi_I(v)|_p^\sim = |v|_p^2$, then π_I is continuous. Its image $h_{I+}^p = \pi_I(h^p)$ is the positive octant in h_I^p. When there is no ambiguity, we write $I(v) = \pi_I(v)$.

Consider the mapping

$$\Psi : u(x) \mapsto v = (\mathbf{v}_1, \mathbf{v}_2, \dots), \quad \mathbf{v}_n = (v_n, v_{-n})^t \in \mathbb{R}^2,$$

where $v_{\pm n}$ are defined by (1.4.13) and $\{I_n(u)\}$, $\{\varphi_n(u)\}$ are the actions and angles as in Subsection 1.3.2. Clearly $\pi_I \circ \Psi(u) = I(u)$ and $\pi_\varphi \circ \Psi(u) = \varphi(u)$. Below we refer to Ψ as to the *nonlinear Fourier transform.*

Theorem 1.4.2. *(see [38, 37]) The mapping Ψ defines an analytical symplectomorphism $\Psi : (H^0, \omega_2^G) \to (h^0, \sum_{k=1}^\infty dv_k \wedge dv_{-k})$ with the following properties :*

(i) For any $p \in [-1, +\infty)$, it defines an analytic diffeomorphism $\Psi : H^p \mapsto h^p$.

(ii) (Percival's identity) If $v = \Psi(u)$, then $|v|_0 = ||u||_0$.

(iii) (Normalisation) The differential $d\Psi(0)$ is the operator $\sum u_s e_s \mapsto v$, where $v_s = |2\pi s|^{-1/2} u_s$ for each s.

(iv) The function $\hat{H}(v) = \mathcal{H}(\Psi^{-1}(v))$ has the form $\hat{H}(v) = H_K(I(v))$, where the function $H_K(I)$ is analytic in a suitable neighborhood of the octant h^1_{I+} in h^1_I, such that a curve $u \in C^1(0,T;H^0)$ is a solution of KdV if and only if $v(t) = \Psi(u(t))$ satisfies the equations

$$\dot{\mathbf{v}}_j = \mathbb{J}\frac{\partial H_K}{\partial I_j}(I)\mathbf{v}_j, \quad \mathbf{v}_j = (v_j, v_{-j})^t \in \mathbb{R}^2,\ j \in \mathbb{N}. \tag{1.4.14}$$

The assertion (iii) normalizes Ψ in the following sense. For any $\theta = (\theta_1, \theta_2, \dots) \in \mathbb{T}^\infty$ denote by Φ_θ the operator

$$\Phi_\theta v = v', \quad \mathbf{v}'_j = \bar{\Phi}_{\theta_j}\mathbf{v}_j, \quad \forall j \in \mathbb{N}, \tag{1.4.15}$$

where $\bar{\Phi}_\alpha$ is the rotation of the plane $\mathbb{R}^2$ by the angle α. Then $\Phi_\theta \circ \Psi$ satisfies all assertions of the theorem except (iii). But the properties (i)-(iv) jointly determine Ψ in a unique way.

The theorem above can be viewed as a global infinite dimensional version of the Vey Theorem 1.1.7 for KdV, and eq. (1.4.14) – as a global Birkhoff normal form for KdV. Note that in finite dimension a global Birkhoff normal form exists only for very exceptional integrable equations, which were found during the boom of activity in integrable systems, provoked by the discovery of the method of Lax pair.

Remark 1.4.3. *The map Ψ simultaneously transforms all Hamiltonians of the KdV hierarchy to the Birkhoff normal form. The KdV hierarchy is a collection of hamiltonian functions $\mathcal{J}_l$, $l \geqslant 0$, commuting with the KdV Hamiltonian, and having the form*

$$\mathcal{J}_l(u) = \int \Big(\frac{1}{2}(u^{(l)})^2 + J_{l-1}(u)\Big)dx.$$

Here $J_{-1} = 0$ and $J_{l-1}(u)$, $l \geq 1$, is a polynomial of $u, \dots, u^{(l-1)}$. The functions from the KdV hierarchy form another complete set of KdV integrals. E.g. see [21, 38, 56].

One of the important properties of the nonlinear Fourier transform Ψ that we will use in Chapter 2 is that It is quasi-linear. Precisely,

Theorem 1.4.4. *If $m \geqslant 0$, then the map $\Psi - d\Psi(0) : H^m \to h^{m+1}$ is analytic.*

That is, the non-linear part of Ψ is 1-smoother than its linearisation at the origin. See [52] for a local version of this theorem, applicable as well to other integrable infinite-dimensional systems, and see [39, 40] for the global result. The fact that the global transformation to the normal form (1.4.11) also is quasi-linear, is established in see [46, 47].

1.4.5 Properties of frequency map

Let us denote

$$W(I) = (W_1(I),\ W_2(I), \dots), \quad W_i(I) = \frac{\partial H_K}{\partial I_i}, \quad i \in \mathbb{N}. \tag{1.4.16}$$

This is the *frequency map for KdV*. By Theorem 1.4.2 each its component is an analytic function, defined in the vicinity of h^1_{I+} in h^1_I.

Lemma 1.4.5. *a) For $i, j \geq 1$ we have $\partial^2 W(0)/\partial I_i \partial I_j = -6\delta_{i,j}$.*
b) For any $n \in \mathbb{N}$, if $I_{n+1} = I_{n+2} = \cdots = 0$, then

$$\det\left(\left(\frac{\partial W_i}{\partial I_j}\right)_{1\leqslant i,j\leqslant n}\right) \neq 0.$$

For a) see [10, 38, 47]. For a proof of b) and references to the original works of Krichever and Bikbaev-Kuksin see Section 3.3 of [47].

Let l^∞_i, $i \in \mathbb{Z}$, be the Banach spaces of all real sequences $l = (l_1, l_2, \dots)$ with norms

$$|l|^\infty_i = \sup_{n\geqslant 1} n^i |l_n| < \infty.$$

Denote $\boldsymbol{\kappa} = (\kappa_n)_{n\in\mathbb{N}}$, where $\kappa_n = (2\pi n)^3$. For the following result see [38], Theorem 15.4.

Lemma 1.4.6. *The normalized frequency map $I \mapsto W(I) - \boldsymbol{\kappa}$ is real analytic as a mapping from h^1 to l^∞_{-1}.*

From these two lemmata we known that the Hamiltonian $H_K(I)$ of KdV is non-degenerated in the sense of Kolmogorov and its nonlinear part is more regular than its linear part. These properties are very important to study perturbations of KdV.

Chapitre 2

An averaging theorem for perturbed KdV equations

The results of this chapter is taken from my papers [31] and [32].

Abstract : Consider a perturbed KdV equation :

$$u_t + u_{xxx} - 6uu_x = \epsilon f(u)(x), \quad x \in \mathbb{T} = \mathbb{R}/\mathbb{Z}, \int_{\mathbb{T}} u(x,t)dx = 0, \tag{$*$}$$

where the nonlinear perturbation defines analytic operators $u(\cdot) \mapsto f(u(\cdot))$ in sufficiently smooth Sobolev spaces. For a periodic function $u(x)$, let $I(u) = (I_1(u), I_2(u), \cdots) \in \mathbb{R}_+^\infty$ be the vector, formed by the KdV integrals of motion, calculated for the potential $u(x)$. Assume that the equation $(*)$ has an ϵ-quasi-invariant measure μ and satisfies some additional mild assumptions. Let $u^\epsilon(t)$ be a solution. Then on time intervals of order ϵ^{-1}, as $\epsilon \to 0$, its actions $I(u^\epsilon(t, \cdot))$ can be approximated by solutions of a certain well-posed averaged equation, provided that the initial datum is μ-typical.

2.1 Introduction

We consider a perturbed Korteweg-de Vries (KdV) equation with zero mean-value periodic boundary condition :

$$\dot{u} + u_{xxx} - 6uu_x = \epsilon f(u)(x), \quad x \in \mathbb{T} = \mathbb{R}/\mathbb{Z}, \quad \int_{\mathbb{T}} u(x,t)dx = 0. \tag{2.1.1}$$

Here $\epsilon f(u(\cdot))$ is a nonlinear perturbation, specified below. For any $p \in \mathbb{R}$ we denote by H^p the Sobolev space of order p, formed by real-valued periodic functions with zero mean-value, provided with the homogeneous norm $||\cdot||_p$. Particularly, if $p \in \mathbb{N}$ we have

$$H^p = \Big\{u \in L^2(\mathbb{T}) : ||u||_p < \infty, \int_{\mathbb{T}} u dx = 0\Big\}, \qquad ||u||_p^2 = \int_{\mathbb{T}} \Big|\frac{\partial^p u}{\partial x^p}\Big|^2 dx.$$

For any p, the operator $\frac{\partial}{\partial x}$ defines a linear isomorphism : $\frac{\partial}{\partial x} : H^p \to H^{p-1}$. Denoting by $(\frac{\partial}{\partial x})^{-1}$ its inverse, we provide the spaces H^p, $p \geqslant 0$, with a symplectic structure by means of the 2-form Ω :

$$\Omega(u_1, u_2) = -\Big\langle (\frac{\partial}{\partial x})^{-1} u_1, u_2 \Big\rangle, \tag{2.1.2}$$

where $\langle\cdot,\cdot\rangle$ is the scalar product in $L^2(\mathbb{T})$. Then in any space H^p, $p \geqslant 1$, the KdV equation $(2.1.1)_{\epsilon=0}$ may be written as a Hamiltonian system with the Hamiltonian $\mathcal{H}$, given by $\mathcal{H}(u) = \int_{\mathbb{T}} \left(\frac{1}{2}u_x^2 + u^3\right)dx$. That is, KdV may be written as

$$\dot{u} = \frac{\partial}{\partial x}\nabla\mathcal{H}(u).$$

It is well-known that KdV is integrable. It means that the function space H^p admits analytic symplectic coordinates $v = (\mathbf{v}_1, \mathbf{v}_2, \cdots) = \Psi(u(\cdot))$, where $\mathbf{v}_j = (v_j, v_{-j}) \in \mathbb{R}^2$, such that the quantities $I_j = \frac{1}{2}|\mathbf{v}_j|^2$, $j \geqslant 1$, are actions (integrals of motion), while $\varphi_j = \text{Arg } \mathbf{v}_j$, $j \geqslant 1$, are angles. In the (I,φ)-variables, KdV takes the integrable form

$$\dot{I} = 0, \quad \dot{\varphi} = W(I), \tag{2.1.3}$$

where $W(I) \in \mathbb{R}^\infty$ is the frequency vector (see [38]). For any $p \geqslant 0$, the integrating transformation Ψ, called the nonlinear Fourier transform, defines an analytic isomorphism $\Psi : H^p \to h^p$, where

$$h^p = \Big\{v = (\mathbf{v}_1, \mathbf{v}_2, \cdots) : |v|_p^2 = \sum_{j=1}^{+\infty}(2\pi j)^{2p+1}|\mathbf{v}_j|^2 < \infty,\ \mathbf{v}_j \in \mathbb{R}^2,\ j \in \mathbb{N}\Big\}.$$

We introduce the weighted l^1-space h_I^p,

$$h_I^p = \Big\{I = (I_1, I_2, \dots) \in \mathbb{R}^\infty : |I|_p^\sim = 2\sum_{j=1}^{+\infty}(2\pi j)^{2p+1}|I_j| < \infty\Big\},$$

and the mapping π_I :

$$\pi_I : h^p \to h_I^p, \quad (\mathbf{v}_1, \dots) \mapsto (I_1, \dots),\ I_j = \frac{1}{2}\mathbf{v}_j^t\mathbf{v}_j,\ j \in \mathbb{N}. \tag{2.1.4}$$

Obviously, π_I is continuous, $|\pi_I(v)|_p^\sim = |v|_p^2$ and its image $h_{I+}^p = \pi_I(h^p)$ is the positive octant of h_I^p.

We wish to study the long-time behavior of solutions for equation (2.1.1). Accordingly, fix some

$$\zeta_0 \geqslant 0, \quad p \geqslant 3, \quad T > 0,$$

and assume

Assumption A : *(i) For $u_0 \in H^p$, there exists a unique solution $u(\cdot) \in C([0,T], H^p)$ of (2.1.1) with $u(0) = u_0$. It satisfies*

$$||u||_p \leqslant C(T, p, ||u_0||_p), \quad 0 \leqslant t \leqslant T\epsilon^{-1}.$$

(ii) There exists a $p' = p'(p) < p$ such that for $q \in [p', p]$, the perturbation term defines an analytic mapping

$$H^q \to H^{q+\zeta_0}, \quad u(\cdot) \mapsto f(u)(\cdot).$$

We are mainly concerned with the behavior of the actions $I(u(t)) \in \mathbb{R}_+^\infty$ on time interval $[0, T\epsilon^{-1}]$. For this end, write the perturbed KdV (2.1.1), using slow time $\tau = \epsilon t$ and the v-variables :

$$\frac{dv}{d\tau} = \epsilon^{-1} d\Psi(u)V(u) + P(v). \tag{2.1.5}$$

Here $V(u) = -u_{xxx} + 6uu_x$ is the vector field of KdV and $P(v)$ is the perturbation term, written in the v-variables. In the action-angle variables (I, φ) this equation reads :

$$\frac{dI}{d\tau} = F(I, \varphi), \quad \frac{d\varphi}{d\tau} = \epsilon^{-1} W(I) + G(I, \varphi). \tag{2.1.6}$$

Here $I \in \mathbb{R}^\infty$ and $\varphi \in \mathbb{T}^\infty$, where $\mathbb{T}^\infty := \{\theta = (\theta_j)_{j \geqslant 1}, \theta_j \in \mathbb{T}\}$ is the infinite-dimensional torus, endowed with the Tikhonov topology. The two functions $F(I, \varphi)$ and $G(I, \varphi)$ represent the perturbation term f, written in the action-angle variables, see below (2.2.3) and (2.2.4).

It is well established that for a perturbed integrable finite-dimensional system,

$$\dot{I} = \epsilon f(I, \varphi), \quad \dot{\varphi} = W(I) + \epsilon g(I, \varphi), \quad \epsilon << 1,$$

where $I \in \mathbb{R}^n$, $\varphi \in \mathbb{T}^n$, on time intervals of order ϵ^{-1} the actions $I(t)$ may be well approximated by solutions of the averaged equation :

$$\dot{J} = \epsilon \langle f \rangle (J), \quad \langle f \rangle (J) = \int_{\mathbb{T}^n} f(J, \varphi) d\varphi,$$

provided that the initial data $(I(0), \varphi(0))$ are typical (see [57]). This assertion is known as the *averaging principle*. But in the infinite dimensional case, there is no similar general result. Several theorems are available for different situations, mainly in the context of perturbations of linear equations, see [5] and references therein. When the unperturbed system is nonlinear (like KdV), results are rare. In [53, 49], S. Kuksin and A. Piatniski proved that the averaging principle holds for the randomly perturbed KdV equation of the form :

$$\dot{u} - \epsilon u_{xx} + u_{xxx} - 6uu_x = \sqrt{\epsilon}\eta(t, x), \quad x \in \mathbb{S}^1, \quad \int u dx = \int \eta dx = 0, \tag{2.1.7}$$

where the force η is a white noise in t, is smooth in x and is non-degenerate. Our goal in this work is to justify the averaging principle for the KdV equation with deterministic perturbations, using the Anosov scheme (see [57]), exploited earlier in the finite dimensional situation. The main technical difficulty to achieve this goal comes from the fact that to perform the scheme one has to use a measure in the function space which is 'quasi-invariant' under the flow of the perturbed equation. We now give the exact definition of it.

Let S_ϵ^τ, $0 \leqslant \tau \leqslant T$, be the flow-maps of equation (2.1.5) on h^p and denote

$$B_p^v(M) = \Big\{ v \in h^p : \ |v|_p \leqslant M \Big\}.$$

Definition 2.1.1. *1) A measure μ on h^p is called regular if for any analytic function g on h^p such that $g \not\equiv 0$, we have $\mu(\{v \in h^p : g(v) = 0\}) = 0$.*

2) A measure μ on h^p is said to be ϵ-quasi-invariant for equation (2.1.5) on the ball $B_p^v(M)$ if it is regular, $0 < \mu(B_p^v(M)) < \infty$ and there exists a constant $C(T, M)$ such that for any Borel set $A \subset B_p^v(M)$, we have

$$e^{-\tau C(T,M)}\mu(A) \leqslant \mu(S_\epsilon^\tau(A)) \leqslant e^{\tau C(T,M)}\mu(A). \tag{2.1.8}$$

Similarly, these definitions can be carried to measures on the space H^p and the flow-maps of equation (2.1.1) on H^p.

The ϵ-quasi-invariant measure is needed to guarantee that a small 'bad' set which we have to prohibit for a solution of the perturbed equation at a time $t > 0$ corresponds to a small set of initial data.

Inspired by finite averaging theory, we consider an averaged equation for the actions $I(\cdot)$:

$$\frac{dJ}{d\tau} = \langle F\rangle(J), \quad \langle F\rangle(J) = \int_{\mathbb{T}^\infty} F(J, \varphi)d\varphi, \tag{2.1.9}$$

where $d\varphi$ is the Haar measure on $\mathbb{T}^\infty$. It turns out that $\langle F\rangle(J)$ defines a Lipschitz vector filed in h_I^p (see (2.4.18) below). So equation (2.1.9) is well-posed, at least locally. Our task is to study the relation between the actions $I(\tau)$ of solutions for equation (2.1.6) and solutions $J(\tau)$ of equation (2.1.9), for $\tau \in [0, T]$.

The main result of this chapter is the following theorem, in which $u^\epsilon(t)$ denotes solutions for equation (2.1.1), $v^\epsilon(\tau) = \Psi\Big(u^\epsilon(\epsilon^{-1}\tau)\Big)$ denotes solutions for (2.1.5) and $I(v^\epsilon), \varphi(v^\epsilon)$ are their action-angle variables. By Assumption A, for $\tau \in [0, T]$,

$$|I(v^\epsilon(\tau))|_p^\sim \leqslant C_1(|I(v^\epsilon(0))|_p^\sim).$$

Theorem 2.1.2. *Fix any $M > 0$. Suppose that assumption A holds and equation (2.1.1) has an ϵ-quasi-invariant measure μ on $B_p^v(M)$. Then*

(i) For any $\rho > 0$ and any $q < p + \frac{1}{2}\min\{\zeta_0, 1\}$, there exists $\delta_\rho > 0$, $\epsilon_{\rho,q} > 0$ and a Borel subset $\Gamma_{\rho,q}^\epsilon \subset B_p^v(M)$ such that

$$\lim_{\epsilon\to 0}\mu(B_p^v(M)\setminus\Gamma_{\rho,q}^\epsilon) = 0, \tag{2.1.10}$$

and for $\epsilon \leqslant \epsilon_{\rho,q}$, we have that if $v^\epsilon(0) \in \Gamma_{\rho,q}^\epsilon$, then

$$|I(v^\epsilon(\tau)) - J(\tau)|_q^\sim \leqslant \rho, \quad \textit{for} \quad 0 \leqslant \tau \leqslant \min\{T, \mathcal{T}(I_0^\epsilon)\}. \tag{2.1.11}$$

Here $I_0^\epsilon = I(v^\epsilon(0))$, $J(\cdot)$ is the unique solution of the averaged equation (2.1.9) with any initial data $J_0 \in h_I^p$, satisfying $|J_0 - I_0^\epsilon|_q^\sim \leqslant \delta_\rho$, and

$$\mathcal{T}(I_0^\epsilon) = \min\Big\{\tau : |J(\tau)|_p^\sim \geqslant C_1(|I_0^\epsilon|_p^\sim) + 1\Big\}.$$

(ii) Let $\lambda_\epsilon^{v_0}$ be the probability measure on $\mathbb{T}^\infty$, defined by the relation

$$\int_{\mathbb{T}^\infty} f(\varphi)\, d\lambda_\epsilon^{v_0}(d\varphi) = \frac{1}{T}\int_0^T f(\varphi(v^\epsilon(\tau))d\tau, \quad \forall f \in C(\mathbb{T}^\infty),$$

where $v_0 = v^\epsilon(0) \in B_p(M)$. Then the averaged measure

$$\lambda_\epsilon := \frac{1}{\mu\Big(B_p(M)\Big)}\int_{B_p(M)} \lambda_\epsilon^{v_0} d\mu(v_0)$$

converges weakly, as $\epsilon \to 0$, to the Haar measure $d\varphi$ on $\mathbb{T}^\infty$.

Remark 2.1.3. *1) Assume that an ϵ-quasi-invariant measure μ depends on ϵ, i.e. $\mu = \mu_\epsilon$. Then the same conclusion holds with μ replaced by μ_ϵ, if μ_ϵ satisfies some consistency conditions. See subsection 2.4.3.*

2) Item (ii) of Assumption A may be removed if the perturbation is hamiltonian. See the end of subsection 2.4.1.

Toward the existence of ϵ-quasi-invariant measures, let us consider a class of Gaussian measures μ_0 on the Hilbert space h^p :

$$\mu_0 := \prod_{j=1}^{\infty} \frac{(2\pi j)^{1+2p}}{2\pi\sigma_j} \exp\{-\frac{(2\pi j)^{1+2p}|\mathbf{v}_j|^2}{2\sigma_j}\} d\mathbf{v}_j, \tag{2.1.12}$$

where $d\mathbf{v}_j$, $j \geqslant 1$, is the Lebesgue measure on $\mathbb{R}^2$. We recall that (2.1.12) is a well-defined probability measure on h^p if and only if $\sum \sigma_j < \infty$(see [11]). It is regular in the sense of Definition 2.1.1 and is non-degenerated in the sense that its support equals to h^p (see [11, 12]). From (2.1.3), it is easy to see that this kind of measures are invariant for KdV.

For any $\zeta_0' > 1$, we say the measure μ_0 is ζ_0'-*admissible* if the σ_j in (2.1.12) satisfies $0 < j^{-\zeta_0'}/\sigma_j < const$ for all $j \in \mathbb{N}$.

Theorem 2.1.4. *If Assumption A holds and*
$(ii)'$ There exists $\zeta_0' > 1$ such that the operator defined by

$$h^p \to h^{p+\zeta_0'} : \ v \mapsto P(v)$$

(see (2.1.5)) is analytic. Then every ζ_0'-admissible measure μ_0 is ϵ-quasi-invariant for equation (2.1.1) on h^p.

However, the conditions $(ii)'$ is not easy to verify due to the complexity of the nonlinear Fourier transform. So we give here another sufficient condition for existence of ϵ-quasi-invariant measure by restricting Assumption A.

As is known, for solutions of KdV, there are countably many conservation laws $\mathcal{J}_n(u)$, $n \geqslant 0$, where $\mathcal{J}_0(u) = \frac{1}{2}||u||_0^2$ and

$$\mathcal{J}_n(u) = \int_{\mathbb{T}} \Big\{\frac{1}{2}(\partial_x^n u)^2 + c_n u(\partial_x^{n-1}u)^2 + \mathcal{Q}_n(u, \dots, \partial_x^{n-2}u)\Big\} dx,$$

for $n \geqslant 1$, where c_n are real constants, and $\mathcal{Q}_n$ are polynomial in their arguments (see e.g. [38]). Let μ_n be the Gibbs measures on the space H^n, generated by the conservation law $\mathcal{J}_n(u)$. They are invariant for KdV ([75]). We have the following :

Theorem 2.1.5. *Let $p \in \mathbf{N}$. Then if Assumption A holds with $\zeta_0 \geqslant 2$, then the Gibbs measure μ_p is ϵ-quasi-invariant for the perturbed KdV (0.0.1).*

We point out straightly that this condition is not optimal (see Remark 2.5.11).

Note that μ_0 (2.1.12) also is a Gibbs measure for KdV, written in the Birkhoff coordinates (1.4.14), since formally it may be written as $\mu_0 = Z^{-1}\exp\{-\langle Qv, v\rangle\}dv$, where $\langle Qv, v\rangle = \sum c_j|\mathbf{v}_j|^2$ is an integrals of motion for KdV (the statistical sum $Z = \infty$, so indeed this is a formal expression). Some recent results (see, e.g. [75, 14, 18]) show that the Gibbs measure is an efficient tool to study nonlinear partial differential equations. There are mainly two kinds of applications : the recurrent properties given by the Poincaré Recurrent Theorem and the almost sure (in the sense of the Gibbs measure) global well-posedness for 'rough' initial data. Here we give a new application of the Gibbs measure. We show that in the averaging theory for perturbed KdV, the Gibbs measure plays a role which the Lebesgue measure plays in the classic finite dimensional averaging theory. This indicates that the Gibbs measure is important not only for the study of the original PDEs but also for the study of its perturbations.

Concerning the validity of Assumption A, particularly, we have :

Proposition 2.1.6. *The Assumption A holds for the perturbed KdV equation :*

$$u_t + u_{xxx} - 6uu_x = \epsilon f(x), \tag{2.1.13}$$

where $f(x)$ is a smooth function on the circle with zero mean value.

The equation (2.1.13) can be viewed as a model for shallow water wave propagation under small external force.

This chapter is organized as follows : Section 2.2 is about some important properties of the nonlinear Fourier transform and the action-angle form of the perturbed KdV (2.1.1). We discuss the averaged equation in Section 2.3. The Theorem 2.1.2 is proved in Section 2.4. We will discuss the existence of ϵ-quasi-invariant measures in Section 2.5. Finally in Section 2.6, we prove Proposition 2.1.6.

Agreements. Analyticity of maps $B_1 \to B_2$ between Banach spaces B_1 and B_2, which are the real parts of complex spaces B_1^c and B_2^c, is understood in the sense of Fréchet. All analytic maps that we consider possess the following additional property : for any R, a map extends to a bounded analytical mapping in a complex ($\delta_R > 0$)-neighborhood of the ball $\{|u|_{B_1} < R\}$ in B_1^c. We call such analytic maps uniformly analytic.

2.2 The perturbed KdV in action-angle variables

First we recall some results on the integrability of the KdV equation $(0.1)_{\epsilon=0}$ which have been discussed in Section 1.4.

2.2.1 Nonlinear Fourier transform for KdV

Theorem 2.2.1. *(see [38]) There exists an analytic diffeomorphism $\Psi : H^0 \mapsto h^0$ and an analytic functional K on h^1 of the form $K(v) = \tilde{K}(I(v))$, where the function $\tilde{K}(I)$ is analytic in a suitable neighborhood of the octant h^1_{I+} in h^1_I, with the following properties :*

(i) For any $p \in [-1, +\infty)$, the mapping Ψ defines an analytic diffeomorphism $\Psi : H^p \mapsto h^p$.

(ii) The differential $d\Psi(0)$ is the operator $\sum u_s e_s \mapsto v, v_s = |2\pi s|^{-1/2} u_s$.

(iii) A curve $u \in C^1(0, T; H^0)$ is a solution of the KdV equation $(2.1.1)_{\epsilon=0}$ if and only if $v(t) = \Psi(u(t))$ satisfies the equation

$$\dot{\mathbf{v}}_j = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{\partial \tilde{K}}{\partial I_j}(I)\mathbf{v}_j, \quad \mathbf{v}_j = (v_j, v_{-j})^t \in \mathbb{R}^2,\ j \in \mathbb{N}. \tag{2.2.1}$$

The coordinates $v = \Psi(u)$ are called the *Birkhoff coordinates*, and the form (1.1) of KdV is its *Birkhoff normal form*

Since the maps Ψ and Ψ^{-1} are analytic, then for $m = 0, 1, 2 \dots$, we have

$$||d^j\Psi(u)||_m \leqslant P_m(||u||_m), \quad ||d^j\Psi^{-1}(v)||_m \leqslant Q_m(|v|_m), \quad j = 0, 1, 2,$$

where P_m and Q_m are continuous functions.

A remarkable property of the nonlinear Fourier transform Ψ is its quasi-linearity. It means :

Theorem 2.2.2. *(see [52, 39]) If $p \geqslant 0$, then the map $\Psi - d\Psi(0) : H^p \to h^{p+1}$ is analytic.*

We denote

$$W(I) = (W_1, W_2, \dots),\ \ W_k(I) = \frac{\partial \tilde{K}}{\partial I_k}(I),\ \ k = 1, 2, \dots.$$

Lemma 2.2.3. *(see [47], appendix 6) For any $n \in \mathbb{N}$, if $I_{n+1} = I_{n+2} = \cdots = 0$, then*

$$det\Big((\frac{\partial W_i}{\partial I_j})_{1 \leqslant i,j \leqslant n} \Big) \neq 0.$$

Let l^∞_{-1} be the Banach space of all real sequences $l = (l_1, l_2, \dots)$ with the norm

$$|l|_{-1} = \sup_{n \geqslant 1} n^{-1}|l_n| < \infty.$$

Denote $\boldsymbol{\kappa} = (\kappa_n)_{n\geqslant 1}$, where $\kappa_n = (2\pi n)^3$.

Lemma 2.2.4. *(see [38], Thoerem 15.4) The normalized frequency map*

$$I \mapsto W(I) - \boldsymbol{\kappa}$$

is real analytic as a map from h^1_{I+} to l^∞_{-1}.

2.2.2 Equation (2.1.1) in the Birkhoff coordinates.

For $k = 1, 2 \dots$ we denote :

$$\Psi_k : H^m \to \mathbb{R}^2, \quad \Psi_k(u) = \mathbf{v}_k,$$

where $\Psi(u) = v = (\mathbf{v}_1, \mathbf{v}_2, \dots)$. Let $u(t)$ be a solution of equation (2.1.1). Passing to the slow time $\tau = \epsilon t$ and denoting $\dot{}$ to be $\frac{d}{d\tau}$, we get

$$\dot{\mathbf{v}}_k = d\Psi_k(u)(\epsilon^{-1}V(u)) + \mathcal{P}_k(v), \quad k \geqslant 1, \tag{2.2.2}$$

where $V(u) = -u_{xxx} + 6uu_x$ and $\mathcal{P}_k(v) = d\Psi_k(\Psi^{-1}(v))(f(\Psi^{-1}(v)))$. Since the action $I_k(v) = \frac{1}{2}|\Psi_k|^2$ is an integral of motion for KdV equation $(2.1.1)_{\epsilon=0}$, we have

$$\dot{I}_k = (\mathcal{P}_k(v), \mathbf{v}_k) := F_k(v). \tag{2.2.3}$$

Here and below $(\cdot,\cdot)$ indicates the scalar product in $\mathbb{R}^2$.

For $k \geqslant 1$ defines the angle $\varphi_k = \arctan(\frac{v_{-k}}{v_k})$ if $\mathbf{v}_k \neq 0$ and $\varphi_k = 0$ if $\mathbf{v}_k = 0$. Using equation (2.2.2), we get

$$\dot{\varphi}_k = \epsilon^{-1}W_k(I) + |\mathbf{v}_k|^{-2}(d\Psi_k(u)f(x,u), \mathbf{v}_k^{\perp}), \quad \text{if} \quad \mathbf{v}_k \neq 0, \tag{2.2.4}$$

where $\mathbf{v}_k^{\perp} = (-v_{-k}, v_k)$. Denoting for brevity, the vector field in equation (1.4) by $\epsilon^{-1}W_k(I) + G_k(v)$, we rewrite the equation for the pair $(I_k, \varphi_k)(k \geqslant 1)$ as

$$\begin{aligned} \dot{I}_k &= F_k(v) = F_k(I,\varphi), \\ \dot{\varphi}_k &= \epsilon^{-1}W_k(I) + G_k(v). \end{aligned} \tag{2.2.5}$$

We set

$$F(I,\varphi) = (F_1(I,\varphi), F_2(I,\varphi), \dots).$$

Denote

$$\bar{\zeta}_0 = \min\{1, \zeta_0\}.$$

For any $q \in [p', p]$, define a map $\mathcal{P}$ as

$$\mathcal{P}: \ h^q \to h^{q+\bar{\zeta}_0}, \quad v \mapsto (\mathcal{P}_1(v), \dots).$$

Clearly, $\mathcal{P}(v) = d\Psi\Big(\Psi^{-1}(v)\Big)\Big(f(\Psi^{-1}v)\Big)$. Then Theorem 2.2.2 and Assumption A imply that the map $\mathcal{P}$ is analytic. Using (2.2.3), for any $k \in \mathbb{N}$, we have

$$(2\pi k)^{2q+1+\bar{\zeta}_0}|F_k(v)| \leqslant (2\pi k)^{2q+1}|\mathbf{v}_k|^2 + (2\pi k)^{2q+1+2\bar{\zeta}_0}|\mathcal{P}_k(v)|^2.$$

Therefore,

$$|F(I,\varphi)|^{\sim}_{q+\bar{\zeta}_0/2} \leqslant |v|_q^2 + |\mathcal{P}(v)|^2_{q+\bar{\zeta}_0} \leqslant C(|v|_q). \tag{2.2.6}$$

In the following lemma P_k and P_k^j are some fixed continuous functions.

Lemma 2.2.5. *For $k, j \in \mathbb{N}$ and each $q \in [p', p]$, we have :*

(i) The function $F_k(v)$ is analytic in each space h^q.

(ii) For any $\delta > 0$, the function $G_k(v)\chi_{\{I_k \geqslant \delta\}}$ is bounded by $\delta^{-1/2}P_k(|v|_q)$.

(iii) For any $\delta > 0$, the function $\frac{\partial F_k}{\partial I_j}(I,\varphi)\chi_{\{I_j \geqslant \delta\}}$ is bounded by $\delta^{-1/2}P_k^j(|v|_q)$.

(iv) The function $\frac{\partial F_k}{\partial \varphi_j}(I,\varphi)$ is bounded by $P_k^j(|v|_q)$, and for any $n \in \mathbb{N}$ and $(I_1, \dots, I_n) \in \mathbb{R}^n_+$, the function $\varphi \mapsto F_k(I_1, \varphi_1, \dots, I_n, \varphi_n, 0, \dots)$ is smooth on $\mathbb{T}^n$.

Démonstration. Items (i) and (ii) follow directly from Theorem 1.1. Items (iii) and (iv) follow from item (i) and the chain-rule :

$$\begin{aligned} \frac{\partial F_k}{\partial \varphi_j} &= \sqrt{2I_j}\Big(\frac{\partial F_k}{\partial v_{-j}}\cos(\varphi_j) - \frac{\partial F_k}{\partial v_j}\sin(\varphi_j)\Big), \\ \frac{\partial F_k}{\partial I_j} &= (\sqrt{2I_j})^{-1}\Big(\frac{\partial F_k}{\partial v_j}\cos(\varphi_j) + \frac{\partial F_k}{\partial v_{-j}}\sin(\varphi_j)\Big). \end{aligned}$$

$\square$

We denote

$$\Pi_I : h^p \to h_I^p, \quad \Pi_I(v) = I(v),$$
$$\Pi_{I,\varphi} : h^p \to h_I^p \times \mathbb{T}^\infty, \quad \Pi_{I,\varphi}(v) = (I(v), \varphi(v)).$$

Abusing notation, we will identify v with $(I, \varphi) = \Pi_{I,\varphi}(v)$.

Definition 2.2.6. *We say that a curve* $\big(I(\tau), \varphi(\tau)\big)$, $\tau \in [0, \tau_1]$, *is a regular solution of equation (2.2.5), if there exists a solution* $u(\cdot) \in H^p$ *of equation (2.1.1) such that*

$$\Pi_{I,\varphi}\Big(\Psi(u(\epsilon^{-1}\tau))\Big) = (I(\tau), \varphi(\tau)), \quad \tau \in [0, \tau_1].$$

Note that if $(I(\tau), \varphi(\tau))$ is a regular solution, then each $I_j(\tau)$ is a C^1-function, while $\varphi_j(\tau)$ may be discontinuous at points τ, where $I_j(\tau) = 0$.

2.3 Averaged equation

For a function f on a Hilbert space H, we write $f \in Lip_{loc}(H)$ if

$$|f(u_1) - f(u_2)| \leqslant P(R)||u_1 - u_2||, \quad \text{if} \quad ||u_1||, ||u_2|| \leqslant R, \tag{2.3.1}$$

for a suitable continuous function P which depends on f. By the Cauchy inequality, any analytic function on H belongs to $Lip_{loc}(H)$ (see Agreements). In particularly, for any $k \geqslant 1$,

$$W_k(I) \in Lip_{loc}(h_I^q), \;\; q \geqslant 1, \quad \text{and} \quad F_k(v) \in Lip_{loc}(h^q), \;\; q \in [p'(p), p]. \tag{2.3.2}$$

Let $f \in Lip_{loc}(h^{p_0})$ for some $p_0 \geqslant 0$ and $v \in h^{p_1}$, $p_1 > p_0$. Denoting by Π^M, $M \geqslant 1$, the projection

$$\Pi^M : h^0 \to h^0, \quad (\mathbf{v}_1, \mathbf{v}_2, \dots) \mapsto (\mathbf{v}_1, \dots, \mathbf{v}_M, 0, \dots),$$

we have $|v - \Pi^M v|_{p_0} \leqslant (2\pi M)^{-(p_1 - p_0)}|v|_{p_1}$. Accordingly,

$$|f(v) - f(\Pi^M v)| \leqslant P(|v|_{p_1})(2\pi M)^{-(p_1 - p_0)}. \tag{2.3.3}$$

The torus $\mathbb{T}^\infty$ acts on the space h^0 by the linear transformations Φ_θ, $\theta \in \mathbb{T}^\infty$, where $\Phi_\theta : (I, \varphi) \mapsto (I, \varphi + \theta)$. For a function $f \in Lip_{loc}(h^p)$, we define the averaging in all angles as

$$\langle f \rangle(v) = \int_{\mathbb{T}^\infty} f(\Phi_\theta(v)) d\theta,$$

where $d\theta$ is the Haar measure on $\mathbb{T}^\infty$. Clearly, the average $\langle f \rangle$ is independent of φ. Thus $\langle f \rangle$ can be written as $\langle f \rangle(I)$.

Extend the mapping π_I to a complex mapping $h^p \otimes \mathbb{C} \to h_I^p \otimes \mathbb{C}$, using the same formulas (2.1.4). Obviously, if $\mathcal{O}$ is a complex neighbourhood of h^p, then $\pi_I^c(\mathcal{O})$ is a complex neighbourhood of h_I^p.

Lemma 2.3.1. *(See [53], Lemma 4.2) Let $f \in Lip_{loc}(h^p)$, then*

(i) The function $\langle f\rangle(v)$ satisfy (2.3.1) with the same function P as f and take the same value at the origin.

(ii) This function is smooth (analytic) if f is. If $f(v)$ is analytic in a complex neighbourhood $\mathcal{O}$ of h^p, then $\langle f\rangle(I)$ is analytic in the complex neighbourhood $\pi_I^c(\mathcal{O})$ of h_I^p.

For any $\bar q \in [p', p]$, we consider the mapping defined by

$$\langle F\rangle : h_I^{\bar q} \to h_I^{\bar q+\bar\zeta_0/2}, \quad J \mapsto \langle F\rangle(J),$$

where $\langle F\rangle(J) = (\langle F_1\rangle(J), \langle F_2\rangle(J), \dots)$.

Corollary 2.3.2. *For every $\bar q \in [p', p]$, the mapping $\langle F\rangle$ is analytic as a map from the space $h_I^{\bar q}$ to $h_I^{\bar q+\bar\zeta_0/2}$.*

Démonstration. The mapping $\mathcal{P}(v)$ extends analytically to a complex neighbourhood $\mathcal{O}$ of $h^{\bar q}$ (see Agreements). Then by (2.2.3), the functions $F_j(v)$, $j \in \mathbb{N}$ are analytic in $\mathcal{O}$. Hence it follows from Lemma 2.1 that for each $j \in \mathbb{N}$, the function $\langle F_j\rangle$ is analytic in the complex neighbourhood $\pi_I^c(\mathcal{O})$ of $h_I^{\bar q}$. By (2.2.6), the mapping $\langle F\rangle$ is locally bounded on $\pi_I^c(\mathcal{O})$. It is well known that the analyticity of each coordinate function and the locally boundness of the maps imply the analyticity of the maps (see, e.g. [3]). This finishes the proof of the corollary. $\square$

We recall that a vector $\omega \in \mathbb{R}^n$ is called *non-resonant* if

$$\omega \cdot k \neq 0, \quad \forall\ k \in \mathbb{Z}^n \setminus \{0\}.$$

Denote by $C^{0+1}(\mathbb{T}^n)$ the set of all Lipschitz functions on $\mathbb{T}^n$. The following lemma is a version of the classical Weyl theorem.

Lemma 2.3.3. *Let $f \in C^{0+1}(\mathbb{T}^n)$ for some $n \in \mathbb{N}$. For any $\delta > 0$ and any non-resonant vector $\omega \in \mathbb{R}^n$, there exists $T_0 > 0$ such that if $T \geqslant T_0$, then*

$$\left|\frac{1}{T}\int_0^T f(x_0+\omega t)dt - \langle f\rangle\right| \leqslant \delta,$$

uniformly in $x_0 \in \mathbb{T}^n$.

Démonstration. Let us write $f(x)$ as the Fourier series $f(x) = \sum f_k e^{ik\cdot x}$. Since the Fourier series of a Lipschitz function converges uniformly (see [71]), for any $\epsilon > 0$ we may find $R = R_\epsilon$ such that $\left|\sum_{|k|>R} f_k e^{ik\cdot x}\right| \leqslant \frac{\epsilon}{2}$ for all x. Now it is enough to show that

$$\left|\frac{1}{T}\int_0^T f_R(x_0+\omega t)dt - f_0\right| \leqslant \frac{\epsilon}{2}, \quad \forall T \geqslant T_\epsilon, \tag{2.3.4}$$

for a suitable T_ϵ, where $f_R(x) = \sum_{|k|\leqslant R} f_k e^{ik\cdot x}$. Observing that

$$\left|\frac{1}{T}\int_0^T e^{ik\cdot(x_0+\omega t)}dt\right| \leqslant \frac{2}{T|k\cdot\omega|},$$

for each nonzero k. Therefore the l.h.s of (2.3) is smaller than

$$\frac{2}{T}\Big(\inf_{|k|\leqslant R}|k\cdot\omega|\Big)^{-1}\sum_{|k|\leqslant R}|f_k|.$$

The assertion of the lemma follows. $\square$

2.4 Proof of the main theorem

In this section we prove Theorem 2.1.2 by developing a suitable infinite-dimensional version of the Anosov scheme (see [57]), and by studying the behavior of the regular solutions of equation (2.2.5) and the corresponding solutions of (2.1.1).

Assume $u(0) = u_0 \in H^p$. So

$$\Pi_{I,\varphi}(\Psi(u_0)) = (I_0, \varphi_0) \in h^p_{I+} \times \mathbb{T}^\infty. \tag{2.4.1}$$

We denote

$$B^I_p(M) = \{I \in h^p_{I+} : |I|^\sim_p \leqslant M\}.$$

Without loss of generality, we assume that $T = 1$. Fix any $M_0 > 0$. Let

$$(I_0, \varphi_0) \in B^I_p(M_0) \times \mathbb{T}^\infty := \Gamma_0,$$

that is,

$$v_0 = \Psi(u_0) \in B^v_p(\sqrt{M_0}).$$

We pass to the slow time $\tau = \epsilon t$. Let $(I(\cdot), \varphi(\cdot))$ be a regular solution of the system (2.2.5) with $(I(0), \varphi(0)) = (I_0, \varphi_0)$. We will also write it as $(I^\epsilon(\cdot), \varphi^\epsilon(\cdot))$ when we want to stress the dependence on ϵ. Then by assumption A, there exists $M_1 \geqslant M_0$ such that

$$I(\tau) \in B^I_p(M_1), \quad \tau \in [0, 1]. \tag{2.4.2}$$

By (2.2.6), we know that

$$|F(I, \varphi)|^\sim_1 \leqslant C_{M_1}, \quad \forall\, (I, \varphi) \in B^I_p(M_1) \times \mathbb{T}^\infty, \tag{2.4.3}$$

where the constant C_{M_1} depends only on M_1.

We denote $I^m = (I_1, \dots, I_m, 0, 0, \dots)$, $\varphi^m = (\varphi_1, \dots, \varphi_m, 0, 0, \dots)$, and $W^m(I) = (W_1(I), \dots, W_m(I), 0, 0, \dots)$, for any $m \in \mathbb{N}$.

2.4.1 Proof of assertion (i)

Fix any

$$n_0 \in \mathbb{N} \quad \text{and} \quad \rho > 0.$$

By (2.2), there exists $m_0 \in \mathbb{N}$ such that

$$|F_k(I, \varphi) - F_k(I^{m_0}, \varphi^{m_0})| \leqslant \rho, \quad \forall (I, \varphi) \in B^I_p(M_1) \times \mathbb{T}^\infty, \tag{2.4.4}$$

where $k = 1, \cdots, n_0$.

From now on, we always assume that

$$(I, \varphi) \in \Gamma_1 := B^I_p(M_1) \times \mathbb{T}^\infty, \quad \text{i.e.} \quad v \in B^v_p(\sqrt{M_1}),$$

and identify $v \in h^p$ with $(I, \varphi) = \Pi_{I,\varphi}(v)$.

By Lemma 2.2.5, we have

$$\begin{aligned}
&|G_j(I,\varphi)| \leqslant \frac{C_0(j,M_1)}{\sqrt{I_j}},\\
&|\frac{\partial F_k}{\partial I_j}(I,\varphi)| \leqslant \frac{C_0(k,j,M_1)}{\sqrt{I_j}},\\
&|\frac{\partial F_k}{\partial \varphi_j}(I,\varphi)| \leqslant C_0(k,j,M_1).
\end{aligned} \tag{2.4.5}$$

From Lemma 2.2.4 and Lemma 2.3.1, we know that

$$\begin{aligned}
&|W_j(I)-W_j(\bar I)| \leqslant C_1(j,M_1)|I-\bar I|_1,\\
&|\langle F_k\rangle(I)-\langle F_k\rangle(\bar I)| \leqslant C_1(k,j,M_1)|I-\bar I|_1.
\end{aligned} \tag{2.4.6}$$

By (2.3.2) we get

$$|F_k(v^{m_0})-F_k(\bar v^{m_0})| \leqslant C_2'(k,M_1)|v^{m_0}-\bar v^{m_0}|_p \leqslant C_2(k,m_0,M_1)|v^{m_0}-\bar v^{m_0}|_\infty, \tag{2.4.7}$$

where $|\cdot|_\infty$ is the l^∞-norm (here we have used the fact that norms in finite dimensional space are equivalent).

We denote

$$C_{M_1}^{n_0,m_0} = m_0\cdot \max\{C_0,C_1,C_2 : 1\leqslant j\leqslant m_0, 1\leqslant k\leqslant n_0\}.$$

Below we define a number of sets, depending on various parameters. All of them also depend on ρ and n_0, but this dependence is not indicated. For any $\delta>0$ and $T_0>0$, we define a subset

$$E(\delta,T_0)\subset\Gamma_1$$

as the collection of all $(I,\varphi)\in\Gamma_1$ such that for every $T\geqslant T_0$, we have,

$$\left|\frac{1}{T}\int_0^T [F_k(I^{m_0},\varphi^{m_0}+W^{m_0}(I)s)-\langle F_k\rangle(I^{m_0})]ds\right| \leqslant \delta, \quad \text{for} \quad k=1,\dots,n_0. \tag{2.4.8}$$

Let $\mathcal{S}_\epsilon^\tau$ be the flow generated by regular solutions of the system (1.5). We define two more groups of sets.

$$\Delta(\tau)=\Delta(\tau,\epsilon,\delta,T_0,I,\varphi) := \{\tau_1\in[0,\tau] : \mathcal{S}_\epsilon^{\tau_1}(I,\varphi)\notin E(\delta,T_0)\}.$$

$$N(\beta)=N(\beta,\epsilon,\delta,T_0) := \{(I,\varphi)\in\Gamma_0 : \mathrm{Mes}[\Delta(1,\epsilon,\delta,T_0,I,\varphi)]\leqslant\beta\}.$$

Here and below $\mathrm{Mes}[\cdot]$ stands for the Lebesgue measure in $\mathbb{R}$. We will indicate the dependence of the set $N(\beta)$ on n_0 and ρ as $N_{n_0,\rho}(\beta)$, when necessary.

By continuity, $E(\delta,T_0)$ is a closed subset of Γ_1 and $\Delta(\tau)$ is an open subset of $[0,\tau]$.

Lemma 2.4.1. *For $k=1,\dots,n_0$, the I_k-component of any regular solution of (2.2.5) with initial data in $N(\beta,\epsilon,\delta,T_0)$ can be written as :*

$$I_k(\tau)=I_k(0)+\int_0^\tau \langle F_k\rangle(I(s))ds+\Xi(\tau),$$

where for any $\gamma \in (0,1)$ the function $|\Xi(\tau)|$ is bounded on $[0,1]$ by

$$\begin{aligned}
&4C_{M_1}^{n_0,m_0}\Big\{\Big[2(\gamma+2T_0C_{M_1}\epsilon)^{1/2}\Big](T_0\epsilon+\beta+1)\\
&+\Big[\frac{T_0C_{M_1}\epsilon}{\gamma^{1/2}}+T_0C_{M_1}\epsilon+(\frac{T_0\epsilon}{2\gamma^{1/2}}+\frac{\epsilon C_{M_1}T_0^2}{3})\Big](T_0\epsilon+\beta+1)\Big\}\\
&+2C_{M_1}\beta+2\rho+2\delta+2C_{M_1}(T_0\epsilon+\beta).
\end{aligned}$$

Démonstration. For any $(I,\varphi)\in N(\beta)$, we consider the corresponding set $S(\tau)$. It is composed of open intervals of total length less than $\min\{\beta,\tau\}$. Thus at most $[\beta/(T_0\epsilon)]$ of them have length greater than or equal to $T_0\epsilon$. We denote these long intervals by (a_i,b_i), $1\leqslant i\leqslant d$, $d\leqslant\beta/(T_0\epsilon)$ and denote by $C(\tau)$ the complement of $\cup_{1\leqslant i\leqslant d}(a_i,b_i)$ in $[0,\tau]$.

By (3.5.6), we have

$$\int_0^\tau F_k(I(s),\varphi(s))ds=\int_{C(\tau)}F_k(I^{m_0}(s),\varphi^{m_0}(s))ds+\xi_1(\tau),$$

where $|\xi_1(\tau)|\leqslant C_{M_1}\beta+\rho\tau$.

The set $C(\tau)$ is composed of segments $[b_{i-1},a_i]$ (if necessary, we set $b_0=0$, and $a_{d+1}=\tau$). We proceed by dividing each segment $[b_{i-1},a_i]$ into shorter segments by points τ_j^i, where $b_i=\tau_1^i<\tau_2^i<\cdots<\tau_{n_i}^i=a_i$. The points τ_j^i lie outside the set $S(\tau)$ and $T_0\epsilon\leqslant t_{j+1}^i-t_j^i\leqslant 2T_0\epsilon$ except for the terminal segment containing the end points a_i, which may be shorter than $T_0\epsilon$.

This partition is constructed as follows :

— If $a_i-b_{i-1}\leqslant 2T_0\epsilon$, then we keep the whole segment with no subdivisions. ($\tau_1^i=b_{i-1}$,$\tau_2^i=a_i$).

— If $a_i-b_{i-1}>2T_0\epsilon$, we divide the segment in the following way :

  a) If $b_{i-1}+2T_0\epsilon$ does not belong to $S(\tau)$, we chose $t_2^i=b_{i-1}+2T_0\epsilon$, and continue by subdividing $[\tau_2^i,a_i]$;

  b) if $b_{i-1}+2T_0\epsilon$ belongs to $S(\tau)$, then there are points in $[b_{i-1}+T_0\epsilon,b_{i-1}+2T_0\epsilon]$ which do not, by definition of b_{i-1}. We set τ_2^i equal to one of these points and continue by subdividing $[\tau_2^i,a_i]$.

We will adopt the notation : $h_j^i=\tau_{j+1}^i-\tau_j^i$ and $s(i,j)=[\tau_j^i,\tau_{j+1}^i]$. So

$$C(\tau)=\bigcup_{i=1}^{d}\bigcup_{j=1}^{n_i-1}s(i,j),\ T_0\leqslant h_j^i=|s(i,j)|\leqslant 2T_0\epsilon,\ j\leqslant n_i-2.$$

By its definition, $C(\tau)$ contains at most $[\beta/(T_0\epsilon]+1$ segments $[b_{i-1},a_i]$, thus $C(\tau)$ contains at most $[\beta/(T_0\epsilon)]+1$ terminal subsegments of length less than $T_0\epsilon$. Since all other segments have length no less than $T_0\epsilon$ and $\tau\leqslant 1$, the number of these segments is not greater than $[\epsilon T_0]^{-1}$. So the total number of subsegments $s(i,j)$ is bounded by $1+[(\epsilon T_0)^{-1}]+[\beta/(T_0\epsilon)]$.

For each segment $s(i,j)$ we define a subset $\Lambda(i,j)$ of $\{1,2,\cdots,m_0\}$ in the following way :

$$l \in \Lambda(i,j) \quad \Longleftrightarrow \quad \exists \tau \in s(i,j), \quad I_l(\tau) < \gamma.$$

If $l \in \Lambda$, then by (3.5.4) we have

$$|I_l(\tau)| < 2T_0 C_{M_1}\epsilon + \gamma, \quad \tau \in s(i,j). \tag{2.4.9}$$

For $I = (I_1, I_2, \dots)$ and $\varphi = (\varphi_1, \varphi_2, \dots)$ we set

$$\lambda_{i,j}(I) = \hat{I}, \quad \lambda_{i,j}(\varphi) = \hat{\varphi},$$

where $\hat{\varphi} = (\hat{\varphi}_1, \hat{\varphi}_2, \dots)$ and $\hat{I} = (\hat{I}_1, \hat{I}_2, \dots)$ are defined by the following relation :

$$\text{If} \quad l \in \Lambda(i,j), \quad \text{then} \quad \hat{I}_l = 0, \ \ \hat{\varphi}_l = 0, \quad \text{else} \quad \hat{I}_l = I_l, \ \ \hat{\varphi}_l = \varphi_l.$$

We also denote $\lambda_{i,j}(I,\varphi) = (\lambda_{i,j}(I), \lambda_{i,j}(\varphi))$ and when the segment $s(i,j)$ is clearly indicated, we write for short $\lambda_{i,j}(I,\varphi) = (\hat{I}, \hat{\varphi})$.

Then on $s(i,j)$, using (3.5.9) and (2.4.9) we obtain

$$\begin{aligned}
&\int_{s(i,j)} \Big| F_k\Big(I^{m_0}(s), \varphi^{m_0}(s)\Big) - F_k\Big(\lambda_{i,j}\Big(I^{m_0}(s), \varphi^{m_0}(s)\Big)\Big)\Big| ds \\
&\leqslant \int_{s(i,j)} C_{M_1}^{n_0,m_0} \Big| I^{m_0}(s) - \lambda_{i,j}\Big(I^{m_0}(s)\Big)\Big|^{1/2} ds \\
&\leqslant 2C_{M_1}^{n_0,m_0}(\gamma + 2T_0 C_{M_1}\epsilon)^{1/2} T_0 \epsilon.
\end{aligned} \tag{2.4.10}$$

In Proposition 1-5 below, $k = 1, \dots, n_0$.

Proposition 1.

$$\int_{C(\tau)} F_k\Big(I^{m_0}(s), \varphi^{m_0}(s)\Big) ds = \sum_{i,j} \int_{s(i,j)} F_k\Big(I^{m_0}(\tau_j^i), \varphi^{m_0}(s)\Big) ds + \xi_2(\tau),$$

where

$$|\xi_2| \leqslant 4C_{M_1}^{n_0,m_0}\Big[(\gamma + 2T_0 C_{M_1}\epsilon)^{1/2} + \gamma^{-1/2} T_0 C_{M_1}\epsilon\Big](T_0\epsilon + \beta + 1). \tag{2.4.11}$$

Démonstration. We may write $\xi_2(\tau)$ as

$$\begin{aligned}
\xi_2(t) &= \sum_{i,j} \int_{s(i,j)} \Big[F_k\Big(I^{m_0}(s), \varphi^{m_0}(s)\Big) - F_k\Big(I^{m_0}(\tau_j^i), \varphi^{m_0}(s)\Big)\Big] ds \\
&:= \sum_{i,j} I(i,j).
\end{aligned}$$

For each $s(i,j)$, we have

$$\begin{aligned}
&\int_{s(i,j)} \Big| F_k\Big(\hat{I}^{m_0}(s), \hat{\varphi}^{m_0}(s)\Big) - F_k\Big(\hat{I}^{m_0}(\tau_j^i), \hat{\varphi}^{m_0}(s)\Big)\Big| ds \\
&\leqslant \int_{s(i,j)} \gamma^{-1/2} C_{M_1}^{n_0,m_0} \Big| \hat{I}^{m_0}(s) - \hat{I}^{m_0}(\tau_j^i)\Big| ds \\
&\leqslant 2\gamma^{-1/2} C_{M_1}^{n_0,m_0} C_{M_1} T_0^2 \epsilon^2.
\end{aligned} \tag{2.4.12}$$

We replace the integrand $F_k(I^{m_0}, \varphi^{m_0})$ by $F_k(\hat{I}^{m_0}, \hat{\varphi}^{m_0})$. Using (3.5.11) and (3.5.14) we obtain that

$$I(i,j) \leqslant 4C_{M_1}^{n_0,m_0}\Big[(\gamma + 2T_0C_{M_1}\epsilon)^{1/2} + \gamma^{-1/2}T_0C_{M_1}\epsilon\Big]T_0\epsilon.$$

The inequality (3.5.13) follows. □

On each subsegment $s(i,j)$, we now consider the unperturbed linear dynamics $\varphi_j^i(\tau)$ of the angles $\varphi^{m_0} \in \mathbb{T}^{m_0}$:

$$\varphi_j^i(\tau) = \varphi^{m_0}(\tau_j^i) + \epsilon^{-1}W^{m_0}(I(t_j^i))(\tau - \tau_j^i) \in \mathbb{T}^{m_0}, \quad \tau \in s(i,j).$$

Proposition 2.

$$\sum_{i,j}\int_{s(i,j)} F_k\Big(I^{m_0}(\tau_j^i), \varphi^{m_0}(s)\Big)ds = \sum_{i,j}\int_{s(i,j)} F_k\Big(I^{m_0}(\tau_j^i), \varphi_j^i(s)\Big)ds + \xi_3(\tau),$$

where

$$\begin{aligned} |\xi_3(\tau)| &\leqslant 4C_{M_1}^{n_0,m_0}(\gamma + 2T_0C_{M_1}\epsilon)^{1/2}(T_0\epsilon + \beta + 1) \\ &+ (C_{M_1}^{n_0,m_0})^2\Big(\frac{2T_0\epsilon}{\gamma} + \frac{4\epsilon C_{M_1}T_0^2}{3}\Big)(T_0\epsilon + \beta + 1). \end{aligned} \tag{2.4.13}$$

Démonstration. For each $s(i,j)$ we have

$$\begin{aligned} &\int_{s(i,j)} \Big|\lambda_{i,j}\Big(\varphi^{m_0}(s) - \varphi_j^i(s)\Big)\Big|ds \\ &\leqslant \int_{s(i,j)}\int_{\tau_j^i}^{s} \Big|\lambda_{i,j}\Big(G^{m_0}(I(s'), \varphi(s')) + \epsilon^{-1}W^{m_0}(I(s')) - \epsilon^{-1}W^{m_0}(I(\tau_j^i))\Big)\Big|ds'ds \\ &\leqslant \int_{s(i,j)}\int_{\tau_j^i}^{s} C_{M_1}^{n_0,m_0}\Big[\gamma^{-1/2} + \epsilon^{-1}|I(s') - I(\tau_j^i))|_1\Big]ds'ds \\ &\leqslant \int_{s(i,j)} C_{M_1}^{n_0,m_0}\Big[\gamma^{-1/2}(s - \tau_j^i) + \frac{1}{2}C_{M_1}\epsilon^{-1}(s - \tau_j^i)^2\Big]ds \\ &\leqslant C_{M_1}^{n_0,m_0}\Big(\frac{2T_0^2\epsilon^2}{\sqrt{\gamma}} + \frac{4C_{M_1}T_0^3\epsilon^2}{3}\Big). \end{aligned}$$

Here the first inequality comes from equation (2.2.4), and using (3.5.7) and (3.5.8) we can get the second inequality. The third one follows from (3.5.4).

Using again (3.5.7), we get

$$\begin{aligned} &\int_{s(i,j)} \Big[F_k\Big(\lambda_{i,j}\Big(I^{m_0}(\tau_j^i), \varphi^{m_0}(s)\Big)\Big) - F\Big(\lambda_{i,j}\Big(I^{m_0}(\tau_j^i), \varphi_j^i(s)\Big)\Big)\Big]ds \\ &\leqslant \int_{s(i,j)} C_{M_1}^{n_0,m_0}\Big|\lambda_{i,j}\Big(\varphi^{m_0}(s) - \varphi_j^i(s)\Big)\Big|ds \\ &\leqslant (C_{M_1}^{n_0,m_0})^2\Big(\frac{2T_0^2\epsilon^2}{\sqrt{\gamma}} + \frac{4C_{M_1}T_0^3\epsilon^2}{3}\Big). \end{aligned}$$

Therefore (3.5.15) holds for the same reason as (3.5.13). □

We will now compare the integral $\int_{s(i,j)} F_k(I^{m_0}(\tau_j^i), \varphi_j^i(s))ds$ with the average value $\langle F_k(I^{m_0}(\tau_j^i))\rangle h_j^i$.

Proposition 3.

$$\sum_{i,j} \int_{s(i,j)} F_k\Big(I^{m_0}(\tau_j^i), \varphi_j^i(s)\Big)ds = \sum_{i,j} h_j^i \langle F_k\rangle\Big(I^{m_0}(\tau_j^i)\Big) + \xi_4(\tau),$$

where

$$|\xi_4(\tau)| \leqslant 2\delta + 2C_{M_1}(T_0\epsilon + \beta). \tag{2.4.14}$$

Démonstration. We divide the set of segments $s(i,j)$ into two subsets Δ_1 and Δ_2. Namely, $s(i,j) \in \Delta_1$ if $h_j^i \geqslant T_0\epsilon$ and $s(i,j) \in \Delta_2$ otherwise.

(i) $s(i,j) \in \Delta_1$. In this case, by (2.4.8), we have

$$\Big| \int_{s(i,j)} \Big[F_k\Big(I^{m_0}(\tau_j^i), \varphi_j^i(s)\Big) - \langle F_k\rangle\Big(I^{m_0}(\tau_j^i)\Big)\Big]ds\Big| \leqslant \delta h_j^i.$$

So

$$\sum_{s(i,j)\in\Delta_1} \Big| \int_{s(i,j)} F_k\Big(I^{m_0}(\tau_j^i), \varphi_j^i(s)\Big)ds - \langle F_k\rangle\Big(I^{m_0}(\tau_j^i)\Big)h_j^i\Big| \leqslant \delta \sum_{s(i,j)\in\Delta_1} h_j^i \leqslant 2\delta.$$

(ii) $s(i,j) \in \Delta_2$. Now, using (3.5.4) we get

$$\Big| \int_{s(i,j)} F_k\Big(I^{m_0}(\tau_j^i), \varphi_j^i(s)\Big)ds - \langle F_k\rangle\Big(I^{m_0}(\tau_j^i)\Big)h_j^i\Big| \leqslant 2C_{M_1}h_j^i \leqslant 2C_{M_1}T_0\epsilon.$$

Since $\mathrm{Card}(\Delta_2) \leqslant (1 + \beta/(T_0\epsilon))$, then

$$\sum_{s(i,j)\in\Delta_2} \Big| \int_{s(i,j)} F\Big(I^{m_0}(\tau_j^i), \varphi_j^i(s)\Big)ds - \langle F_k\rangle\Big(I^{m_0}(\tau_j^i)\Big)h_j^i\Big| \leqslant 2C_{M_1}(\beta + T_0\epsilon).$$

This implies the inequality (3.5.16). □

Proposition 4.

$$\sum_{i,j} h_j^i \langle F_k\rangle\Big(I^{m_0}(\tau_j^i)\Big) = \int_{C(\tau)} \langle F_k\rangle\Big(I^{m_0}(s)\Big)ds + \xi_5(\tau),$$

where

$$|\xi_5(\tau)| \leqslant 4\epsilon C_{M_1}C_{M_1}^{n_0,m_0}T_0(T_0\epsilon + \beta + 1). \tag{2.4.15}$$

Démonstration. Indeed, as

$$|\xi_5(\tau)| = \Big|\sum_{i,j} \int_{s(i,j)} \Big[\langle F_k\rangle\Big(I^{m_0}(s)\Big) - \langle F_k\rangle\Big(I^{m_0}(\tau_j^i)\Big)\Big]ds\Big|,$$

using (3.5.4) and (3.5.8) we get

$$\begin{aligned}|\xi_5(\tau)| &\leqslant \sum_{i,j} \int_{s(i,j)} C_{M_1}^{n_0,m_0}|I^{m_0}(s) - I^{m_0}(\tau_j^i)|ds \\ &\leqslant \epsilon \sum_{i,j} C_{M_1}C_{M_1}^{n_0,m_0}(h_j^i)^2 \leqslant 4\epsilon C_{M_1}C_{M_1}^{n_0,m_0}T_0(T_0\epsilon + \beta + 1).\end{aligned}$$

□

Finally,

Proposition 5.

$$\int_{C(\tau)} \langle F_k\rangle \Big(I^{m_0}(s)\Big)ds = \int_0^\tau \langle F_k\rangle \Big(I(s)\Big)ds + \xi_6(\tau),$$

and $|\xi_6(\tau)|$ is bounded by $C_{M_1}\beta + \rho t$. $\square$

Gathering the estimates in Propositions 1-5, we obtain

$$\begin{aligned} I_k(\tau) &= I_k(0) + \int_0^\tau F_k\Big(I(s), \varphi(s)\Big)ds \\ &= I_k(0) + \epsilon \int_0^\tau \langle F_k\rangle \Big(I(s)\Big)ds + \Xi(\tau), \end{aligned}$$

where

$$\begin{aligned} |\Xi(\tau)| &\leqslant \sum_{i=1}^{6} |\xi_i(t)| \\ &\leqslant 4C_{M_1}^{n_0,m_0}\Big[2(\gamma + 2T_0C_{M_1}\epsilon)^{1/2} + \frac{T_0C_{M_1}\epsilon}{\gamma^{1/2}} + T_0C_{M_1}\epsilon \\ &\quad + \Big(\frac{T_0\epsilon}{2\gamma^{1/2}} + \frac{\epsilon C_{M_1}T_0^2}{3}\Big)\Big](T_0\epsilon + \beta + 1) + 2C_{M_1}\beta \\ &\quad + 2\rho + 2\delta + 2C_{M_1}(T_0\epsilon + \beta), \quad \tau \in [0,1]. \end{aligned}$$

Lemma 2.4.1 is proved. $\square$

Corollary 2.4.2. *For any $\bar\rho > 0$, with a suitable choice of ρ, γ, δ, T_0, β, the function $|\Xi(\tau)|$ in Lemma 2.4.1 can be made smaller than $\bar\rho$, if ϵ is small enough.*

Démonstration. We choose

$$\gamma = \epsilon^\alpha,\ T_0 = \epsilon^{-\sigma},\ \beta = \frac{\bar\rho}{9C_{M_1}},\ \delta = \rho = \frac{\bar\rho}{9}$$

with

$$1 - \frac{\alpha}{2} - \sigma > 0,\ 0 < \sigma < \frac{1}{2}.$$

Then for ϵ sufficiently small we have

$$|\Xi(\tau)| < \bar\rho.$$

$\square$

Now let μ be a regular ϵ-quasi-invariant measure and $\{S_\epsilon^\tau, \tau \geqslant 0\}$ be the flow of equation (2.1.5) on h^p. Below we follow the arguments, invented by Anosov for the finite dimensional averaging (e.g. see in [57]).

Consider the measure $\mu_1 = d\mu dt$ on $h^p \times \mathbb{R}$. Define the following subset of $h^p \times \mathbb{R}$:

$$\mathcal{B}^\epsilon := \{(v,\tau) : v \in \Gamma_0,\ \ \tau \in [0,1] \text{ and } S_\epsilon^\tau v \in \Gamma_1 \setminus E(\delta, T_0)\}.$$

Then by (2.1.8), there exists $C(M_1)$ such that

$$\mu_1(\mathcal{B}^\epsilon)=\int_0^1\mu\Big(\Gamma_0\cap S_\epsilon^{-\tau}\Big(\Gamma_1\setminus E(\delta,T_0)\Big)\Big)d\tau\leqslant e^{C(M_1)}\mu(\Gamma_1\setminus E(\delta,T_0)).$$

For any $v\in\Gamma_0$, denote $\bar\Delta(v)=\Delta(1,\epsilon,\delta,T_0,I,\varphi)$, where $(I,\varphi)=\Pi_{I,\varphi}(v)$. Then by the Fubini theorem, we have

$$\mu_1(\mathcal{B}^\epsilon)=\int_{\Gamma_0}Mes(\bar\Delta(v))d\mu(v).$$

Using Chebyshev's inequality, we obtain

$$\mu\Big(\Gamma_0\setminus N(\beta,\epsilon,\delta,T_0)\Big)\leqslant\frac{e^{C(M_1)}}{\beta}\mu\Big(\Gamma_1\setminus E(\delta,T_0)\Big).\tag{2.4.16}$$

By the definition of $E(\delta,T_0)$, we know that

$$E(\delta,T_0)\subset E(\delta,T_0'),\quad \text{if}\quad T_0'\geqslant T_0.\tag{2.4.17}$$

We set $E^\infty(\delta):=\cup_{T_0\geqslant1}E(\delta,T_0)$. Define

$$\mathcal{RES}(m_0)=\Big\{(I,\varphi)\in\Gamma_1:\ \exists k\in\mathbb{Z}^{m_0}\ \text{ such that }\ k_1W_1(I)+\cdots+k_{m_0}W_{m_0}(I)=0\Big\}.$$

Since the measure μ is regular, then by Lemma 2.2.3, we have that $\mu(\mathcal{RES}(m_0))=0$. If $(I',\varphi')\in\Gamma_1\setminus\mathcal{RES}(m_0)$, then the vector $W^{m_0}(I')\in\mathbb{R}^{m_0}$ is non-resonant. Due to Lemma 2.3.3, we know that there exists $T_0'>0$ such that for $T\geqslant T_0'$, the inequalities (2.4.8) hold. Therefore $(I',\varphi')\in E(\delta,T_0')\subset E^\infty(\delta)$. Hence

$$\Gamma_1\setminus E^\infty(\delta)\subset\mathcal{RES}.$$

So we have $\mu(E^\infty(\delta))=\mu(\Gamma_1)$. Since $\mu(E^\infty(\delta))=\lim_{T_0\to\infty}E(\delta,T_0)$ due to (2.4.17), then for any $\nu>0$, there exists $T_0'>0$ such that for each $T_0\geqslant T_0'$, we have

$$\mu(E^\infty\setminus E(\delta,T_0))\leqslant\nu.$$

So the r.h.s of the inequality (2.4.16) can be made arbitrary small if T_0 is large enough.

Fix some $0<\sigma<1/2$, we have proved the following lemma.

Lemma 2.4.3. *Fix any $\delta>0$, $\bar\rho>0$. Then for every $\nu>0$ we can find $\epsilon(\nu)>0$ such that, if $\epsilon<\epsilon(\nu)$, then*

$$\mu\Big(\Gamma_0\setminus N(\frac{\bar\rho}{9C_{M_1}})\Big)<\nu,$$

where $N(\frac{\bar\rho}{9C_{M_1}})=N(\frac{\bar\rho}{9C_{M_1}},\epsilon,\delta,\epsilon^{-\sigma})$.

We now are in a position to prove assertion (i) of Theorem 2.1.2.

By Corollary 2.3.2, for each $q \in [p', p]$, there exists $C_3(q, M_1)$ such that for any $J_1, J_2 \in B_{\bar{q}}^I(M_1+1)$ (see Agreements),

$$|\langle F\rangle(J_1) - \langle F\rangle(J_2)|_q^\sim \leqslant |\langle F\rangle(J_1) - \langle F\rangle(J_2)|_{q+\bar{\zeta}_0/2}^\sim \leqslant C_3(\bar{q}, M_1)|J_1 - J_2|_q^\sim. \quad (2.4.18)$$

Since the mapping $\langle F\rangle : h_I^p \to h_I^p$ is locally Lipschitz by (2.4.18), then using Picard's theorem, for any $J_0 \in B_p^I(M_1)$ there exists a unique solution $J(\cdot)$ of the averaged equation (2.1.9) with $J(0) = J_0$. We denote

$$\mathcal{T}(J_0) := \inf\{\tau > 0 : |J(\tau)|_p > M_1 + 1\} \leqslant \infty.$$

For any $\bar{\rho} > 0$ and $q < p + \zeta_0$, there exist n_1 such that

$$\begin{aligned} &|F(I,\varphi) - F^{n_1}(I,\varphi)|_q < \frac{\bar{\rho}}{8} e^{-C_3(M_1)}, \quad (I,\varphi) \in B_p^I(M_1+1) \times \mathbb{T}^\infty, \\ &|\langle F\rangle(J) - \langle F\rangle^{n_1}(J)|_q < \frac{\bar{\rho}}{8} e^{-C_3(M_1)}, \quad J \in B_p^I(M_1+1). \end{aligned} \quad (2.4.19)$$

Here

$$C_3(M_1) = \begin{cases} C_3(p, M_1) & \text{if} \quad q > p, \\ C_3(q, M_1) & \text{if} \quad q \leqslant p. \end{cases}$$

Find ρ_0 from the relation

$$8 \sum_{j=1}^{n_1} j^{1+2q} \rho_0 = \bar{\rho} e^{-C_3(M_1)}.$$

By Lemma 2.4.1 and Corollary 2.4.2, there exists $\epsilon_{\bar{\rho},q}$ such that if $\epsilon \leqslant \epsilon_{\bar{\rho},q}$, then for initial data in the subset $\Gamma_{\bar{\rho}} = N_{n_1,\rho_0}(\frac{\rho_0}{9C_{M_1}\epsilon}, \epsilon, \frac{\rho_0}{9}, \epsilon^{-\sigma})$ we have for $k = 1, \cdots, n_1$,

$$I_k^\epsilon(\tau) = I_k^\epsilon(0) + \int_0^\tau \langle F_k\rangle(I^\epsilon(s))ds + \xi_k(\tau), \quad |\xi_k(\tau)| < \rho_0, \quad \tau \in [0,1], \quad (2.4.20)$$

Therefore, by (2.4.19) and (2.4.20), for $(I^\epsilon(0), \varphi^\epsilon(0)) \in \Gamma_{\bar{\rho}}$, $J(0) \in B_p(M_1+1)$ and $|\tau| \leqslant \min\{1, \mathcal{T}(J(0))\}$,

$$\begin{aligned} &|I^\epsilon(\tau) - J(\tau)|_q^\sim - |I^\epsilon(0) - J(0)|_q^\sim \\ &\leqslant \int_0^\tau |F(I^\epsilon(s))ds - \int_0^\tau \langle F\rangle(J(s))|_q^\sim ds \\ &\leqslant \int_0^\tau |F^{n_1}(I^\epsilon(s)) - \langle F\rangle^{n_1}(J(s))|_q^\sim ds + \frac{\rho}{4} e^{-C_3(M_1)}. \\ &\leqslant \int_0^\tau |\langle F\rangle(I^\epsilon(s)) - \langle F\rangle(J(s))|_q^\sim ds + \frac{\rho}{2} e^{-C_3(M_1)}. \end{aligned}$$

Using (2.4.18), we obtain

$$|I^\epsilon(\tau) - J(\tau)|_q^\sim \leqslant |I^\epsilon(0) - J(0)|_q^\sim + \int_0^\tau C_3(M_1)|I^\epsilon(s) - J(s)|_q^\sim ds + \xi_0(\tau),$$

where $|\xi_0(\tau)| \leqslant \frac{\bar{\rho}}{2} e^{-C_3(M_1)}$. By Gronwall's lemma, if

$$|I^\epsilon(0) - J(0)|_q^\sim \leqslant \delta = e^{-C_3(M_1)}\bar{\rho},$$

then

$$|I(\tau) - J(\tau)|_q \leqslant 2\bar{\rho}, \quad |\tau| \leqslant \min\{1, \mathcal{T}(J(0))\}.$$

This establishes inequality (2.1.11). Assuming that $\bar{\rho} << 1$, we get from the definition of $\mathcal{T}(J(0))$ that $\mathcal{T}(J(0))$ is bigger than 1, if $\zeta_0 > 0$ and $q \geqslant p$. From Lemma 3.3 we know that for any $\nu > 0$, if ϵ small enough, then $\mu(\Gamma_0 \setminus \Gamma_{\bar{\rho}}) < \nu$. This completes the proof of the assertion (i) of Theorem 2.1.2.

Proof of statement (2) of Remark 2.1.3. If the perturbation is hamiltonian with Hamiltonian H, then $F = -\nabla_\varphi H$. Therefore the averaged vector filed $\langle F \rangle = 0$. For any $\rho > 0$ and any $q < p$, there exists n_2 such that

$$|I - I^{n_2}|_q^\sim < \rho/4, \quad \forall I \in B_p(M).$$

By similar argument, we can obtain that, there exists a subset $\Gamma^\epsilon_{\rho,n_2} \subset \Gamma_0$, satisfying (2.1.10), such that for initial data $(I^\epsilon(0), \varphi^\epsilon(0)) \in \Gamma^\epsilon_{\rho,n_2}$, and for $\tau \in [0,1]$, we have

$$|I^{\epsilon,n_2}(\tau) - I^{\epsilon,n_2}(0)|_q^\sim \leqslant \rho/4.$$

So

$$|I^\epsilon(\tau) - I^\epsilon(0)|_q^\sim \leqslant \rho \quad \text{for} \quad \tau \in [0,1].$$

In this argument we do not require $\zeta_0 \geqslant 0$. So item (ii) of Assumption A is not needed if the perturbation is hamiltonian.

2.4.2 Proof of the assertion (ii)

We fix $\alpha < 1/4$. For any $(m,n) \in \mathbb{N}^2$ denote

$$\mathcal{B}_m(\epsilon) := \Big\{(I,\varphi) \in \Gamma_1 : \inf_{k \leqslant m} |I_k| < \epsilon^\alpha\Big\},$$
$$\mathcal{R}_{m,n}(\epsilon) := \bigcup_{|L| \leqslant n, L \in \mathbb{Z}^m \setminus \{0\}} \Big\{(I,\varphi) \in \Gamma_1 : |W(I) \cdot L| < \epsilon^\alpha\}\Big\}.$$

Then let

$$\Upsilon_{m,n}(\epsilon) = \Big(\bigcup_{m_0 \leqslant m} \mathcal{R}_{m_0,n}(\epsilon)\Big) \cup \mathcal{B}_m(\epsilon), \tag{2.4.21}$$

and for any $(I_0, \varphi_0) \in \Gamma_0$, denote

$$S(\epsilon, m, n, I_0, \varphi_0) = \{\tau \in [0,1] : (I^\epsilon(\tau), \varphi^\epsilon(\tau)) \in \Upsilon_{m,n}(\epsilon)\}.$$

Fix $m \in \mathbb{N}$, take a bounded Lipschitz function g defined on the torus $\mathbb{T}^m$ such that $Lip(g) \leqslant 1$ and $|g|_{L_\infty} \leqslant 1$. Let $\sum_{s \in \mathbb{Z}^m} g_s e^{is\cdot\varphi}$ be its Fourier series. Then for any $\rho > 0$, there exists n, such that if we denote $\bar{g}_n = \sum_{|s| \leqslant n} g_s e^{is\cdot\varphi}$, then

$$\Big|g(\varphi) - \bar{g}_n(\varphi)\Big| < \frac{\rho}{2}, \quad \forall \varphi \in \mathbb{T}^m.$$

As the measure μ is regular and $\Upsilon_{m,n}(\epsilon_1) \subset \Upsilon_{m,n}(\epsilon_2)$ if $\epsilon_1 \leqslant \epsilon_2$, then

$$\mu(\Upsilon_{m,n}(\epsilon)) \to 0, \quad \epsilon \to 0.$$

Since the measure μ is ϵ-quasi-invariant, then following the same argument that proves Lemma 2.4.3, we have there exists subset $\Lambda_\rho^\epsilon \subset \Gamma_0$ such that for initial data $(I_0, \varphi_0) \in \Lambda_\rho^\epsilon$, we have $\mathrm{Mes}\Big(S(\epsilon, m, n, I_0, \varphi_0)\Big) \leqslant \rho/4$, and if ϵ is small enough, then $\mu(\Gamma_0 \setminus \Lambda_\rho^\epsilon) < \mu(\Gamma_0)\rho/4$. Due to Lemma 2.3.3, if $(I^\epsilon(\cdot), \varphi^\epsilon(\cdot))$ stays long enough time outside the subset $\Upsilon_{m,n}(\epsilon)$, then the time average of $\bar{g}(\varphi^{\epsilon,m}(\tau))$ can be well approximated by its space average. Following an similar argument of Lemma 2.4.1, we could obtain that for ϵ small enough, for initial data $(I_0, \varphi_0) \in \Lambda_\rho^\epsilon$, we have

$$\Big| \int_{\mathbb{T}^m} \bar{g}(\varphi) d\lambda_\epsilon^{I_0,\varphi_0} - \int_{\mathbb{T}^m} \bar{g} d\varphi \Big| = \Big| \int_0^1 \bar{g}\Big(\varphi^{\epsilon,m}(\tau)\Big) d\tau - \int_{\mathbb{T}^m} \bar{g}(\varphi) d\varphi \Big| < \rho/2. \tag{2.4.22}$$

So if ϵ is small enough, then

$$\begin{aligned}
&\Big| \int_{\mathbb{T}^m} g(\varphi)\lambda_\epsilon - \int_{\mathbb{T}^m} g(\varphi) d\varphi \Big| \\
&\leqslant \frac{1}{\mu(\Gamma_0)} \Big\{ \Big| \int_{(I_0,\varphi_0)\in\Lambda_\rho^\epsilon} \Big[\int_{\mathbb{T}^m} g(\varphi) d\lambda_\epsilon^{I_0,\varphi_0} - \int_{\mathbb{T}^m} g(\varphi) d\varphi \Big] d\mu(I_0,\varphi_0) \Big| \\
&+ \Big| \int_{(I_0,\varphi_0)\in\Gamma_0\setminus\Lambda_\rho^\epsilon} \Big[\int_{\mathbb{T}^m} g(\varphi) d\lambda_\epsilon^{I_0,\varphi_0} - \int_{\mathbb{T}^m} g(\varphi) d\varphi \Big] d\mu(I_0,\varphi_0) \Big| \Big\} \leqslant 2\rho.
\end{aligned}$$

That is ,

$$\Big| \int g(\varphi)\lambda_\epsilon - \int g(\varphi) d\varphi \Big| \longrightarrow 0 \quad \text{as} \quad \epsilon \to 0, \tag{2.4.23}$$

for any Lipschitz function as above. Hence, the probability measure λ_ϵ converges weakly to Haar measure $d\varphi$ (see [22]). This proves the assertion (ii).

2.4.3 Consistency conditions

Assume the ϵ-quasi-invariant measure μ is dependent on ϵ, i.e. $\mu = \mu_\epsilon$. Using again the Anosov arguments, we have for the measure μ_ϵ that

$$\mu_\epsilon\Big(N(\beta, \epsilon, \delta, T_0)\Big) \leqslant \frac{e^{C_\epsilon(M_1)}}{\beta} \mu_\epsilon\Big(\Gamma_1 \setminus E(\delta, T_0)\Big).$$

It is easy to see that assertion (i) of Theorem 2.1.2 holds, with μ replace by μ_ϵ, if following consistency conditions are satisfied :

1) For any $\delta > 0$, $\mu_\epsilon(\Gamma_1 \setminus E(\delta, \epsilon^{-\sigma}))$ go to zero with ϵ.

2) $C_\epsilon(M_1)$ is uniformly bounded with respect to ϵ

In subsection 2.4.2, we can see that for assertion (ii) of Theorem 2.1.2 to hold, one more condition should be added to the family $\{\mu_\epsilon\}_{\epsilon\in(0,1)}$. That is,

3) For any m, $n \in \mathbb{N}$, $\mu_\epsilon\Big(\Upsilon_{m,n}(\epsilon)\Big)$ (see (2.4.21)) goes to zero with ϵ.

2.5 On existence of ϵ-quasi-invariant measures

In this section we will provide two sufficient conditions to the existence of ϵ-quasi-invariant measures for perturbed KdV equation (2.1.1).

2.5.1 Quasi-invariant measures on the space H^p

We will first prove that if Assumption A holds with $\zeta_0 \geqslant 2$, then there exist ϵ-quasi-invariant measures for the perturbed KdV (2.1.1) on the space H^p, where p is an integer not less than 3. Through out this section, we suppose that $\zeta_0 = 2$, $3 \leqslant p \in \mathbb{N}$ and $p' = 0$. Our presentation closely follows Chapter 4 of the book [75].

Let η_p be the centered Gaussian measure on H^p with correlation operator ∂_x^{-2}. Since ∂_x^{-2} is an operator of trace type, then η_p is a well-defined probability measure on H^p.

As is known, for solutions of KdV, there are countably many conservation laws $\mathcal{J}_n(u)$, $n \geqslant 0$, where $\mathcal{J}_0(u) = \frac{1}{2}||u||_0^2$ and

$$\mathcal{J}_n(u) = \int_{\mathbb{T}} \Big\{\frac{1}{2}(\partial_x^n u)^2 + c_n u(\partial_x^{n-1}u)^2 + \mathcal{Q}_n(u, \dots, \partial_x^{n-2}u)\Big\}dx, \tag{2.5.1}$$

for $n \geqslant 1$, where c_n are real constants, and $\mathcal{Q}_n$ are polynomial in their arguments (see, p.209 in [38] for exact form of the conservation laws). By induction we get from these relations that

$$||u||_n^2 \leqslant 2\mathcal{J}_n + C(\mathcal{J}_{n-1}, \dots, \mathcal{J}_0), \quad n \geqslant 1, \tag{2.5.2}$$

where C vanishes with $u(\cdot)$.

Now set $J_p = \mathcal{J}_{p+1}(u) - \frac{1}{2}||u||_{p+1}^2$. From the form (2.5.1), we know that the functional J_p is bounded on bounded sets in H^p. We consider the measure μ_p defined by

$$\mu_p(\Omega) = \int_\Omega e^{-J_p(u)} d\eta_p(u), \tag{2.5.3}$$

for every Borel set $\Omega \subset H^p$. This measure is regular in the sense of Definition 2.1.1 and non-degenerated in the sense that its support contains the whole space H^p (see, e.g. Chapter 9 in [11]). Moreover, it is invariant for KdV [75].

The main result of this section is the following theorem :

Theorem 2.5.1. *The measure μ_p is ϵ-quasi-invariant for perturbed KdV equation (2.1.1) on the space H^p.*

To prove this theorem, we follow a classical procedure based on finite dimensional approximation (see, e.g. [75]).

Let us firstly write equation (2.1.1) using the slow time $\tau = \epsilon t$,

$$\dot{u} = \epsilon^{-1}(-u_{xxx} + 6uu_x) + f(u), \tag{2.5.4}$$

where $\dot{u} = \frac{du}{d\tau}$. By Assumption A, for each $u_0 \in B_p^u(M)$, the equation (2.5.4) has a unique solution $u(\cdot) \in C([0,T], H^p)$ and $||u(\tau)||_p \leqslant C(||u_0||_p, T)$ for all $\tau \in [0,T]$.

Denote $\mathbb{L}_m$ the subspace of H^p, spanned by the basis vectors $\{e_1, e_{-1}, \dots, e_m, e_{-m}\}$. Let $\mathbb{P}_m$ be the orthogonal projection of H^p onto $\mathbb{L}_m$ and $\mathbb{P}_m^\perp = Id - \mathbb{P}_m$. For any $u \in H^p$, denote $u^m = \mathbb{P}_m u \in \mathbb{L}_m$. We will identify $\mathbb{P}_\infty$ with Id and u^∞ with u. Consider the problem

$$\dot{u}^m = \epsilon^{-1}\Big[- u^m_{xxx} + 6\mathbb{P}_m(u^m u^m_x)\Big] + \mathbb{P}_m(f(u^m)), \quad u^m(x,0) = \mathbb{P}_m u_0(x). \tag{2.5.5}$$

Clearly, for each $u_0 \in H^p$ this problem has a unique solution $u^m(\cdot) \in C([0,T'], \mathbb{L}_m)$ for some $T' > 0$.

Proposition 2.5.2. *Let $u_0 \in H^p$ and $u_0^m \in \mathbb{L}_m$ such that u_0^m strongly converge to u_0 in H^p as $m \to +\infty$. Then as $m \to +\infty$,*

$$u^m(\cdot) \to u(\cdot) \quad in \quad C([0,T], H^p),$$

where $u(\cdot)$ is the solution of equation (2.1.1) with initial datum $u(0) = u_0$ and $u^m(\cdot)$ is the solution of problem (2.5.5) with initial condition $u^m(0) = u_0^m$.

In this result, as well as in the Lemmas 2.5.5-2.5.7 below, the rate of convergence depends on the small parameter ϵ.

We shall prepare several lemmas to prove this proposition. For any $n, m \in \mathbb{N}$, we have for the solution $u^m(\tau)$ of problem(2.5.5)

$$\begin{aligned} \frac{d}{d\tau}\mathcal{J}_n(u^m(\tau)) &= \Big\langle \nabla_u \mathcal{J}_n(u^m), \dot{u}^m(\tau) \Big\rangle \\ &= \Big\langle \nabla_u \mathcal{J}_n(u^m), \epsilon^{-1}[-u^m_{xxx} + \mathbb{P}_m(u^m u^m_x)] + \mathbb{P}_m[f(u^m)] \Big\rangle \end{aligned}$$

Here ∇_u stands for the L_2-gradient with respect to u. Since $\mathcal{J}_n$ is a conservation law of KdV, then $\langle \nabla_u \mathcal{J}_n(u^m), -u^m_{xxx} + u^m u^m_x \rangle = 0$. So

$$\frac{d}{d\tau}\mathcal{J}_n(u^m) = -\epsilon^{-1}\Big\langle \nabla_u \mathcal{J}_n(u^m), \mathbb{P}_m^{\perp}(u^m u^m_x) \Big\rangle + \Big\langle \nabla_u \mathcal{J}_n(u^m), \mathbb{P}_m[f(u^m)] \Big\rangle. \tag{2.5.6}$$

We denote the first term in the right hand side by $\epsilon^{-1}\mathcal{E}_n(u^m)$ and the second term by $\mathcal{E}_n^f(u^m)$.

Lemma 2.5.3. *There exist continuous functions $\gamma_n(R,s)$ and $\gamma_n'(R,s)$ on $\mathbb{R}_+^2 = \{(R,s)\}$ such that they are non-decreasing in the second variable s, vanish if $s = 0$, and*

$$|\mathcal{E}_n^f(u^m)| \leqslant \gamma_n'(||u^m||_{n-1}, ||u^m||_{n-1}), \tag{2.5.7}$$

$$\begin{aligned} |\mathcal{E}_n(u^m)| \leqslant \gamma_n\Big(&||u^m||_{n-1}, \\ &\max_{\substack{0 \leqslant i,j \leqslant n-1, \\ i+j \neq 2n-2}} ||\mathbb{P}_m^{\perp}[\partial_x^i u^m \partial_x^j u^m]||_0 + ||\mathbb{P}_m^{\perp}(u^m u^m_x)||_1\Big). \end{aligned} \tag{2.5.8}$$

for all $n = 3, 4, \ldots$. For $n = 2$ equality (2.5.7) still holds, and

$$|\mathcal{E}_2(u^m)| \leqslant C_2(||u^m||_1)||u^m||_2^2 + C_2'(||u^m||_1). \tag{2.5.9}$$

Démonstration. Since $f(u)$ is 2-smoothing, from (2.5.1) and (2.5.6) we know that

$$|\mathcal{E}_n^f(u^m)| \leqslant \gamma_n'(||u^m||_{n-1}, ||u^m||_{n-1}),$$

where $\gamma_n'(\cdot,\cdot)$ is a continuous function satisfying the requirement in the statement of the lemma.

For the quantity $\mathcal{E}_n(u^m)$, by (2.5.1) and (2.5.6) we have

$$\begin{aligned}
\mathcal{E}_n(u^m) = \int_{\mathbb{T}} \Big\{ & 6(-1)^n(\partial_x^{2n}u^m)\mathbb{P}_m^{\perp}(u^m u_x^m) + 6c_n\mathbb{P}_m^{\perp}(u^m u_x^m)(\partial_x^{n-1}u^m)^2 \\
& + (-1)^{n-1}12c_n\partial_x^{n-1}(u^m\partial_x^{n-1}u^m)\mathbb{P}_m^{\perp}(u^m u_x^m) \\
& + 6\sum_{i=0}^{n-2}\frac{\partial\mathcal{Q}_n(u^m,\dots,\partial_x^{n-2}u^m)}{\partial(\partial_x^i u^m)}\partial_x^i\mathbb{P}_m^{\perp}(u^m u_x^m)\Big\}dx \\
= 0 + \int_{\mathbb{T}} \Big\{ & 6c_n\mathbb{P}_m^{\perp}(u^m u_x^m)(\partial_x^{n-1}u^m)^2 \\
& + 12c_n\mathbb{P}_m^{\perp}(u^m\partial_x^{n-1}u^m)[\partial_x\sum_{i=0}^{n-3}C_{n-2}^i\partial_x^{n-2-i}u^m\partial_x^{i+1}u^m] \\
& + 6\sum_{i=0}^{n-2}\frac{\partial\mathcal{Q}_n(u^m,\dots,\partial_x^{n-2}u^m)}{\partial(\partial_x^i u^m)}\partial_x^i\mathbb{P}_m^{\perp}(u^m u_x^m)\Big\}dx.
\end{aligned}$$

Hence we prove the assertion of the lemma. □

Lemma 2.5.4. *For every $u_0 \in H^p$, there exist $\tau_1(||u_0||_0) > 0$ and a continuous non-decreasing ϵ-depending function $\beta_p^\epsilon(s)$ on $[0, +\infty)$ such that the value $||u^m(\tau)||_p$ are bounded by the quantity $\beta_n^\epsilon(||u_0||_p)$, uniformly in $m = 1, 2, \dots$ and $\tau \in [0, \tau_1]$.*

Démonstration. Let $M = \max\{||u_0||_0, 1\}$. It is easy to verify that

$$\frac{d}{d\tau}||u^m||_0^2 = 2\langle u^m, \mathbb{P}_m(f(u^m))\rangle \leqslant 2||u^m||_0^2 + C(2M),$$

if $||u^m||_0 \leqslant 2M$. Therefore for a suitable $\tau_1 = \tau_1(||u_0||_0) > 0$ and all $\tau \in [0, \tau_1]$, we have $||u^m(\tau)||_0 \leqslant 2M$. For the quantity $\mathcal{J}_1(u^m)$ and $\tau \in [0, \tau_1]$,

$$\frac{d}{d\tau}\mathcal{J}_1(u^m) = \langle \nabla_u\mathcal{J}_1(u^m), \mathbb{P}_m f(u^m)\rangle \leqslant C_1(2M).$$

Therefore $\mathcal{J}_1(u^m(\tau)) \leqslant C_1\tau + \mathcal{J}_1(u^m(0))$. So $||u^m(\tau)||_1 \leqslant \beta_1(||u_0||_1)$. Similarly, by Lemma 2.5.3 and inequality (2.5.2), we have for $\tau \in [0, \tau_1]$,

$$\frac{d}{d\tau}\mathcal{J}_2(u^m(\tau)) \leqslant \epsilon^{-1}C_2[\beta_1(||u_0||_1)]\mathcal{J}(u^m(\tau)) + C_2''[\epsilon^{-1}, \beta_1(||u_0||_1)].$$

By Gronwall's lemma and relation (2.5.2), we obtain $||u^m(\tau)||_2 \leqslant \beta_2^\epsilon(||u_0||_2)$. In the view of Lemma 2.5.3, we have

$$\mathcal{J}_n(u^m(\tau)) \leqslant \mathcal{J}_n(u^m(0)) + \tau C_n[\epsilon^{-1}, \beta_{n-1}^\epsilon(||u_0||_{n-1})],$$

for $n = 3, \dots, p$. Hence $\max_{\tau\in[0,\tau_1]}||u^m(\tau)||_p \leqslant \beta_p^\epsilon(||u_0||_p)$. □

Below, we will denote by τ_1 the quantity $\min\{\tau_1(||u_0||_0), T\}$.

Lemma 2.5.5. *As $m \to \infty$,*

$$||u^m(\tau) - u(\tau)||_{p-1} \to 0,$$

uniformly in $\tau \in [0, \tau_1]$.

Démonstration. Denote $w = u^m - u$. Using that $\langle \partial_x^j u^m, \mathbb{P}_m^\perp u' \rangle = 0$ for any j and each $u' \in H^0$, we get :

$$\begin{aligned}
&\frac{1}{2}\frac{d}{d\tau}||w||_{p-1}^2 \\
&= \Big\langle \partial_x^{p-1} w,\ \partial_x^{p-1}\Big[\epsilon^{-1}\Big(-w_{xxx} + 6\mathbb{P}_m(u^m u_x^m) - 6uu_x\Big) + \mathbb{P}_m(f(u^m)) - f(u)\Big]\Big\rangle \\
&= 3\epsilon^{-1}\Big\langle \partial_x^{p-1} w,\ \partial_x^p\Big[(u^m)^2 - u^2\Big]\Big\rangle + 3\epsilon^{-1}\Big\langle \mathbb{P}_m^\perp(\partial_x^{p-1}u), \partial_x^p[(u^m)^2]\Big\rangle \\
&\quad + \Big\langle \partial_x^{p-1} w,\ \partial_x^{p-1}\Big[\mathbb{P}_m(f(u^m)) - f(u)\Big]\Big\rangle.
\end{aligned}$$

Using Sobolev embedding and integration by part, we have

$$\begin{aligned}
\Big\langle \partial_x^{p-1} w, \partial_x^p\Big[(u^m)^2 - u^2\Big]\Big\rangle &= \sum_{i=0}^p C_p^i \int_{\mathbb{T}} \partial_x^{p-1} w \partial_x^{p-i} w \partial_x^i (u^m + u) dx \\
&\leqslant -\int_{\mathbb{T}} \partial_x(u^m + u)(\partial_x^{p-1} w)^2 dx + \sum_{i=1}^p C_p^i ||w||_{p-1}^2 ||u^m + u||_p \\
&\leqslant C(||u||_p, ||u^m||_p)||w||_{p-1}^2
\end{aligned}$$

Therefore,

$$\frac{d}{d\tau}||w||_{p-1}^2 \leqslant C_1(\epsilon^{-1}, ||u^m||_n)||\mathbb{P}_m^\perp u||_p + ||\mathbb{P}_m^\perp f(u)||_p + C_2(\epsilon^{-1}, ||u||_n, ||u^m||_n)||w||_{p-1}^2.$$

Since $||\mathbb{P}_m^\perp(u)||_p$ and $||\mathbb{P}_m^\perp(f(u))||_p$ go to zero as $m \to \infty$ for each $\tau \in [0, \tau_1]$ and as they are uniformly bounded on $[0, \tau_1]$ by Lemma 2.5.4, we have for $\tau \in [0, \tau_1]$,

$$||w||_{p-1}^2(\tau) = ||w(0)||_{p-1}^2 + \int_0^\tau C(\epsilon^{-1}, ||u_0||_p)||w||^2 ds + a_m(\epsilon^{-1}, \tau),$$

where $a_m(\epsilon^{-1}, \tau) \to 0$ as $m \to \infty$. So the assertion of the lemma follows form Gronwall's lemma. □

Lemma 2.5.6. *Let $\tau^m \in [0, \tau_1]$ such that $\tau^m \to \tau^0 \in [0, \tau_1]$, then*

$$||u^m(\tau^m) - u(\tau_0)||_p \to 0 \quad \textit{as} \quad m \to \infty.$$

Démonstration. We firstly prove $\mathcal{J}_p(u^m(\tau^m)) \to \mathcal{J}_p(u(\tau_0)$ as $m \to \infty$. Indeed, for $m \leqslant +\infty$, by (2.5.6), we have

$$\mathcal{J}_p(u^m(\tau^m)) = \mathcal{J}_p(u^m(0)) + \int_0^{\tau^m} [\epsilon^{-1}\mathcal{E}_p(u^m(s)) + \mathcal{E}_p^f(u^m(s))]ds$$

Since $f(u)$ is 2-smoothing, the second term in the integrand is continuous in H^{p-1}. So, in the view of Lemma 2.5.5, we only need to prove that the first term goes to zero as $m \to \infty$. Due to Lemma 2.5.3, we only need to show that uniformly in $\tau \in [0, \tau_1]$,

$$\lim_{m\to\infty} ||\mathbb{P}_m^\perp(\partial_x^i u^m(\tau)\partial_x^j u^m(\tau)||_0 + \lim_{m\to\infty} ||\mathbb{P}_m^\perp u^m(\tau) u^m(\tau)_x||_1 = 0,$$

where $0 \leqslant i, j \leqslant p-1$ and $i+j \neq 2p-2$. For the first term in the left hand side, we have

$$\begin{aligned}&||\mathbb{P}_m^{\perp}(\partial_x^i u^m(\tau)\partial_x^j u^m(\tau))||_0 \\ &\leqslant ||\mathbb{P}_m^{\perp}(\partial_x^i u^m(\tau)\partial_x^j u^m(\tau) - \partial_x^i u(\tau)\partial_x^j u(\tau))||_0 + ||\mathbb{P}_m^{\perp}(\partial_x^i u(\tau)\partial_x^j u(\tau))||_0.\end{aligned} \tag{2.5.10}$$

By Lemma 2.5.5, the first term in the r.h.s of (2.5.10) goes to zero as $m \to \infty$, uniformly in $\tau \in [0, \tau_1]$. Since $u(\cdot) \in C([0,\tau_1], H^p)$, then $\partial_x^i u(\cdot)\partial_x^j u(\cdot) \in C([0,\tau_1], H^0)$. Therefore, the quantity $||\mathbb{P}_m^{\perp}(\partial_x^i u(\tau)\partial_x^j u(\tau)||_0 \to 0$ as $m \to \infty$, uniformly in $\tau \in [0, \tau_1]$.

In the same way $\lim_{m\to\infty} ||\mathbb{P}_m^{\perp} u^m u_x^m||_1 = 0$. Therefore, we have

$$\lim_{m\to\infty} \mathcal{J}_p(u^m(\tau^m)) = \mathcal{J}_p(u(0)) + \int_0^{\tau^0} \langle \nabla_u \mathcal{J}_p(u(s)), f(u(s))\rangle ds = \mathcal{J}_p(u(\tau^0)).$$

Since the quantity $\mathcal{J}_p(u) - ||u||_p^2/2$ is continuous in H^{p-1}, we have

$$\lim_{m\to\infty} ||u^m(\tau^m)||_p = ||u(\tau_0)||_p.$$

The assertion of the lemma follows from the fact that weak convergence plus norm convergence imply strong convergence. □

Lemma 2.5.7. *As $m \to \infty$, $||u^m(\tau) - u(\tau)||_p \to 0$ uniformly for $\tau \in [0, \tau_1]$.*

Démonstration. Assume the contrary holds. Then there exists $\delta > 0$ such that for each $m \in \mathbb{N}$, there exists $\tau^m \in [0, \tau_1]$ satisfying

$$||u^m(\tau^m) - u(\tau^m)||_p \geqslant \delta. \tag{2.5.11}$$

Take a subsequence $\{m_k\}$ such that $\tau^{m_k} \to \tau^0 \in [0, \tau_1]$ as $m_k \to \infty$. By Lemma 4.6, we have

$$\begin{aligned}&\lim_{m_k\to\infty} ||u^{m_k}(\tau^{m_k}) - u(\tau^{m_k})||_p \\ =&\lim_{m_k\to\infty} (||u^{m_k}(\tau^{m_k}) - u(\tau^0)||_p + ||u(\tau^{m_k}) - u(\tau^0)||_p) = 0.\end{aligned}$$

This contradicts with inequality (2.5.11). So the assertion of the Lemma holds. □

If $T = \tau_1$, Proposition 2.5.2 is proved. Otherwise, we just need to iterate above procedure by letting the initial datum be $u(\tau_1)$. This finishes the proof of Proposition 2.5.2.

With Proposition 2.5.2, we will need two more results to prove Theorem 2.5.1.

Proposition 2.5.8. *For each $u_0 \in H^p$ and any $\nu > 0$, there exists $\delta > 0$ such that*

$$||u^m(\tau) - u_1^m(\tau)||_p < \nu,$$

uniformly for all $m = 1, 2, \dots, \tau \in [0, T]$ and every solution $u_1^m(\cdot)$ of problem (2.5.5) with initial data $u_1^m(0)$ satisfying

$$||u^m(0) - u_1^m(0)||_p < \delta,$$

(here $u^m(\cdot)$ is the solution of (2.5.5) with initial data $\mathbb{P}_m u_0$).

Démonstration. Assume the contrary. Then there exists $\nu > 0$ such that for each $\delta > 0$, there exists $m \in \mathbb{N}$, $u_1 \in \mathbb{L}_m$ and $\tau^m \in [0, T]$ satisfying

$$||u_1^m(\tau^m) - u^m(\tau^m)||_p \geqslant \nu \quad \text{and} \quad ||u_1^m(0) - u^m(0)||_p < \delta. \tag{2.5.12}$$

Hence there exists subsequence m_k such that $||u_1^{m_k} - \mathbb{P}_{m_k} u_0||_p \to 0$ as $m_k \to \infty$. Therefore $\lim_{m_k \to \infty} ||u_1^{m_k} - u_0||_p = 0$. By Proposition 2.5.2, we known that

$$|u_1^{m_k}(\tau^{m_k}) - u^{m_k}(\tau^{m_k})||_p \leqslant ||u_1^{m_k}(\tau^{m_k}) - u(\tau^{m_k})||_p + ||u^{m_k}(\tau^{m_k}) - u(\tau^{m_k})||_p \to 0,$$

as $m_k \to \infty$. This contradicts with the first inequality of (2.5.12). Proposition 2.5.8 is proved. $\square$

Lemma 2.5.9. *Let $u_0 \in H^p$. Then for any $\delta > 0$, there exist $r > 0$ and $m_0 > 0$ such that for each $m \geqslant m_0$ and $\bar{u}(0) \in \dot{B}_p^u(u_0, r)$, the quantity*

$$\epsilon^{-1} |\mathcal{E}_{p+1}(\bar{u}^m(\tau))| \leqslant \delta,$$

for all $\tau \in [0, T]$.

Démonstration. In the view of Lemma 2.5.3, we only need to show for each $\delta_\epsilon > 0$, there exist $r > 0$ and $m_0 > 0$ such that for every $\bar{u}_0 \in \dot{B}_p^u(u_0, r)$, and $m \geqslant m_0$, we have for $\tau \in [0, T]$,

$$\max_{0 \leqslant i,j \leqslant p, i+j \neq 2p} ||\mathbb{P}_m^{\perp}[\partial_x^i \bar{u}^m(\tau) \partial_x^j \bar{u}^m(\tau)]||_0 + ||\mathbb{P}_m^{\perp}(\bar{u}^m(\tau)\bar{u}_x^m(\tau))||_1 < \delta_\epsilon. \tag{2.5.13}$$

Here $\bar{u}^m(\tau)$ is the solution of problem (2.5.5) with initial datum $\bar{u}^m(0) = \mathbb{P}_m \bar{u}_0$.

For the first term, we have

$$\begin{aligned}
&||\mathbb{P}_m^{\perp}[\partial_x^i \bar{u}^m(\tau) \partial_x^j \bar{u}^m(\tau)]||_0 \\
&\leqslant ||\partial_x^i u^m(\tau) \partial_x^j u^m(\tau) - \partial_x^i \bar{u}^m(\tau) \partial_x^j \bar{u}^m(\tau)||_0 \\
&\quad + ||\partial_x^i u^m(\tau) \partial_x^j u^m(\tau) - \partial_x^i u(\tau) \partial_x^j u(\tau)||_0 + ||\mathbb{P}_m^{\perp}[\partial_x^i u(\tau) \partial_x^j u(\tau)]||_0.
\end{aligned}$$

By Proposition 2.5.2 and the fact that $\partial_x^i u(\cdot) \partial_x^j u(\cdot) \in C([0, T], H^0)$, the second and the third terms on the right hand side of this inequality converge to zero as $m \to \infty$, uniformly in $\tau \in [0, T]$. From Proposition 2.5.8, we know that there exists $r > 0$ such that the first term is smaller than $\delta_\epsilon/2$ for all $\bar{u} \in \dot{B}_r^p(u_0)$ and uniformly in all $m \in \mathbb{N}$ and $\tau \in [0, T]$. Estimating in this way the term $||\mathbb{P}_m^{\perp}(\bar{u}^m \bar{u}_x^m)||_1$, we obtain inequality (2.5.13). Hence we prove the assertion of the lemma. $\square$

We now begin to prove Theorem 4.1.

Consider the following Gaussian measure η_p^m on the subspaces $\mathbb{L}_m \subset H^p$:

$$\begin{aligned}
d\eta_p^m := &\prod_{i=1}^m (2\pi)^{2p} i^{2p+1} \exp - \frac{(2\pi i)^{2p+2}(\hat{u}_i^2 + \hat{u}_{-i}^2)}{2} d\hat{u}_i d\hat{u}_{-i} \\
&= c(m) \exp \frac{-||u^m||_{p+1}^2}{2} d\hat{u}_1 d\hat{u}_{-1} \dots d\hat{u}_m d\hat{u}_{-m},
\end{aligned}$$

where $u^m := \sum_{i=1}^m(\hat{u}_ie_i + \hat{u}_{-i}e_{-i}) \in \mathbb{L}_m$ and $d\hat{u}_{\pm i}$, $i \in \mathbb{N}$, is the Lebesgue measure on $\mathbb{R}$. Obviously, η_p^m is a Borel measure on $\mathbb{L}_m$. Then we have obtained a sequence of Borel measure $\{\eta_p^m\}$ on H^p (see, e.g. [75]). We set

$$\mu_p^m(\Omega) = \int_\Omega e^{-J_p(u)}d\eta_p^m,$$

for every Borel set $\Omega \in H^p$. Then μ_p^m are well defined Borel measure on H^p. Clearly

$$d\mu_p^m = c(m)e^{-\mathcal{J}_{p+1}(u^m)}d\hat{u}_1d\hat{u}_{-1}\dots d\hat{u}_md\hat{u}_{-m}.$$

Lemma 2.5.10. *([75]) The sequence of Borel measures μ_p^m in H^p converges weakly to the measure μ_p as $m \to \infty$.*

Rewrite the system (2.5.5) in the variables

$$\hat{u}^m = (\hat{u}_1, \hat{u}_{-1}, \dots, \hat{u}_m, \hat{u}_{-m}),$$

where $u^m = \sum_j^m(\hat{u}_je_j + \hat{u}_{-j}e_{-j})$:

$$\frac{d}{d\tau}\hat{u}_j = -2\pi j\epsilon^{-1}\frac{\partial\mathcal{J}_1(\hat{u}^m)}{\partial\hat{u}_{-j}} + f_j(\hat{u}^m), \quad j = \pm 1, \dots, \pm m. \tag{2.5.14}$$

where $\mathbb{P}_mf(\hat{u}^m) = \sum_{j=1}^m(f_j(\hat{u}^m)e_j + f_{-j}(\hat{u}^m)e_{-j})$. Let S_m^τ, $\tau \in [0,T]$, be the flow map of equation (2.5.14). For any Borel set $\Omega \subset H^p$, let $S_m^\tau(\Omega) = S_m^\tau(\mathbb{P}_m(\Omega))$. By the Liouville Theorem and (2.5.6), we have

$$\frac{d}{d\tau}\mu^m(S_m^\tau(\Omega)) = \int_{S_m^\tau(\Omega)}\Big[\epsilon^{-1}\mathcal{E}_{p+1}(u^m) + \mathcal{E}_{p+1}^f(u^m) + \sum_{i=-m,i\neq 0}^m\frac{\partial f_i}{\partial\hat{u}_i})\Big]d\mu^m. \tag{2.5.15}$$

Denote S_ϵ^τ, $\tau \in [0,T]$, to be the flow map of equation (2.5.4) on the space H^p. Fix any $M > 0$, by Assumption A, there exists M_1 such that

$$S_\epsilon^\tau(B_p^u(M)) \subset B_p^u(M_1).$$

Since $f(u)$ is 2-smoothing, then by Cauchy inequality, $|\partial f_i/\partial\hat{u}_i| = \mathcal{O}(i^{-2})$. So we have

$$|\mathcal{E}_{p+1}^f(u^m) + \sum_{i=-m,i\neq 0}^m\frac{\partial f_i}{\partial\hat{u}_i}(u^m)| \leqslant C(M_1), \ \ \forall m \in \mathbb{N} \ \ \text{and} \ \ \forall u^m \in B_p^u(M_1). \tag{2.5.16}$$

Now fix $\tau_0 \in [0,T]$. Take an open set $\Omega \subset B_p^u(M)$. For any $\delta > 0$, there exists compact set $K \subset \Omega$ such that $\mu_p(\Omega \setminus K) < \delta$. Let $K_1 = S_\epsilon^{\tau_0}(K)$. Then the set K_1 also is compact and $K_1 \subset S_\epsilon^{\tau_0}(\Omega) = \Omega_1$. Define

$$\alpha = \min\{dist(K,\partial\Omega); dist(K_1,\partial\Omega_1)\},$$

where $dist(A,B) = \inf_{u\in A,v\in B}||u-v||_p$ and ∂A is the boundary of the set $A \subset H^p$.

Clearly $\alpha > 0$. By Proposition 4.8 and Lemma 4.9, for each $u_0 \in K$, there exists a $m_{u_0} > 0$ and an open ball $\dot{B}_p^u(u_0, r_{u_0})$ of radius $r_{u_0} > 0$ such that

$$||u^m(s) - \bar{u}^m(s)||_p \leqslant \alpha/3 \quad \text{and} \quad |\epsilon^{-1}\mathcal{E}_{p+1}(\bar{u}^m)| \leqslant C(M_1)/2, \tag{2.5.17}$$

for all $\bar{u} \in B_p^u(u_0, r_{u_0})$, $m \geqslant m_0$ and $s \in [0, \tau_0]$. Let $B_1, \ldots, B_l$ be the finite covering of the compact set K by such balls. Let

$$D = \cup_{i=1}^{l} B_i \quad \text{and} \quad \Omega_{\alpha/3} := \{u \in \Omega_1 | \ dist(u, \partial\Omega_1) \geqslant \alpha/3\}.$$

By Proposition 4.2,

$$S_m^{\tau_0}(D) \subset \Omega_{\alpha/3},$$

for all large enough $m \in \mathbb{N}$. From inequalities (2.5.15), (2.5.16) and (2.5.17), we know that if m is sufficiently large, then

$$e^{-3C(M_1)\tau_0/2}\mu_p^m(D) \leqslant \mu_p^m(S^{\tau_0}(D)) \leqslant e^{3C(M_1)\tau_0/2}\mu_p^m(D).$$

By Lemma 4.10, we have

$$\begin{aligned}
\mu_p(\Omega) &\leqslant \mu_p(D) + \delta \leqslant \liminf_{m\to\infty} \mu_p^m(D) + \delta \\
&\leqslant \liminf_{m\to\infty} e^{3C(M_1)\tau_0/2}\mu_p^m(S_m^{\tau_0}(D)) + \delta \leqslant \limsup_{m\to\infty} e^{3C(M_1)\tau_0/2}\mu_p^m(\Omega_{\alpha/3}) + \delta \\
&\leqslant e^{3C(M_1)\tau_0/2}\mu_p(\Omega_1) + \delta.
\end{aligned}$$

Here we have used the Portemanteau theorem. Since δ was chosen arbitrarily, it follows that

$$\mu_p(\Omega) \leqslant e^{3C(M_1)\tau_0/2}\mu_p(S_\epsilon^{\tau_0}(\Omega)).$$

Similarly, $\mu_p(S^{\tau_0}(\Omega)) \leqslant e^{3C(M_1)\tau_0/2}\mu_p(\Omega)$. As $\tau_0 \in [0, T]$ is fixed arbitrarily, Theorem 4.1 is proved.

Remark 2.5.11. *The measure μ_p is also ϵ-quasi-invariant for the following perturbed KdV equations on H^p :*

$$\dot{u} + \epsilon^{-1}(u_{xxx} - 6uu_x) = \partial_x u, \tag{2.5.18}$$
$$\dot{u} + \epsilon^{-1}(u_{xxx} - 6uu_x) = \partial_x^{-1} u. \tag{2.5.19}$$

Indeed, consider the following finite dimensional system corresponding to equation (2.5.18) as in problem (2.5.5) :

$$\dot{u}^m = \epsilon^{-1}\Big[-u_{xxx}^m + 6\mathbb{P}_m(u^m u_x^m)] + \partial_x u^m, \quad u^m(0) = \mathbb{P}_m u_0. \tag{2.5.20}$$

Let us investigate the quantity $\frac{d}{d\tau}\mathcal{J}_n(u^m)$, $n \geqslant 3$, for equation (2.5.20) :

$$\frac{d}{d\tau}\mathcal{J}_n(u^m) = \epsilon^{-1}\mathcal{E}_n(u^m) + \langle \nabla_u \mathcal{J}_n(u^m), \partial_x u^m\rangle.$$

For the first term, see in Lemma 2.5.3. For the second term,

$$\begin{aligned}
D_n := \langle \nabla_u \mathcal{J}_n(u^m), \partial_x u^m\rangle = \int_{\mathbb{T}} \Big\{ & \partial_x^n u^m \partial_x^{n+1} u^m + c_n \partial_x u^m (\partial_x^{n-1} u^m)^2 \\
& + 2c_n u^m \partial^{n-1} u^m \partial^n u^m \\
& + \sum_{i=0}^{n-2} \frac{\partial \mathcal{Q}_n(u^m, \ldots, \partial_x^{n-2} u^m)}{\partial(\partial_x^i u^m)} \partial^{i+1} u^m \Big\} dx.
\end{aligned}$$

Notice that the first term in right hand side vanishes. For the second and the third terms,

$$\int_{\mathbb{T}} c_n[\partial_x u^m(\partial_x^{n-1}u^m)^2 + 2u^m\partial_x^{n-1}u^m\partial_x^n u^m]dx = c_n\int_{\mathbb{T}} d[u^m(\partial_x^{n-1}u^m)^2] = 0.$$

So we have

$$|D_n| \leqslant C(||u^m||_{n-1}). \tag{2.5.21}$$

Note that equation (2.5.20) can be written as a Hamiltonian system in coordinates $\hat{u}^m = (\hat{u}_1, \hat{u}_{-1}, \dots, \hat{u}_m, \hat{u}_{-m})$:

$$\frac{d}{d\tau}\hat{u}_j = -2\pi j\epsilon^{-1}\frac{\partial H_1(\hat{u}^m)}{\partial \hat{u}_{-j}}, \quad j = \pm 1, \dots, \pm m, \tag{2.5.22}$$

where the Hamiltonian $H_1(u) = \mathcal{J}_1(u) - \frac{\epsilon}{2}\int_{\mathbb{T}} u^2 dx$. Therefore the divergence for the vector field of equation (2.5.22) is zero. This property and inequality (2.5.21) also hold for equation (2.5.19). Hence the same proof in this section applies to equation (2.5.18) and (2.5.19), which justifies the claim in the Remark 2.5.11.

2.5.2 The ϵ-quasi-invariant measure on the space h^p

Fix $\zeta_0' > 1$ and $p \geqslant 3$, and let μ be a ζ_0'-admissible Gaussian measure on the Hilbert space h^p (see 2.1.12). In this subsection we will discuss how this measure evolves under the flow of the perturbed KdV equation (2.1.1). We follow a classical procedure based on finite dimensional approximations (see e.g. [75]).

We suppose the assumption A holds. Let us write the equation (2.1.1) in the Birkhoff normal form, using the slow time $\tau = \epsilon t$:

$$\frac{d}{d\tau}\mathbf{v}_j = \epsilon^{-1}\mathcal{J}W_j(I)\mathbf{v}_j + \mathbf{X}_j(v), \quad j \in \mathbb{N}, \tag{2.5.23}$$

where $\mathbf{X}_j = (X_j,\ X_{-j})^t \in \mathbb{R}^2$ and $\mathcal{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Let $X(v) = (\mathbf{X}_1(v), \dots)$, then

$$X(v) = d\Psi\Big(\Psi^{-1}(v)\Big)\Big(f\Big(\Psi^{-1}(v)\Big)\Big).$$

We assume additionally that :

The mapping defined by $h^p \to h^{p+\zeta_0'} : v \mapsto X(v)$ is analytic.

For any $n \in \mathbb{N}$, we consider the $2n$-dimensional subspace $\pi_n(h^p)$ of h^p with coordinates $v^n = (\mathbf{v}_1, \dots, \mathbf{v}_n, 0, \dots)$. On $\pi_n(h^p)$, we define the following finite-dimensional systems :

$$\frac{d}{d\tau}\vec{\omega}_j = \epsilon^{-1}\mathcal{J}W_j(I(\omega^n))\vec{\omega}_j + \mathbf{X}_j(\omega^n), \quad 1 \leqslant j \leqslant n, \tag{2.5.24}$$

where $\vec{\omega}_j = (\omega_j, \omega_{-j})^t \in \mathbb{R}^2$ and $\omega^n = (\vec{\omega}_1, \dots, \vec{\omega}_n, 0, \dots) \in \pi_n(h^p)$. We have the following theorem :

Theorem 2.5.12. *The curve $\omega^n(\cdot)$ converges to $v(\cdot)$ as $n \to \infty$ in $C([-T,T];h^p)$, where $v(\cdot)$ and $\omega^n(\cdot)$ are, respectively, solutions of (2.5.23) and (2.5.24) with initial data $v(0) \in h^p$ and $\omega^n(0) = v^n(0) \in \pi_n(h^p)$.*

The proof of this theorem is long and standard, using finite dimensional approximation. We move the detail of it to Appendix B and directly go to the main theorem of this subsection.

Let $\mathcal{S}_v^\tau$ denote the flow determined by equations (2.5.23) in the space h^p, and

$$B_p^v(M) := \{v \in h^p : |v|_p \leqslant M\}.$$

Theorem 2.5.13. *For any $M_0 > 0$, there exists a constant $C > 0$ which depends only on M_0 and T, such that if A is a open subset of $B_p^v(M_0)$, then for $\tau \in [0, T]$, we have*

$$e^{-C\tau}\mu(A) \leqslant \mu(\mathcal{S}_v^\tau(A)) \leqslant e^{C\tau}\mu(A).$$

Démonstration. From Assumption A, we know that there is constant M_1 which only depends on M_0 and T, such that if $v(0) \in B_p^v(M_0)$, then

$$v(\tau) \in B_p^v(M_1), \quad |\tau| \leqslant T. \tag{2.5.25}$$

For any $n \in \mathbb{N}$, consider the measure $\mu_n = \pi_n \circ \mu$ on the subspace $\pi_n(h^p)$. Since μ is a ζ_0'-admissible Gaussian measure, by (2.1.12) the measure μ_n has the following density with respect to the Lebesgue measure :

$$b_n(v^n) := (2\pi)^{-n}\prod_{j=1}^n (2\pi j)^{1+2p}\sigma_j^{-1}\exp\{-\frac{1}{2}\sum_{j=1}^n \frac{j^{1+2p}|\mathbf{v}_j|^2}{\sigma_j}\}.$$

Let $\mathcal{S}_n^\tau$ be the flow determined by equations (2.5.24) on subspace $\pi_n(h^p)$. For any open set $A_n \subset \pi_n(B_p^v(M_0))$, due to Theorem 2.A.1 in the Appendix 2.A, we have

$$\begin{aligned}
&\frac{d}{d\tau}\mu_n(\mathcal{S}_n^\tau(A_n)) \\
&= \int_{\mathcal{S}_n^\tau(A_n)} \sum_{j=1}^n \Big(\frac{\partial(b_n(v^n)X_j(v^n))}{\partial v_j} + \frac{\partial(b_n(v^n)X_{-j}(v^n))}{\partial v_{-j}}\Big)dv^n \\
&= \int_{\mathcal{S}_n^\tau(A_n)} \sum_{j=1}^n j^{2p+1}\Big(\frac{v_jX_j + v_{-j}X_{-j}}{\sigma_j} + \frac{\partial X_j}{\partial v_j} + \frac{\partial X_{-j}}{\partial v_{-j}}\Big)b_n(v^n)dv^n \\
&:= \int_{\mathcal{S}_n^\tau(A_n)} c^n(v^n)b_n(v^n)dv^n
\end{aligned}$$

Since $j^{-\zeta_0'}/\sigma_j = O(1)$, using the Cauchy's inequality and the assumption that $X(v)$ is ζ_0'-smoothing, there exists a constant C which depends only on M_1, such that

$$|c^n(v^n)| \leqslant C, \quad v^n \in \pi_n(B_p^v(M_1)), \quad \forall n \in \mathbb{N}. \tag{2.5.26}$$

We have

$$e^{-C\tau}\mu_n(A_n) \leqslant \mu_n(\mathcal{S}_n^\tau(A_n)) \leqslant e^{C\tau}\mu_n(A_n), \tag{2.5.27}$$

as long as $\mathcal{S}_n^\tau(A_n) \subset \pi_n(B_p^v(M_1))$.

Since μ_n convergences weakly to μ, the theorem follows from (2.5.25), (2.5.27) and Theorem 2.5.12. $\square$

2.6 Application to a special case

In this section we prove Proposition 2.1.6.

Clearly, we only need to prove the statement (i) of Assumption A.

Let $\mathcal{F} : H^m \to \mathbb{R}$ be a smooth functional (for some $m \geqslant 0$). If $u(t)$ is a solution of (2.1.13), then

$$\frac{d}{dt}\mathcal{F}(u(t)) = \langle \nabla\mathcal{F}(u(t)), -V(u) + \epsilon f(x)\rangle.$$

In particular, if $\mathcal{F}(u)$ is an integral of motion for the KdV equation, then we have $\langle \nabla\mathcal{F}(u(t), V(u)\rangle = 0$, so

$$\frac{d}{dt}\mathcal{F}(u(t)) = \epsilon\langle \nabla\mathcal{F}(u(t)), f(x)\rangle.$$

Since $||u(0)||_0^2$ is an integral of motion, then

$$\frac{d}{dt}||u(t)||_0^2 = 2\epsilon\langle u, f(x)\rangle \leqslant \epsilon(||u||_0^2 + ||f(x)||_0^2).$$

Thus we have

$$||u(t)||_0^2 \leqslant e^{\epsilon t}(||u(0)||_0^2 + \epsilon t||f(x)||_0^2). \tag{2.6.1}$$

The KdV equation has infinitively many integral of motion $\mathcal{J}_m(u)$, $m \geqslant 0$. The integral $\mathcal{J}_m$ can be writen as

$$\mathcal{J}_m(u) = ||u||_m^2 + \sum_{r=3}^{m}\sum_{\mathbf{m}}\int C_{r,\mathbf{m}} u^{(m_1)}\cdots u^{(m_r)}dx,$$

where the inner sum is taken over all integer r-vectors $\mathbf{m} = (m_1, \dots, m_r)$, such that $0 \leqslant m_j \leqslant m-1$, $j = 1, \dots, r$ and $m_1 + \dots + m_r = 4 + 2m - 2r$. Particularly, $\mathcal{J}_0(u) = ||u||_0^2$.

Lets consider

$$I = \int u^{(m_1)}\cdots f^{(m_i)}\cdots u^{(m_{r_1})}dx, \quad m_1 + \dots + m_{r_1} = M,$$

where $r_1 \geqslant 2$, $M \geqslant 1$, and $0 \leqslant m_j \leqslant \mu - 1$. Then, by Hölder's inequality,

$$|I| \leqslant ||u^{(m_1)}||_{L_{p_1}}\cdots||f(x)||_{L_{p_i}}\cdots||u^{m_{r_1}}||_{L_{p_f}}, \quad p_j = \frac{M}{m_j} \leqslant \infty.$$

Applying next the Gagliardo-Nirenberg and the Young inequalities, we obtain that

$$|I| \leqslant \delta||u||_\mu^2 + C_\delta||u||_0^{C_1}, \quad \forall\delta > 0, \tag{2.6.2}$$

where C_δ and C_1 do not depend on u. Below we denote C a positive constant independent of u, not necessary the same in each inequality. Let

$$I_1 := \langle\nabla\mathcal{J}_m(u), f\rangle = \langle u^{(m)}, f^{(m)}\rangle + \sum_{r=3}^{m}\sum_{\mathbf{m}} C'_{r,\mathbf{m}} u^{(m_1)}\cdots f^{(m_i)}\cdots u^{m_r}dx,$$

where $m_1 + \cdots + m_r = 6 + 2m - 2r$. Using (2.6.2) with a suitable δ, we get

$$I_1 \leqslant ||u||_m^2 + C||u||_0^{C_1} \leqslant ||u||_m^2 + C(1 + ||u||_0^{4m}) + ||f||_m^2. \tag{2.6.3}$$

If $u(t) = u(t, x)$ is a solution of equation (2.1.13), then

$$\frac{d}{dt}\mathcal{J}_m(u) = \langle \nabla \mathcal{J}_m(u), \epsilon f \rangle \leqslant \epsilon ||u||_m^2 + \epsilon C(1 + ||u||_0^{4m}) + \epsilon ||f||_m^2,$$

and

$$\frac{1}{2}||u||_m^2 - C(1 + ||u||_0^{4m}) \leqslant \mathcal{J}_m(u) \leqslant 2||u||_m^2 + C(1 + ||u||_0^{4m}).$$

Denote $C_m = C(1 + ||u(0)||_0^{4m}) + C||f||_m^2$, then from (2.6.1) and above, we deduce

$$\frac{d}{dt}(\mathcal{J}_m(u) - C_m) \leqslant \frac{1}{2}\epsilon(\mathcal{J}_m(u) - C_m),$$

thus

$$\mathcal{J}_m(u) - C_m \leqslant e^{\frac{1}{2}\epsilon t}[\mathcal{J}_m(u(0)) - C_m],$$

so

$$||u(t)||_m^2 \leqslant 4||u(0)||_m^2 e^{\frac{1}{2}\epsilon t} + C_m.$$

This prove Proposition 2.1.6. $\square$

2.A Liouville's theorem

Consider the following system of ordinary differential equations :

$$\dot{x} = Y(x), \quad x(0) = x_0 \in \mathbb{R}^n,$$

where $Y(x) = (Y_1(x), \cdots, Y_n(x)) : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable map. Let $F(t, x)$ be a (local) flow determined by this equation.

Theorem 2.A.1. *(Liouville). Let $B(x_1, \cdots, x_n)$ be a continuous differentiable function on $\mathbb{R}^n$. For the Borel measure $d\mu = B(x)dx$ in $\mathbb{R}^n$ and any bounded open set $A \subset \mathbb{R}^n$, we have*

$$\frac{d}{dt}\mu(F(t, A)) = \int_{F(t,A)} \Big[\sum_{i=1}^n \frac{\partial(B(x)Y_i(x))}{\partial x_i}\Big] dx, \quad t \in (-T, T),$$

where $T > 0$ is such that $F(t, x)$ is well defined and bounded for any $t \in (-T, T)$ and $x \in A$.

For $B = const$ this result is well known. For its proof for a non-constant density B see e.g. [75].

2.B Proof of Theorem 2.5.12

In this appendix, we give a detail proof of Theorem 2.5.12.

Fix any $M_0 > 0$. From Assumption A, we know that there exists a constant M_1 such that if $|v(0)|_p \leqslant M_0$, then

$$|v(\tau)|_p \leqslant M_1, \quad \tau \in [0,T]. \tag{2.B.1}$$

The equation (2.5.24) yields that

$$\frac{d}{d\tau}|\omega^n|_p^2 = 2\sum_{j=1}^{n} j^{1+2p}\vec{\omega}_j \cdot \mathbf{X}_j(\omega^n) := \chi^n(\omega^n). \tag{2.B.2}$$

We define

$$\chi(v) := 2\sum_{j=1}^{\infty} j^{1+2p}\mathbf{v}_j \cdot \mathbf{X}_j(v).$$

By smoothing assumption of $X(v)$, we know that there exists a constant $C_1 > 0$ such that

$$|\chi^n(\omega^n)| \leqslant C_1, \quad |\omega^n|_p \leqslant 2M_1, \quad \forall n \in \mathbb{N}. \tag{2.B.3}$$

Denote $\bar{\tau} = M_1/C_1$, then if $|\omega^n(0)|_p \leqslant M_0$, then

$$|\omega^n(\tau)|_p \leqslant 2M_1, \quad \tau \in [-\bar{\tau}, \bar{\tau}], \quad \forall n \in \mathbb{N}. \tag{2.B.4}$$

Lemma 2.B.1. *In the space $C([-\bar{\tau}, \bar{\tau}], h^{p-1})$, we have the convergence*

$$\omega^n(\cdot) \to v(\cdot) \quad \textit{as} \quad n \to \infty.$$

Démonstration. Denote $\vec{\xi}_j = \mathbf{v}_j - \vec{\omega}_j$, $I_v = I(v)$ and $I_{\omega^n} = I(\omega^n)$. Since $\mathcal{J}\mathbf{v}_j = \mathbf{v}_j^{\perp}$, using equations (2.5.23) and (2.5.24), for $1 \leqslant j \leqslant n$, we get

$$\begin{aligned}\frac{d}{d\tau}|\vec{\xi}_j|^2 &= 2(\vec{\xi}_j)^t[\epsilon^{-1}\mathcal{J}(W_j(I_v)\mathbf{v}_j - W_j(I_{\omega^n})\vec{\omega}_j) + \mathbf{X}_j(v) - \mathbf{X}_j(\omega^n)] \\ &= 2\epsilon^{-1}[W_j(I_v) - W_j(I_{\omega^n})]\mathbf{v}_j \cdot (\vec{\omega}_j)^{\perp} + 2(\vec{\xi}_j)^t \cdot (\mathbf{X}_j(v) - \mathbf{X}_j(\omega^n)).\end{aligned}$$

By Lemma 2.2.4 and Cauchy's inequality, we know that

$$\Big|W_j(I(v)) - W_j(I(\omega^n))\Big| \leqslant C_2(M_1) j |v - \omega^n|_{p-1}.$$

Using the smoothing of the mapping $X(v)$, we get that

$$\frac{d}{d\tau}|v - \omega^n|_{p-1}^2 \leqslant C_3(\epsilon, M_1)|v - \omega^n|_{p-1}^2 + a_n(v), \quad \tau \in [-\bar{\tau}, \bar{\tau}],$$

where

$$a_n(v) = \sum_{j=n+1}^{\infty} j^{2p-1}\mathbf{v}_j \cdot \mathbf{X}_j(v).$$

Obviously, $a_n(v) \to 0$ as $n \to \infty$ uniformly for $|v|_p \leqslant M_1$.

The lemma now follows directly from Gronwall's Lemma. □

Lemma 2.B.2. *If $\omega^n(0) \to v(0)$ strongly in h^p and $\tau_n \to \tau$, $\tau_n \in [-\bar{\tau}, \bar{\tau}]$, as $n \to \infty$, then*

$$\lim_{n\to\infty} |v(\tau) - \omega^n(\tau_n)|_p = 0.$$

Démonstration. From (2.B.2) we know that for any $\tau_n \in [-\bar{\tau}, \bar{\tau}]$,

$$|\omega^n(\tau_n)|_p^2 - |\omega^n(0)|_p^2 = \int_0^{\tau_n} \chi^n(\omega^n(s))ds.$$

Since $\omega^n(0) \to v(0)$ strongly in h^p, then using Lemma 2.B.1 we get

$$\begin{aligned}
&|v(\tau)|_p^2 \leqslant \liminf_{n\to\infty} |\omega^n(\tau_n)|_p^2 \leqslant \limsup_{n\to\infty} |\omega^n(\tau_n)|_p^2 \\
&= \limsup_{n\to\infty} \left(|\omega^n(0)|_p^2 + \int_0^{\tau_n} \chi^n(\omega^n(s))ds \right) = |v(0)|_p^2 + \int_0^{\tau} \chi(v(s))ds \\
&= |v(\tau)|_p^2.
\end{aligned}$$

Therefore, $\lim_{n\to\infty} |\omega^n(\tau_n)|_p = |v(\tau)|_p$. Since $\omega^n(\tau_n) \to v(\tau)$ in the space h^{p-1} as $n \to \infty$, then the required convergence follows. □

Lemma 2.B.3. *In the space $C([-\bar{\tau}, \bar{\tau}], h^p)$, $\omega^n(\cdot) \to v(\cdot)$ as $n \to \infty$.*

Démonstration. Suppose this statement is invalid. Then there exists $\delta > 0$ and a sequence $\{\tau^n\}_{n\in\mathbb{N}} \subset [-\bar{\tau}, \bar{\tau}]$ such that

$$|\omega^n(\tau^n) - v(\tau^n)|_p \geqslant \delta.$$

Let $\{\tau^{n_k}\}_{k\in\mathbb{N}}$ be a subsequence of the sequence $\{\tau^n\}_{n\in\mathbb{N}}$ converging to some $\tau^0 \in [-\bar{\tau}, \bar{\tau}]$. But $v(\tau^{n_k}) \to v(\tau^0)$ in h^p as $k \to \infty$, and using Lemma 2.B.2, we can get $\omega^{n_k}(\tau^{n_k}) \to v(\tau^0)$ as $k \to \infty$ in h^p. So we get a contradiction, and Lemma 2.B.3 is proved. □

If $T \leqslant \bar{\tau}$, the theorem is proved, otherwise we iterate the above procedure. This finishes the proof of Theorem 2.5.12. □

Chapitre 3

An averaging theorem for weakly nonlinear PDEs (non-resonant case)

The results of this chapter are taken from my paper [30].

Abstract : Consider nonlinear partial differential equations with small nonlinearities

$$\frac{d}{dt}u + i(-\triangle u + V(x)u) = \epsilon\mathcal{P}(\triangle u, \nabla u, u, x), \quad x \in \mathbb{T}^d. \qquad (*)$$

Let $\{\zeta_1(x), \zeta_2(x), \dots\}$ be the L_2-basis formed by eigenfunctions of the operator $-\triangle + V(x)$. For any complex function u, write it as $u(x) = \sum_{k\geqslant 1} v_k\zeta_k(x)$ and set $I(u) = (I_k(u), k \geqslant 1)$, where $I_k(u) = \frac{1}{2}|v_k|^2$. Then for any solution $u(t,x)$ of the linear equation $(*)_{\epsilon=0}$ we have $I(u(t,\cdot)) = const$. Suppose that the spectrum of the operator $-\triangle + V(x)$ is non-resonant. In this work it is proved that if $(*)$ is well posed on time-intervals $t \lesssim \epsilon^{-1}$ and satisfies there some mild a-priori assumptions, then for any its solution $u^\epsilon(t,x)$, the limiting behavior of the curve $I(u^\epsilon(t,\cdot))$ on time intervals of order ϵ^{-1}, as $\epsilon \to 0$, can be uniquely characterized by solutions of a certain well-posed effective equation.

3.1 Introduction

We consider the Schrödinger equation

$$\frac{d}{dt}u + i(-\triangle u + V(x)u) = 0, \quad x \in \mathbb{T}^d, \qquad (3.1.1)$$

and its nonlinear perturbation :

$$\frac{d}{dt}u + i(-\triangle u + V(x)u) = \epsilon\mathcal{P}(\triangle u, \nabla u, u, x), \quad x \in \mathbb{T}^d, \qquad (3.1.2)$$

where $\mathcal{P} : \mathbb{C}^{d+2}\times\mathbb{T}^d \to \mathbb{C}$ is a smooth function, $1 \leqslant V(x) \in C^n(\mathbb{T}^d)$ is a potential (we will assume that n is sufficiently large) and $\epsilon \in (0, 1]$ is the perturbation parameter. For any $p \in \mathbb{R}$ denote by H^p the Sobolev space of complex-valued periodic functions, provided with the norm $||\cdot||_p$,

$$||u||_p^2 = \Big\langle (-\triangle)^p u,\ u\Big\rangle + \langle u,\ u\rangle, \quad \text{if} \quad p \in \mathbb{N},$$

where $\langle\cdot,\cdot\rangle$ is the real scalar product in $L^2(\mathbb{T}^d)$,

$$\langle u,\ v\rangle = Re\int_{\mathbb{T}^d} u\bar{v}dx, \quad u,\ v\in L^2(\mathbb{T}^d).$$

If $p > \frac{d}{2}+2 = p_d$, then the mapping $H^p\to H^{p-2}$, $u(x)\mapsto \mathcal{P}(\triangle u, \nabla u, u, x)$ is smooth (see below Lemma 3.3.1). For any $T>0$, a curve $u\in C([0,T],H^p)$, $p>p_d$, is called a solution of (3.1.2) in H^p if it is a mild solution of this equation. That is, if the relation obtained by integrating (3.1.2) in t from 0 to s holds for any $0\leqslant s\leqslant T$. We wish to study long-time behaviours of solutions for (3.1.2) and assume :

Assumption A *(a-priori estimate). Fix some $T>0$. For any $p>p_d+2$, there exists $n_1(p)>0$ such that if $n\geqslant n_1(p)$, then for any $0<\epsilon\leqslant 1$, the perturbed equation (3.1.2), provided with initial data*

$$u(0)=u_0\in H^p, \tag{3.1.3}$$

has a unique solution $u(t,x)\in H^p$ such that

$$||u||_p\leqslant C(T,p,||u_0||_p), \quad \textit{for} \quad t\in[0,T\epsilon^{-1}].$$

Here and below the constant C also depends on the potential $V(x)$.

Denote the operator

$$A_V u := -\triangle u + V(x)u.$$

Let $\{\zeta_k\}_{k\geqslant 1}$ and $\{\lambda_k\}_{k\geqslant 1}$ be its real eigenfunctions and eigenvalues, ordered in such a way that

$$1\leqslant\lambda_1\leqslant\lambda_2\leqslant\cdots.$$

We say that a potential $V(x)$ is *non-resonant* if

$$\sum_{k=1}^{\infty}\lambda_k s_k\neq 0, \tag{3.1.4}$$

for every finite non-zero integer vector $(s_1,s_2,\cdots)$. For any complex-valued function $u(x)\in H^p$, we denote by

$$\Psi(u) := v = (v_1,v_2,\cdots), \quad v_j\in\mathbb{C}, \tag{3.1.5}$$

the vector of its Fourier coefficients with respect to the basis $\{\zeta_k\}$, i.e. $u(x)=\sum_{k=1}^{\infty}v_k\zeta_k$. In the space of complex sequences v, we introduce the norms

$$|v|_p^2=\sum_{k\geqslant 1}|v_k|^2\lambda_k^p, \quad p\in\mathbb{R},$$

and define $h^p := \{v : |v|_p<+\infty\}$. Denote

$$I_k=\frac{1}{2}|v_k|^2, \quad \varphi_k=\text{Arg } v_k, \quad k\geqslant 1. \tag{3.1.6}$$

Then $(I,\varphi)\in\mathbb{R}^\infty\times\mathbb{T}^\infty$ are the action-angles for the linear equation (3.1.1). That is, in these variables equation (3.1.1) takes the integrable form

$$\frac{d}{dt}I_k=0, \quad \frac{d}{dt}\varphi_k=\lambda_k, \quad k\geqslant 1. \tag{3.1.7}$$

Abusing notation we will write $v = (I, \varphi)$. Define h_I^p to be the weighted l^1-space

$$h_I^p := \Big\{ I = (I_1, \dots) \in \mathbb{R}^\infty : |I|_p^\sim < +\infty \Big\}, \quad |I|_p^\sim = 2\sum_{i=1}^{\infty} \lambda_i^p |I_i|,$$

and consider the mapping

$$\pi_I : h^p \to h_I^p,\ v \mapsto I, \quad I_j(v) = \frac{1}{2}|v_j|^2, \quad j \geqslant 1.$$

It is continuous and its image is the positive octant $h_{I+}^p = \{I \in h_I^p : I_j \geqslant 0, \forall j\}$.

We mainly concern with the long time behavior of the actions $I(u(t)) \in \mathbb{R}_+^\infty$ of solutions for the perturbed equation (3.1.2) for $t \lesssim \epsilon^{-1}$. For this purpose, it is convenient to pass to the slow time $\tau = \epsilon t$ and write equation (3.1.2) in the action-angle coordinates (I, φ) :

$$\dot{I}_k = F_k(I, \varphi), \quad \dot{\varphi}_k = \epsilon^{-1}\lambda_k + G_k(I, \varphi), \quad k \geqslant 1, \tag{3.1.8}$$

where $I \in \mathbb{R}^\infty$, $\varphi \in \mathbb{T}^\infty$ and $\mathbb{T}^\infty := \{(\theta_i)_{i\in\mathbb{N}} : \theta_i \in \mathbb{T}\}$ is the infinite-dimensional torus endowed with the Tikhonov toppology. The functions F_k and G_k, $k \geqslant 1$ represent the perturbation term $\mathcal{P}$, written in the action-angle coordinates. In the finite dimensional situation, the *averaging principle* is well established for perturbed integrable systems. The principle states that for equations

$$\frac{d}{dt}I = \epsilon f(I, \varphi), \quad \frac{d}{dt}\varphi = W(I) + \epsilon g(I, \varphi),$$

where $I \in \mathbb{R}^M$ and $\varphi \in \mathbb{T}^m$, on time intervals of order ϵ^{-1} the action components $I(t)$ can be well approximated by solutions of the following averaged equation :

$$\frac{d}{dt}J = \epsilon\langle f\rangle(J), \quad \langle f\rangle(J) = \int_{\mathbb{T}^m} f(J, \varphi)d\varphi. \tag{3.1.9}$$

This assertion has been justified under various non-degeneracy assumptions on the frequency vector W and the initial data $(I(0), \varphi(0))$ (see [57]). In this paper we want to prove a version of the averaging principle for the perturbed Schrödinger equation (3.1.2). We define a corresponding averaged equation for (3.1.8) as in (3.1.9) :

$$\dot{J}_k = \langle F_k\rangle(J), \quad \langle F_k\rangle(J) = \int_{\mathbb{T}^\infty} F_k(J, \varphi)d\varphi, \quad k \geqslant 1, \tag{3.1.10}$$

where $d\varphi$ is the Haar measure on $\mathbb{T}^\infty$. But now, in difference with the finite-dimensional case, the well-posedness of equation (3.1.10) is not obvious, since the map $\langle F\rangle(I) = (\langle F_1\rangle(I), \dots)$ is unbounded and the functions $\langle F_k\rangle(I)$, $k \geqslant 1$, may be not Lipschitz with respect to I in h_{I+}^p. In [49], S. Kuksin observed that the averaged equation (3.1.10) may be lifted to a regular 'effective equation' on the variable $v \in h^p$, which transforms to (3.1.10) under the projection π_I. To derive an effective equation, corresponding to our problem, we first use mapping Ψ to write (3.1.2) as a system of equation on the vector $v(\tau)$:

$$\dot{v} = \epsilon^{-1}d\Psi(u)(-iA_V(u)) + P(v). \tag{3.1.11}$$

Here $P(v)$ is the perturbation term $\mathcal{P}$, written in v-variables. This equation is singular when $\epsilon \to 0$. The effective equation for (3.1.11) is a certain regular equation

$$\dot{v} = R(v). \tag{3.1.12}$$

To define the effective vector filed $R(v)$, for any $\theta = (\theta_1, \theta_2, \cdots) \in \mathbb{T}^\infty$ let us denote by Φ_θ the linear operator in the space of complex sequences $(v_1, v_2, \cdots) \in h^p$ which multiplies each component v_j with $e^{i\theta_j}$. Rotation Φ_θ acts on vector fields on the v-space, and $R(v)$ is the result of action of Φ_θ on $P(v)$, averaged in θ :

$$R(v) = \int_{\mathbb{T}^\infty} \Phi_{-\theta} P(\Phi_\theta v) d\theta.$$

The map $R(v)$ is smooth with respect to v in h^p. Again, we understand solutions for equation (3.1.12) in the mild sense.

We now make the second assumption :

Assumption B *(local well-posedness of the effective equation). For any $p > p_d + 2$, there exists $n_2(p) > 0$ such that if $n \geqslant n_2(p)$, then for any initial data $v_0 \in h^p$, there exists $T(|v_0|_p) > 0$ such that the effective equations (3.1.12) has a unique solution $v \in C([0, T(|v_0|_p)], h^p)$. Here $T : \mathbb{R}_+ \to \mathbb{R}_{>0}$ is an upper semi-continuous function.*

The main result of this paper is the following statement, where $v^\epsilon(\tau)$ is the Fourier transform of a solution $u^\epsilon(t, x)$ for the problem (3.1.2), (3.1.3) (existing by Assumption A), written in the slow time $\tau = \epsilon t$:

$$v^\epsilon(\tau) = \Psi\Big(u^\epsilon(\epsilon^{-1}\tau)\Big), \quad \tau \in [0, T].$$

We also assume Assumption B.

Theorem 3.1.1. *For any $p > p_d + 2$, if $n \geqslant \max\{p, n_1(p), n_2(p)\}$, then there exists $I^0(\cdot) \in C([0, T], h_I^p)$ such that for every $q < p$,*

$$I(v^\epsilon(\cdot)) \underset{\epsilon \to 0}{\longrightarrow} I^0(\cdot) \quad in \quad C([0, T], h_I^q).$$

Moreover $I^0(\tau)$, $\tau \in [0, T]$, solves the averaged equation (3.1.10) with initial data $I^0(0) = I(\Psi(u_0))$, and it may be written as $I^0(\tau) = I(v(\tau))$, where $v(\cdot)$ is the unique solution of the effective equation (3.1.12), equal to $\Psi(u_0)$ at $\tau = 0$.

Proposition 3.1.2. *The assumptions A and B hold if (3.1.2) is a complex Ginzburg-Landau equation*

$$\dot{u} + \epsilon^{-1} i(-\triangle u + V(x)u) = \triangle u - \gamma_R f_p(|u|^2)u - i\gamma_I f_q(|u|^2)u, \quad x \in \mathbb{T}^d, \tag{3.1.13}$$

where the constants γ_R, γ_I satisfy

$$\gamma_R, \ \gamma_I > 0, \tag{3.1.14}$$

the functions $f_p(r)$ and $f_q(r)$ are the monomials $|r|^p$ and $|r|^q$, smoothed out near zero, and

$$0 \leqslant p, q < \infty \quad if \quad d = 1, 2 \quad and \quad 0 \leqslant p, q < \min\{\frac{d}{2}, \frac{2}{d-2}\} \quad if \quad d \geqslant 3. \tag{3.1.15}$$

This work is a continuation of the research started in [31], where the author proved a similar averaging principle (not for all but for typical initial data) for a perturbed KdV equation :

$$u_t + u_{xxx} - 6uu_x = \epsilon f(u)(x),\ x \in \mathbb{T},\ \int_{\mathbb{T}} u(t,x)dx = 0, \tag{3.1.16}$$

assuming the perturbation $\epsilon f(u)(\cdot)$ defines a smoothing mapping $u(\cdot) \mapsto f(u)(\cdot)$. This additional assumption is necessary to guarantee the existence of an quasi-invariant measure for the perturbed equation (3.1.16), which plays an essential role in the proof due to the non-linear nature of the unperturbed equation. Since in the present paper we deal with perturbations of a linear equation, this restriction is not needed.

In [50], a result similar to Theorem 3.1.1 was proved for weakly nonlinear stochastic CGL equation (3.1.13). There are many works on long-time behaviors of solutions for nonlinear Schrödinger equations. E.g. the averaging principle was justified in [36] for solutions of Hamiltonian perturbations of (3.1.1), provided that the potential $V(x)$ is non-degenerated and that the initial data $u_0(x)$ is a sum of finitely many Fourier modes. Several long-time stability theorems which are applicable to small amplitude solutions of nonlinear Schrödinger equations were presented in [4, 7, 68, 15]. The results in these works describe the dynamics over a time scale much longer than the $\mathcal{O}(\epsilon^{-1})$ that we consider, precisely, over a time interval of order ϵ^{-m}, with arbitrary m (even of order $\exp \epsilon^{-\delta}$ with $\delta > 0$ in [4, 68, 15]). These results are obtained under the assumption that the frequencies are completely resonant or highly non-resonant (Diophantine-type), by using the normal form techniques near an equilibrium (this is the reason for which they only apply to small amplitude solutions). See [6] and references therein for general theory of normal form for PDEs. In difference with the mentioned works, the research in this paper is based on the classical averaging method for finite dimensional systems, characterizing by the existence of slow-fast variables. It deals with arbitrary solution of equation (3.1.2) with sufficiently smooth initial data. Also note that the non-resonance assumption (3.1.4) is significantly weaker than those in the mentioned works.

Plan of the Chapter. In Section 3.2 we recall some spectral properties of the operator A_V. Section 3.3 is about the action-angle form of the perturbed linear Schrödinger equation (3.1.2). In Section 3.4 we introduce the averaged equation and the corresponding effective equation. Theorem 3.1.1 and Proposition 0.2 are proved in Section 3.5 and Section 3.6.

3.2 Spectral properties of A_V

As in the introduction, $A_V = -\triangle + V(x)$, $x \in \mathbb{T}^d$, where $1 \leqslant V(x) \in C^n(\mathbb{T}^d)$ and $\{\lambda_k\}_{k\geqslant 1}$ are the eigenvalues of A_V. According to Weyl's law, the λ_k, $k \geqslant 1$, satisfy the following asympototics

$$\lambda_k = C_d k^{2/d} + o(k^{2/d}), \quad k \geqslant 1,$$

Fix an L^2-orthogonal basis of eigenfunctions $\{\zeta_k\}_{k\geqslant 1}$ corresponding to the eigenvalues $\{\lambda_k\}_{k\geqslant 1}$, and define the linear mapping Ψ as (3.1.5). For any $m \in \mathbb{N}$, we have

$\langle A_V^m u, u\rangle = |v|_m^2$, where $v = \Psi u$. Noting that $\langle A_V^m u, u\rangle$ is equivalent to $||u||_m^2$ for $m = 1, \dots, n$, since $V(x)$ is C^n-smooth, we have the following :

Lemma 3.2.1. *For every integer $p \in [0, n]$ the linear mapping $\Psi : H^p \to h^p$ is an isomorphism.*

We denote
$$C_{+1}^n(\mathbb{T}^d) := \{V(x) \geqslant 1 : V(x) \in C^n(\mathbb{T}^d)\}.$$
For any finite $M \in \mathbb{N}$ consider the mapping

$$\Lambda^M : C_{+1}^n(\mathbb{T}^d) \to \mathbb{R}^M, \quad V(x) \to (\lambda_1, \cdots, \lambda_M),$$

and define the open domain $E_M \subset C_{+1}^n(\mathbb{T}^d)$,

$$E_M := \{V | \lambda_1 < \lambda_2 < \cdots < \lambda_M\}.$$

The complement of E_M is a real analytic variety in $C^n(\mathbb{T}^d)$ of codimension at least 2, so E_M is connected. The mapping Λ^M is analytic in E_M (see [36]).

Let μ be a Gaussian measure with a non-degenerate correlation operator, supported by the space $C^n(\mathbb{T}^d)$ (see [11]). Then $\mu(C_{+1}^n(\mathbb{T}^d)) > 0$. Fix $s \in \mathbb{Z}^M \setminus \{0\}$. The set
$$Q_s := \{V \in E_M | \Lambda^M(V) \cdot s = 0\},$$
is closed in E_M. Since the analytic function $\Lambda^M(V) \cdot s \not\equiv 0$ on E_M (e.g. see [36]), then $\mu(Q_s) = 0$ (see chapter 9 in [11] and the note [12]). Since this is true for any M and s as above, then we have :

Proposition 3.2.2. *The non-resonant potentials form a subset of $C_{+1}^n(\mathbb{T}^d)$ of full μ-measure.*

Note that this subset also is dense in $C_{+1}^n(\mathbb{T}^d)$ due to the fact that the Gaussian measure μ assigns every open subset of $C_{+1}^n(\mathbb{T}^d)$ with positive measure.

3.3 Equation (3.1.2) in action-angle variables

For $k = 1, 2, \dots,$ we denote :

$$\Psi_k : H^p \to \mathbb{C}, \quad \Psi_k(u) = v_k,$$

(see (3.1.5)). Let $u(t)$ be a solution of equation (3.1.2). Passing to slow time $\tau = \epsilon t$, we get for $v_k = \Psi_k(u(\tau))$ equations

$$\dot{v}_k + i\epsilon^{-1}\lambda_k v_k = \Psi_k(\mathcal{P}(\triangle u, \nabla u, u, x)), \quad k \geqslant 1. \tag{3.3.1}$$

Since $I_k(v) = \frac{1}{2}|\Psi_k|^2$ is an integral of motion for the Schrödinger equation (3.1.1), we have
$$\dot{I}_k = (\Psi_k(\mathcal{P}(\triangle u, \nabla u, u, x)), v_k) := F_k(v), \quad k \geqslant 1 \tag{3.3.2}$$
(Here and below $(\cdot, \cdot)$ indicates the real scalar product in $\mathbb{C}$, i.e. $(u, v) = Re\ u\bar{v}$.)

Denote $\varphi_k = \text{Arg } v_k$, if $v_k \neq 0$, and $\varphi_k = 0$, if $v_k = 0$, $k \geqslant 1$. Using equation (3.3.1), we get

$$\dot{\varphi}_k = \epsilon^{-1}\lambda_k + |v_k|^{-2}(\Psi_k(\mathcal{P}(\triangle u, \nabla u, u, x)), iv_k), \quad \text{if} \quad v_k \neq 0, \quad k \geqslant 1 \tag{3.3.3}$$

Denoting for brevity, the vector field in equation (3.3.3) by $\epsilon^{-1}\lambda_k + G_k(v)$, we rewrite the equation for the pair $(I_k, \varphi_k)(k \geqslant 1)$ as

$$\dot{I}_k = F_k(v) = F_k(I, \varphi), \quad \dot{\varphi}_k = \epsilon^{-1}\lambda_k + G_k(v). \tag{3.3.4}$$

(Note that the second equation has a singularity when $I_k = 0$.) We denote

$$F(I, \varphi) = (F_1(I, \varphi), F_2(I, \varphi), \cdots).$$

The following result is well known, see e.g. Section 5.5.3 in [69].

Lemma 3.3.1. *If $f(x) : \mathbb{C}^m \to \mathbb{C}^N$ is C^∞, then the mapping*

$$M_f : H^p(\mathbb{T}^d, \mathbb{C}^m) \to H^p(\mathbb{T}^d, \mathbb{C}^N), \quad u \mapsto f(u),$$

is C^∞-smooth for $p > d/2$. Moreover, it is bounded and Lipschitz, uniformly on bounded subsets of $H^p(\mathbb{T}^d, \mathbb{C}^m)$.

In the lemma below, P_k and P_k^j are some fixed continuous functions.

Lemma 3.3.2. *For any $j, k \in \mathbb{N}$, we have for any $p > p_d$*

(i)The function $F_k(v)$ is smooth in each space h^p.

(ii) For any $\delta > 0$, the function $G_k(v)\chi_{\{I_k \geqslant \delta\}}$ is bounded by $\delta^{-1/2}P_k(|v|_p)$.

(iii)For any $\delta > 0$, the function $\frac{\partial F_k}{\partial I_j}(I, \varphi)\chi_{\{I_j \geqslant \delta\}}$ is bounded by $\delta^{-1/2}P_k^j(|v|_p)$.

(iv) The function $\frac{\partial F_k}{\partial \varphi_j}(I, \varphi)$ is bounded by $P_k^j(|v|_p)$ and for any $m \in \mathbb{N}$ and any $(I_1, \cdots, I_m) \in \mathbb{R}_+^m$, the fucntion $F_k(I_1, \varphi_1, \cdots, I_m, \varphi_m, 0, \cdots)$ is smooth on $\mathbb{T}^m$.

Démonstration. Item (i) and (ii) follow directly from (3.3.2), (3.3.3), Lemmata 3.2.1 and 3.3.1. Item (iii) and (iv) follow directly from item (i) and the chain rule. □

Denote

$$\Pi_{I,\varphi} : h^p \to h_I^p \times \mathbb{T}^\infty, \quad \Pi_{I,\varphi}(v) = (I(v), \varphi(v)). \tag{3.3.5}$$

Definition 3.3.3. *Let assumption A holds. Then for any $p \geqslant p_d + 2$ and $T > 0$, we call a curve $(I(\tau), \varphi(\tau))$, $\tau \in [0, T]$, a regular solution of equation (3.3.4), if there is a solution $u(t) \in H^p$ of equation (3.1.2) such that*

$$\Pi_{I,\varphi}(\Psi(u(\epsilon^{-1}\tau))) = (I(\tau), \varphi(\tau)) \in h_I^p \times \mathbb{T}^\infty, \quad \tau \in [0, T].$$

Note that if $(I(\tau), \varphi(\tau))$ is a regular solution, then each $I_j(\tau)$ is a C^1-function, while $\varphi_j(\tau)$ may be discontinuous at points τ, where $I_j(\tau) = 0$.

For any $p \geqslant p_d + 2$, let $(I(\tau), \varphi(\tau))$ be a regular solution of (3.3.4) such that $|I(0)|_p \leqslant M_0$. Then by assumption A, for any $\epsilon > 0$ and $T > 0$, we have

$$|I(\tau)|_p^\sim = \frac{1}{2}|v(p)|_p^2 \leqslant C(p, M_0, T), \quad t \in [0, T]. \tag{3.3.6}$$

3.4 Averaged equation and Effective equation

For a function f on a Hilbert space H, we write $f \in Lip_{loc}(H)$ if

$$|f(u_1) - f(u_2)| \leqslant P(R)||u_1 - u_2||, \quad \text{if} \quad ||u_1||, ||u_2|| \leqslant R, \tag{3.4.1}$$

for a suitable continuous function P which depends on f. Clearly, the set of functions $Lip_{loc}(H)$ is an algebra. By Lemma 3.3.1,

$$F_k(v) \in Lip_{loc}(h^p), \quad k \in \mathbb{N},\ p > p_d. \tag{3.4.2}$$

Let $f \in Lip_{loc}(h^p)$ and $v \in h^{p_1}$, where $p_1 > p$. Denoting by Π^M, $M \geqslant 1$ the projection

$$\Pi^M : h^0 \mapsto h^0, \quad (v_1, v_2, \cdots) \mapsto (v_1, \cdots, v_M, 0, \cdots),$$

we have

$$|v - \Pi^M v|_p \leqslant \lambda_M^{-(p_1-p)/2} |v|_{p_1}.$$

Accordingly,

$$|f(v) - f(\Pi^M v)| \leqslant P(|v|_p) \lambda_M^{-(p_1-p)/2} |v|_{p_1}. \tag{3.4.3}$$

We will denote $v^M = (v_1, \ldots, v_M)$ and identify v^M with $(v_1, \ldots, v_M, 0, \ldots)$ if needed. Similar notations will be used for vectors $\theta = (\theta_1, \theta_2, \ldots) \in \mathbb{T}^\infty$ and vectors $I = (I_1, \ldots) \in h_I^p$.

The torus $\mathbb{T}^M$ acts on the space $\Pi_M h^0$ by linear transformations Φ_{θ^M}, $\theta^M \in \mathbb{T}^M$, where $\Phi_{\theta^M} : (I^M, \varphi^M) \mapsto (I^M, \varphi^M + \theta^M)$. Similarly, the tous $\mathbb{T}^\infty$ acts on h^0 by linear transformations $\Phi_\theta : (I, \varphi) \mapsto (I, \varphi + \theta)$ with $\theta \in \mathbb{T}^\infty$.

For a function $f \in Lip_{loc}(h^p)$ and any positive integer N, we define the average of f in the first N angles as

$$\langle f \rangle_N(v) = \int_{\mathbb{T}^N} f\Big((\Phi_{\theta^N} \oplus \mathrm{id})(v)\Big) d\theta^N,$$

and define the averaging in all angles as

$$\langle f \rangle_\varphi(v) = \int_{\mathbb{T}^\infty} f(\Phi_\theta(v)) d\theta,$$

where $d\theta$ is the Haar measure on $\mathbb{T}^\infty$. We will denote $\langle \cdot \rangle_\varphi$ as $\langle \cdot \rangle$ when there is no confusion. The estimate (3.4.3) readily implies that

$$|\langle f \rangle_N(v) - \langle f \rangle(v)| \leqslant P(R) \lambda_N^{-(p_1-p)/2}, \quad \text{if} \quad |v|_{p_1} \leqslant R.$$

Let $v = (I, \varphi)$, then $\langle f \rangle_N$ is a function independent of $\varphi_1, \cdots, \varphi_N$, and $\langle f \rangle$ is independent of φ. Thus $\langle f \rangle$ can be written as $\langle f \rangle(I)$.

Lemma 3.4.1. *Let $f \in Lip_{loc}(h^p)$, then*

(i) Functions $\langle f \rangle_N(v)$ and $\langle f \rangle$ satisfy (3.4.1) with the same function P as f and take the same value at the origin.

(ii) They are smooth if f is. If f is C^∞-smooth, then for any M, $\langle f \rangle(I)$ is a smooth function of the first M components $I_1, \cdots, I_M$ of the vector I.

Démonstration. Item (i) and the first statement of item (ii) is obvious. Notice that $\langle f\rangle(v)=\langle f\rangle(\sqrt{I_1},\dots)$ is even on each variable $\sqrt{I_j}$, $j\geqslant 1$, i.e.

$$\langle f\rangle(\dots,-\sqrt{I_j},\dots)=\langle f\rangle(\dots,\sqrt{I_j},\dots),\quad j\geqslant 1.$$

Now the second statement of item (ii) follows from Whitney's theorem (see Lemma A in the Appendix 3.A). □

Denote $C^{0+1}(\mathbb{T}^n)$ the set of all Lipschitz functions on $\mathbb{T}^n$. The following result is a version of the classical Weyl theorem.

Lemma 3.4.2. *Let $f\in C^{0+1}(\mathbb{T}^n)$ for some $n\in\mathbb{N}$. For any non-resonant vector $\omega\in\mathbb{R}^n$ (see (3.1.4)) and any $\delta>0$, there exists $T_0>0$ such that if $T\geqslant T_0$, $g\in C(\mathbb{T}^n)$ and $|g-f|\leqslant\delta/3$, then we have*

$$\Big|\frac{1}{T}\int_0^T g(x_0+\omega t)dt-\langle g\rangle\Big|\leqslant\delta,$$

uniformly in $x_0\in\mathbb{T}^n$.

Démonstration. It is well known that for any $\delta>0$ and non-resonant vector $\omega\in\mathbb{R}^n$, there exists $T_0>0$ such that

$$\Big|\frac{1}{T}\int_0^T f(x_0+\omega t)dt-\langle f\rangle\Big|\leqslant\delta/3,\quad\forall T\geqslant T_0,$$

(see e.g. Lemma 2.3.3 in Chapter 2). Therefore if $T\geqslant T_0$, $g\in C(\mathbb{T}^n)$ and $|g-f|\leqslant\delta/3$, then

$$\begin{aligned}\Big|\frac{1}{T}\int_0^T g(x_0+\omega t)dt-\langle g\rangle\Big|\leqslant&\Big|\frac{1}{T}\int_0^T f(x_0+\omega t)dt-\langle f\rangle\Big|\\&+\frac{1}{T}\int_0^T|f(x_0+\omega t)-g(x_0+\omega t)|dt+|\langle f\rangle-\langle g\rangle|\leqslant\delta.\end{aligned}$$

This finishes the proof of the lemma. □

We denote $P_k(v)=\Psi_k(\mathcal{P}(\triangle u,\nabla u,u,x))|_{u=\Psi^{-1}v}$, then equations (3.3.4) becomes

$$\dot I_k=(v_k,P_k(v)),\quad\dot\varphi_k=\epsilon^{-1}\lambda_k+G_k(v),\quad k\geqslant 1.\tag{3.4.4}$$

The averaged equations have the form

$$\dot J_k=\langle(v_k,P_k)\rangle_\varphi(J),\quad k\geqslant 1,\tag{3.4.5}$$

i.e.

$$\langle(v_k,P_k)\rangle_\varphi=\int_{\mathbb{T}^\infty}(v_ke^{i\theta_k},P_k(\Phi_\theta v))d\theta=(v_k,R_k(v)),\tag{3.4.6}$$

with

$$R_k(v)=\int_{\mathbb{T}}\Phi_{-\theta_k}P_k(\Phi_\theta)d\theta.\tag{3.4.7}$$

Similar to equation (3.1.2), for any $T > 0$, we call a curve $J \in C([0,T], h_I^p)$ a solution of equation (3.4.5) if for every $s \in [0,T]$ it satisfies the relation, obtained by integrating (3.4.5).

Consider the differential equations

$$\dot{v}_k = R_k(v), \quad k \geqslant 1. \tag{3.4.8}$$

Solutions of this system are defined similar to that of (3.1.2) and (3.4.5). Relation (3.4.6) implies :

Lemma 3.4.3. *If $v(\cdot)$ satisfies (3.4.8), then $I(v)$ satisfies (3.4.5).*

Following [49], we call equations (3.4.8) the *effective equation* for the perturbed equation (3.1.2).

Proposition 3.4.4. *The effective equation is invariant under the rotation Φ_θ. That is, if $v(\tau)$ is a solution of (3.4.8), then for each $\theta \in \mathbb{T}^\infty$, $\Phi_\theta v(\tau)$ also is a solution.*

Démonstration. Applying Φ_θ to (3.4.8) we get that

$$\frac{d}{d\tau}\Phi_\theta v = \Phi_\theta R(v).$$

Relation (3.4.7) implies that operations R and Φ_θ commute. Therefore

$$\frac{d}{d\tau}\Phi_\theta v = R(\Phi_\theta v).$$

The assertion follows. □

3.5 Proof of the Averaging theorem

In this section we prove the Theorem 3.1.1 by studying the behavior of regular solutions of equation (3.3.4). We fix $p \geqslant p_d + 2$, assume $n \geqslant \max\{p, n_1(p), n_2(p)\}$ and consider $u_0 \in H^p$. So

$$\Pi_{I,\varphi}(\Psi(u_0)) = (I_0, \varphi_0) \in h_{I+}^p \times \mathbb{T}^\infty. \tag{3.5.1}$$

We denote

$$B_p(M) = \{I \in h_{I+}^p : |I|_p^\sim \leqslant M\}. \tag{3.5.2}$$

Without loss of generality, we assume $T = 1$. Fix any $M_0 > 0$. Let

$$(I_0, \varphi_0) \in B_p(M_0) \times \mathbb{T}^\infty := \Gamma_0,$$

and let $(I(\tau), \varphi(\tau))$ be a regular solution of system (3.3.4) with $(I(0), \varphi(0)) = (I_0, \varphi_0)$. Then by (3.3.6), there exists $M_1 \geqslant M_0$ such that

$$I(\tau) \in B_p(M_1), \quad \tau \in [0,1]. \tag{3.5.3}$$

All constants below depend on M_1 (i.e. on M_0), and usually this dependence is not indicated. From the definition of the perturbation and Lemma 3.3.1 we know that

$$|\mathcal{F}(I,\varphi)|_{p-2}^{\sim} \leqslant C_{M_1}, \quad \forall (I,\varphi)\in B_p(M_1)\times\mathbb{T}^\infty. \tag{3.5.4}$$

Recall that we identify $I^m=(I_1,\dots,I_m)$ with $(I_1,\dots,I_m,0,\dots)$, etc.

Fix any $n_0\in\mathbb{N}$. By (3.4.2), for every $\rho>0$, there is $m_0\in\mathbb{N}$, depending only on n_0, M_1 and ρ, such that if $m\geqslant m_0$, then

$$|F_k(I,\varphi)-F_k(I^m,\varphi^m)|\leqslant\rho, \quad \forall (I,\varphi)\in B_p(M_1)\times\mathbb{T}^\infty, \tag{3.5.5}$$

where $k=1,\cdots,n_0$.

From now on, we always assume that $(I,\varphi)\in B_p(M_1)\times\mathbb{T}^\infty$.

Since $V(x)$ is non-resonant, then by Lemma 3.3.2 and Lemma 3.4.2, for any $\rho>0$, there exists $T_0=T_0(\rho,n_0)>0$, such that for all $\varphi\in\mathbb{T}^\infty$ and $T\geqslant T_0$,

$$\left|\frac{1}{T}\int_0^T F_k(I^{m_0},\varphi^{m_0}+\Lambda^{m_0}t)dt-\langle F_k\rangle(I^{m_0})\right|<\rho, \tag{3.5.6}$$

where $k=1,\dots,n_0$. Due to Lemma 3.3.2, we have

$$\begin{aligned}
&|G_j(I,\varphi)|\leqslant\frac{C_0(j,M_1)}{\sqrt{I_j}}, \quad \text{if} \quad I_j\neq 0,\\
&|\frac{\partial F_k}{\partial I_j}(I,\varphi)|\leqslant\frac{C_0(k,j,M_1)}{\sqrt{I_j}}, \quad \text{if} \quad I_j\neq0,\\
&|\frac{\partial F_k}{\partial\varphi_j}(I,\varphi)|\leqslant C_0(k,j,M_1).
\end{aligned}\tag{3.5.7}$$

From Lemma 3.1, we know

$$|\langle F_k\rangle(I^{m_0})-\langle F_k\rangle(\bar I^{m_0})|\leqslant C_1(k,m_0,M_1)|I^{m_0}-\bar I^{m_0}|, \tag{3.5.8}$$

and by (3.4.2),

$$|F_k(I^{m_0},\varphi^{m_0})-F_k(\bar I^{m_0},\bar\varphi^{m_0})|\leqslant C_2(k,m_0,M_1)|v^{m_0}-\bar v^{m_0}|, \tag{3.5.9}$$

where $\Pi_{I,\varphi}(v^{m_0})=(I^{m_0},\varphi^{m_0})$ (see (3.3.5)) and $|\cdot|$ is the l^∞-norm. Denote

$$C_{M_1}^{n_0,m_0}=m_0\cdot\max\{C_0,C_1,C_2 : 1\leqslant j\leqslant m_0, 1\leqslant k\leqslant n_0\}.$$

From now on we shall use the slow time $\tau=\epsilon t$.

Lemma 3.5.1. *For $k=1,\dots,n_0$, the I_k-component of any regular solution of (3.3.4) with initial data in Γ_0 can be written as :*

$$I_k(\tau)=I_k(0)+\int_0^\tau\langle F_k\rangle(I(s))ds+\Xi(\tau),$$

where for any $\gamma\in(0,1)$ the function $|\Xi(\tau)|$ is bounded on $[0,1]$ by

$$\begin{aligned}
|\Xi(\tau)|\leqslant C_{M_1}^{n_0,m_0}\Big[&\frac{T_0\epsilon}{2\gamma^{1/2}}+\frac{T_0C_{M_1}\epsilon}{2\gamma^{1/2}}+T_0C_{M_1}\epsilon\\
&+4(\gamma+T_0C_{M_1}\epsilon)^{1/2}\Big](\epsilon T_0+1)+3\rho+3\epsilon C_{M_1}T_0 \quad \tau\in[0,1],
\end{aligned}\tag{3.5.10}$$

where $\rho>0$ is arbitary and $T_0=T_0(\rho,n_0)$ is as (3.5.6).

Démonstration. Let us divide the time interval $[0,\tau]$, $\tau \leqslant 1$, into subinterval $[a_i, a_{i+1}]$, $0 \leqslant i \leqslant d_0$, such that

$$a_0 = 0, a_{d_0} = \tau, \quad a_{d_0} - a_{d_0-1} \leqslant \epsilon T_0,$$

and $a_{i+1} - a_i = \epsilon T_0$, for $0 \leqslant i \leqslant d_0 - 2$. Then $d_0 \leqslant (T_0\epsilon)^{-1} + 1$. For each interval $[a_i, a_{i+1}]$ we define a subset $\Omega(i) \subset \{1, 2, \cdots, m_0\}$ in the following way :

$$l \in \Omega(i) \quad \Longleftrightarrow \quad \exists \tau \in [a_i, a_{i+1}], \quad I_l(\tau) < \gamma.$$

Then if $l \in \Omega(i)$, by (3.5.4) we have

$$|I_l(\tau)| < T_0 C_{M_1}\epsilon + \gamma, \quad \tau \in [a_i, a_{i+1}].$$

For $I = (I_1, I_2, \cdots)$ and $\varphi = (\varphi_1, \varphi_2, \cdots)$ we set

$$\kappa_i(I) = \hat{I}, \quad \kappa_i(\varphi) = \hat{\varphi},$$

where the vectors $\hat{I}$ and $\hat{\varphi}$ are defined as follows :

$$\text{If} \quad l \in \Omega(i), \quad \text{then} \quad \hat{I}_l = 0, \hat{\varphi}_l = 0, \quad \text{else} \quad \hat{I}_l = I_l,\ \hat{\varphi}_l = \varphi_l.$$

We abbreviate $\kappa_i(I,\varphi) = (\kappa_i(I), \kappa_i(\varphi))$.

Below, $k = 1, \ldots, n_0$.

Then on $[a_i, a_{i+1}]$, noting $|v^{m_0} - \kappa_i(v^{m_0}| = \sqrt{2}|I^{m_0} - \kappa_i(I^{m_0})|^{1/2}$, and using (3.5.9) we have

$$\begin{aligned}
&\int_{a_i}^{a_{i+1}} \left| F_k\Big(I^{m_0}(s), \varphi^{m_0}(s)\Big) - F_k\Big(\kappa_i\Big(I^{m_0}(s), \varphi^{m_0}(s)\Big)\Big)\right| ds \\
&\leqslant \int_{a_i}^{a_{i+1}} C_{M_1}^{n_0,m_0}\sqrt{2}\left|I^{m_0}(s) - \kappa_i\Big(I^{m_0}(s)\Big)\right|^{1/2} ds \\
&\leqslant \epsilon\sqrt{2}T_0 C_{M_1}^{n_0,m_0}(\gamma + T_0 C_{M_1}\epsilon)^{1/2}.
\end{aligned} \tag{3.5.11}$$

By (3.5.5), we have

$$\int_0^\tau F_k(I(s),\varphi(s))ds = \int_0^\tau F_k(I^{m_0}(s), \varphi^{m_0}(s))ds + \xi_1(\tau), \tag{3.5.12}$$

where $|\xi_1(\tau)| \leqslant \rho\tau$.

Proposition 1.

$$\int_0^\tau F_k\Big(I^{m_0}(s), \varphi^{m_0}(s)\Big)ds = \sum_{i=0}^{d_0}\int_{a_i}^{a_{i+1}} F_k\Big(I^{m_0}(a_i), \varphi^{m_0}(s)\Big)ds + \xi_2(\tau),$$

where

$$|\xi_2| \leqslant \frac{1}{2}C_{M_1}^{n_0,m_0}\Big[4\sqrt{2}(\gamma + T_0 C_{M_1}\epsilon)^{1/2} + \gamma^{-1/2}T_0 C_{M_1}\epsilon\Big](\epsilon T_0 + 1). \tag{3.5.13}$$

Démonstration. We may write $\xi_2(\tau)$ as

$$\xi_2(\tau) = \sum_{i=0}^{d_0-1} \int_{a_i}^{a_{i+1}} \Big[F_k\Big(I^{m_0}(s), \varphi^{m_0}(s)\Big) - F_k\Big(I^{m_0}(a_i), \varphi^{m_0}(s)\Big)\Big]ds := \sum_{i=0}^{d_0-1} \tilde{I}_i.$$

For each i, by (3.5.4) and (3.5.7) we have

$$\begin{aligned}
&\int_{a_i}^{a_{i+1}} |F_k(\kappa_i(I^{m_0}(s)), \varphi^{m_0}(s)) - F_k(\kappa_i(I^{m_0}(a_i), \varphi^{m_0}(s))|ds \\
&\leqslant \int_{a_i}^{a_{i+1}} \gamma^{-1/2} C_{M_1}^{n_0,m_0} |\kappa_i(I^{m_0}(s) - I^{m_0}(a_i))|ds \\
&\leqslant \frac{1}{2} C_{M_1}^{n_0,m_0} C_{M_1} T_0^2 \gamma^{-1/2} \epsilon^2.
\end{aligned} \tag{3.5.14}$$

Replacing the integrand $F_k(I^{m_0}, \varphi^{m_0})$ by $F_k(\kappa_i(I^{m_0}, \varphi^{m_0}))$, using (3.5.11) and (3.5.14), we have

$$\tilde{I}_i \leqslant \frac{1}{2} C_{M_1}^{n_0,m_0} [4\sqrt{2}\epsilon T_0(\gamma + T_0 C_{M_1}\epsilon)^{1/2} + \gamma^{-1/2} T_0^2 C_{M_1}\epsilon^2].$$

The inequality (3.5.13) follows. □

On each subsegment $[a_i, a_{i+1}]$, we now consider the unperturbed linear dynamics $\tilde{\varphi}_i(\tau)$ of the angles $\varphi^{m_0} \in \mathbb{T}^{m_0}$:

$$\tilde{\varphi}_i(\tau) = \varphi^{m_0}(a_i) + \epsilon^{-1}\Lambda^{m_0}(\tau - a_i) \in \mathbb{T}^{m_0}, \quad \tau \in [a_i, a_{i+1}].$$

Proposition 2.

$$\int_0^\tau F_k\Big(I^{m_0}(a_i), \varphi^{m_0}(s)\Big)ds = \sum_{i=0}^{d_0-1} \int_{a_i}^{a_{i+1}} F_k\Big(I^{m_0}(a_i), \tilde{\varphi}_i(s)\Big)ds + \xi_3(\tau),$$

where

$$|\xi_3(\tau)| \leqslant [2\sqrt{2} C_{M_1}^{n_0,m_0} (\gamma + T_0 C_{M_1}\epsilon)^{1/2} + \frac{T_0\epsilon}{2\gamma} (C_{M_1}^{n_0,m_0})^2](1 + \epsilon T_0). \tag{3.5.15}$$

Démonstration. On each $[a_i, a_{i+1}]$, notice that

$$\begin{aligned}
&\int_{a_i}^{a_{i+1}} \Big|\kappa_i\Big(\varphi^{m_0}(s) - \tilde{\varphi}_i(s)\Big)\Big|ds \leqslant \int_{a_i}^{a_{i+1}} \int_{a_i}^{s} \Big|\kappa_i\Big(\epsilon G^{m_0}(I(s'), \varphi(s'))\Big)\Big|ds'ds \\
&\leqslant \int_{a_i}^{a_{i+1}} \int_{a_i}^{s} C_{M_1}^{n_0,m_0} \epsilon\gamma^{-1/2} ds'ds \leqslant \frac{T_0^2\epsilon^2}{2\gamma^{1/2}} C_{M_1}^{n_0,m_0}.
\end{aligned}$$

Here the first inequality comes from equation (3.3.4), and using (3.5.7) we can get the second inequality. Therefore, using again (3.5.7), we have

$$\begin{aligned}
&\int_{a_i}^{a_{i+1}} \Big[F_k\Big(\kappa_i\Big(I^{m_0}(a_i), \varphi^{m_0}(s)\Big)\Big) - F\Big(\kappa_i\Big(I^{m_0}(a_i), \tilde{\varphi}_i(s)\Big)\Big)\Big]ds \\
&\leqslant \int_{a_i}^{a_{i+1}} C_{M_1}^{n_0,m_0} \Big|\kappa_i\Big(\varphi^{m_0}(s) - \tilde{\varphi}_i(s)\Big)\Big|ds \\
&\leqslant \frac{T_0^2\epsilon^2}{2\gamma^{1/2}} (C_{M_1}^{n_0,m_0})^2
\end{aligned}$$

Therefore (3.5.15) holds for the same reason as (3.5.13). □

We will now compare the integrals $\int_{a_i}^{a_{i+1}} F_k(I^{m_0}(a_i), \tilde{\varphi}_i(s))ds$ with the average values $\langle F_k(I^{m_0}(a_i))\rangle \epsilon T_0$.

Propositon 3.

$$\sum_{i=0}^{d_0-1} \int_{a_i}^{a_{i+1}} F_k\Big(I^{m_0}(a_i), \tilde{\varphi}_i(s)\Big)ds = \sum_{i=1}^{d_0-1} T_0 \langle F_k\rangle \Big(I^{m_0}(a_i)\Big) + \xi_4(\tau),$$

where

$$|\xi_4(\tau)| \leqslant \rho + 2C_{M_1}\epsilon T_0. \tag{3.5.16}$$

Démonstration. For $0 \leqslant i \leqslant d_0 - 2$, by (3.5.6)

$$\Big|\int_{a_i}^{a_{i+1}} \Big[F_k\Big(I^{m_0}(a_i), \tilde{\varphi}_i(s)\Big) - \langle F_k\rangle\Big(I^{m_0}(a_i)\Big)\Big]ds\Big| \leqslant \epsilon\rho T_0.$$

So

$$\sum_{i=0}^{d_0-2} \Big|\int_{a_i}^{a_{i+1}} F_k\Big(I^{m_0}(a_i), \tilde{\varphi}_i(s)\Big)ds - \langle F_k\rangle\Big(I^{m_0}(a_i)\Big)T_0\Big| \leqslant (d_0-1)\epsilon\rho T_0.$$

Moreover,

$$\Big|\int_{a_{d_0-1}}^{\tau} \Big[F_k\Big(I^{m_0}(a_i), \tilde{\varphi}_i(s)\Big) - \langle F_k\rangle\Big(I^{m_0}(a_i)\Big)\Big]ds\Big| \leqslant 2C_{M_1}\epsilon T_0.$$

This implies the inequality (3.5.16). □

Proposition 4.

$$\sum_{i=1}^{d_0-1} (a_{i+1}-a_i)\langle F_k\rangle\Big(I^{m_0}(a_i)\Big) = \int_0^{\tau} \langle F_k\rangle\Big(I^{m_0}(s)\Big)ds + \xi_5(\tau),$$

where

$$|\xi_5(\tau)| \leqslant \epsilon C_{M_1} C_{M_1}^{n_0,m_0} T_0(\epsilon T_0 + 1). \tag{3.5.17}$$

Démonstration. Indeed, as

$$|\xi_5(\tau)| = \Big|\int_0^{\tau} \Big[\langle F_k\rangle\Big(I^{m_0}(s)\Big)ds - \sum_{i=1}^{d_0-1}(a_{i+1}-a_i)\langle F_k\rangle\Big(I^{m_0}(a_i)\Big)\Big|,$$

then using (3.5.4) and (3.5.8) we get

$$\begin{aligned}|\xi_5(\tau)| &\leqslant \sum_{i=0}^{d_0-1}\int_{s(i,j)} C_{M_1}^{n_0,m_0}|I^{m_0}(s) - I^{m_0}(a_i)|ds \\ &\leqslant \epsilon^2 \sum_{i=0}^{d_0-1} C_{M_1}C_{M_1}^{n_0,m_0}(T_0)^2 \leqslant \epsilon C_{M_1}C_{M_1}^{n_0,m_0}T_0(\epsilon T_0+1).\end{aligned}$$

□

Finally, we have obvious

Proposition 5.

$$\int_0^\tau \langle F_k\rangle\Big(I^{m_0}(s)\Big)ds = \int_0^\tau \langle F_k\rangle\Big(I(s)\Big)ds + \xi_6(\tau),$$

and $|\xi_6(\tau)|$ is bounded by $\rho\tau$.

Gathering the estimates in Propositions 1-5, we obtain

$$I_k(\tau) = I_k(0) + \int_0^\tau F_k\Big(I(s),\varphi(s)\Big)ds = I_k(0) + \int_0^\tau \langle F_k\rangle\Big(I(s)\Big)ds + \Xi(\tau),$$

where $|\Xi(\tau)| \leqslant \sum_{i=1}^6 |\xi_i(\tau)|$ satisfies (3.5.10). Lemma 4.1 is proved. □

Corollary 3.5.2. *For any $\bar\rho > 0$, with a suitable choice of ρ, γ and T_0, the function $|\Xi(\tau)|$ in Lemma 3.5.1 can be made less than $\bar\rho$, if ϵ is small enough.*

Démonstration. We choose $\gamma = \epsilon^\alpha, T_0 = \epsilon^{-\sigma}, \rho = \frac{\bar\rho}{9}$ with

$$1 - \alpha/2 - \sigma > 0,\ 0 < \sigma < 1.$$

Then for ϵ small enough, we have $|\Xi(\tau)| < \bar\rho$. □

For any $(I_0, \varphi_0) \in \Gamma_0$, let the curve $(I^\epsilon(\tau), \varphi^\epsilon(\tau)) \in h_I^p \times \mathbb{T}^\infty$, $\tau \in [0,1]$, be a regular solution of the equation (3.4.4) such that $(I^\epsilon(0), \varphi^\epsilon(0)) = (I_0, \varphi_0)$.

Lemma 3.5.3. *The family of curves $\{I^\epsilon(\tau), \tau \in [0,1]\}_{0<\epsilon<1}$ is pre-compact in $C([0,1], h_I^{p-2})$. Moreover every limiting (as $\epsilon \to 0$) curve $I^0(\tau)$, $\tau \in [0,1]$ is a solution of the averaged equation (3.4.5), satisfying*

$$|I^0(\tau)|_p^\sim \leqslant M_1, \quad \tau \in [0,1].$$

Démonstration. Due to (3.3.6) and (3.5.4), we know that for any $\epsilon \in (0,1)$,

$$|I^\epsilon(\tau)|_p^\sim \leqslant M_1, \quad |\frac{d}{d\tau} I(\tau)|_{p-2}^\sim \leqslant C_{M_1}, \quad \tau \in [0,1].$$

Then by the Arzelà-Ascoli theorem, we have that the set $\mathcal{I} := \{I^\epsilon(\tau), \tau \in [0,1]\}_{0<\epsilon<1}$ is pre-compact in $C([0,1], h_I^{p-2})$. Let $\{\rho_m\}_{m\in\mathbb{N}}$ be a sequence such that $\rho_m \searrow 0$. From Lemma 3.5.1 and Corollary 3.5.2, there is $\epsilon_m > 0$ such that if $\epsilon \leqslant \epsilon_m$, then for $k = 1, \dots, m$, we have

$$\begin{aligned} &I_k^\epsilon(\tau) = I_k^\epsilon(0) + \int_0^\tau \langle F_k\rangle(I^\epsilon(s))ds + \Xi_k(\tau), \\ &|\Xi_k(\tau)| \leqslant \rho_m, \quad \tau \in [0,1]. \end{aligned} \tag{3.5.18}$$

Let $I^0 = I^0(\tau)$, $\tau \in [0,1]$ be a limiting curve of the set $\mathcal{I}$ as $\epsilon \to 0$. Then we have

$$I^0 \in C([0,1], h_I^{p-2}) \quad \text{and} \quad |I^0(\tau)|_p^\sim \leqslant M_1, \quad \tau \in [0,1].$$

By (3.5.18), the curve $I^0(\cdot)$ solves the averaged equation (3.4.5). □

For any $\theta \in \mathbb{T}^\infty$ and any vector $I \in h^p_{I+}$ we set

$$V_\theta(I) = (V_{\theta_1}(I_1), V_{\theta_2}(I_2), \dots),$$

where $\theta = (\theta_1, \theta_2, \dots)$ and $V_{\theta_j}(I_j) = \sqrt{2I_j}\cos(\theta_j) + i\sqrt{2I_j}\sin(\theta_j)$, for every $j \geqslant 1$. Then $\varphi_j(V_{\theta_j}) \equiv \theta_j$, and for each $\theta \in \mathbb{T}^\infty$ the map $I \to V_\theta(I)$ is a right inverse of the map $v \to I(v)$. For any vector I we denote

$$I^{>N} = (I_{N+1}, I_{N+2}, \dots), \quad V_\theta^{>N}(I) = (V_{\theta_{N+1}}(I_{N+1}), V_{\theta_{N+2}}(I_{N+2}), \dots).$$

Lemma 3.5.4. *(Lifting) Let $I^0(\tau) = (I^0_k(\tau), k \geqslant 1) \in h^p_{I+}$, $\tau \in [0,1]$, be a solution of the averaged equation (3.4.5), constructed in Lemma 3.5.3. Then, for any $\theta \in \mathbb{T}^\infty$, there is a solution $v(\cdot)$ of the effective equation (3.4.8) such that*

$$I(v(\tau)) = I^0(\tau), \quad \tau \in [0,1], \quad \textit{and} \quad v(0) = V_\theta(I^0(0)). \tag{3.5.19}$$

Démonstration. [1] For any $m \in \mathbb{N}$, consider the non-autonomous finite dimensional systems

$$\dot{I}_k = \langle F_k\rangle\Big(I_1, \cdots, I_m, \Big(I^0(\tau)\Big)^{>m}\Big), \quad k = 1, \cdots, m, \tag{3.5.20}$$

$$\dot{v}_k = R_k\Big(v_1, \dots, v_m, V_\theta^{>m}(I^0(\tau))\Big), \quad k = 1, \dots, m. \tag{3.5.21}$$

Obviously, $(I^0_1(\tau), \dots, I^0_m(\tau))$, $\tau \in [0,1]$ solves system (3.5.20). It is its unique solution with initial data $(I^0_1(0), \dots, I^0_m(0))$, since by Lemma 3.4.1 the function $\langle F_k\rangle$ is smooth with respect to the variables $(I_1, \dots, I_m)$.

For $\bar{v}_0 = (V_{\theta_1}(I^0_1(0)), \dots, V_{\theta_m}(I^0_m(0)))$, system (3.5.21) has a unique solution $v^m(\tau)$, defined for $\tau \in [0, T')$, with $v^m(0) = \bar{v}_0$, where $T' \leqslant 1$ and $v^m(\tau) \xrightarrow{\tau \to T'} \infty$ if $T' < 1$. Due to equality (3.4.6), $I(v^m)(\tau)$ solves system (3.5.20) in time interval $[0, T')$. Since $I(v^m(0)) = (I^0_1(0), \cdots, I^0_m(0))$, therefore $T' = 1$ and

$$I(v^m(\tau)) \equiv (I^0_1(\tau), \dots, I^0_m(\tau)) \quad \text{for} \quad 0 \leqslant \tau \leqslant 1.$$

Now denote

$$V_m(\tau) = (v^m(\tau), V_\theta^{>m}(\tau)), \quad \tau \in [0,1].$$

For the same reason as in the proof of Lemma 3.5.3, the family $\{V_m(\tau), \tau \in [0,1]\}_{m\in\mathbb{N}}$ is pre-compact in $C([0,1], h^{p-2})$ and

$$V_m(0) = V_\theta(I^0(0)), \quad I(V_m(\tau)) = I^0(\tau), \quad \tau \in [0,1], \quad m \in \mathbb{N}.$$

So any limiting (as $m \to \infty$) curve $v(\cdot)$ of the family$\{V_m(\tau), \tau \in [0,1]\}_{m\in\mathbb{N}}$ is a solution of the effective equation (3.4.8), satisfying equalities (3.5.19). The lemma is proved. □

1. This argument is a simplified version of the proof of Theorem 3.1 in [49]

Lemma 3.5.5. *(uniqueness) Under the same assumptions of Lemma 3.5.3, we have* $I^0(\cdot) \in C([0,1], h_I^p)$ *and for every* $q < p$,

$$I^\epsilon(\cdot) \underset{\epsilon\to 0}{\longrightarrow} I^0(\cdot) \quad in \quad C([0,1], h_I^q). \tag{3.5.22}$$

Démonstration. Let $I^0(\cdot)$ and $J^0(\cdot)$ be two limiting curves of the family $\{I^\epsilon(\cdot)\}_{0<\epsilon<0}$, as $\epsilon \to 0$, in $C([0,1], h_I^{p-2})$. Then by Lemma 3.5.4, for any $\theta \in \mathbb{T}^\infty$, there are solutions $v_I(\cdot)$, $v_J(\cdot)$ of the effective equation (3.4.8) such that for $0 \leqslant \tau \leqslant 1$,

$$I(v_I(\tau)) = I^0(\tau), \quad I(v_J(\tau)) = J(\tau), \quad v_I(0) = v_J(0) = v_0 = V_\theta(I_0).$$

Due to assumption B, for initial data v_0 the effective equation (3.4.8) has a unique solution $v_E(\cdot) \in C\Big(\big[0, T(|v_0|_p)\big), h^p\Big)$. Therefore

$$v_I(\tau) = v_J(\tau) = v_E(\tau). \tag{3.5.23}$$

This relation holds for $\tau \leqslant 1$ if $T(|v_0|_p) > 1$ and for $\tau < T(|v_0|_p)$ if $T(|v_0|_p) < 1$. But if $T(|v_0|_p) < 1$, then $|v_E(\tau)|_p \to \infty$ as $\tau \to T(|v_0|_p)$. By the construction in Lemmata 3.5.3 and 3.5.4, we know $|v_I(\tau)|_p^2 \leqslant M_1$ for $\tau \in [0,1]$. Together with (3.5.23) we have that $T(|v_0|_p) > 1$. Hence $I^0 = J^0$, $I^0 \in C([0,1], h_I^p)$ and

$$I^\epsilon(\cdot) \underset{\epsilon\to 0}{\to} I^0(\cdot) \quad \text{in} \quad C([0,1], h_I^{p-2}). \tag{3.5.24}$$

For any $q < p$, assume that the convergence (3.5.22) do not holds. Then there exists $\delta > 0$ and sequences $\epsilon_n, \tau_n \in [0,1]$ such that

$$\epsilon_n \to 0 \quad \text{as} \quad n \to \infty \quad \text{and} \quad |I^{\epsilon_n}(\tau_n) - I^0(\tau_n)|_q^\sim \geqslant \delta. \tag{3.5.25}$$

Takes subsequence $\{n_k\}$ such that $\tau_{n_k} \to \tau_0$ as $n_k \to \infty$. Since the sequence $\{I^{\epsilon_{n_k}}(\tau_{n_k})\}$ is pre-compact in h_I^q, and by (3.5.24), its limiting point as $n_k \to \infty$ equals $I^0(\tau_0)$, so we have $I^{\epsilon_{n_k}}(\tau_{n_k})$ converges to $I^0(\tau_0)$ in h_I^q as n_k goes to ∞. This contradicts with (3.5.25). So we completes the proof of Lemma 3.5.5 and also the proof of Theorem 3.1.1. □

3.6 Application to complex Ginzburg-Landau equations

In this section we prove that assumptions A and B hold for equation (3.1.13), satisfying (3.1.14) and (3.1.15).

3.6.1 Verification of Assumption A

In this subsection, we denote by $|\cdot|_s$ the L^s-norm. Let $u(\tau)$ be a solution of equation (3.1.13) such that $u(0,x) = u_0$. Then

$$\begin{aligned}\frac{d}{d\tau}||u(\tau)||_0^2 &= 2\langle u, \dot{u}\rangle = 2\langle u, -\epsilon^{-1} i A_V u + \triangle u - \gamma_R |u|^{2p} u - i\gamma_I |u|^{2q} u\rangle,\\ &= -2||u||_1^2 + 2||u||_0^2 - 2\gamma_R |u|_{2p+2}^{2p+2}.\end{aligned}$$

Since $||u||_0^2 \leqslant |u|^2_{2p+2}$, then relation $||u(\tau_1)||_0 > \gamma_R^{-1/2p} = B_2$ implies that

$$\frac{d}{d\tau}||u(\tau_1)||_0^2 < 0.$$

So for any $T > 0$ we have

$$||u(T)||_0 \leqslant \min\{B_2,\ e^T||u_0||_0\}. \tag{3.6.1}$$

Now we rewrite equation (3.1.13) as follows :

$$\dot u + \epsilon^{-1} i(\triangle u + V(x)u + \epsilon\gamma_I |u|^{2q}u) = \triangle u - \gamma_R |u|^{2p}u. \tag{3.6.2}$$

For any $k \in \mathbb{N}$, denote

$$||u||_k^{\smile 2} = \langle A_V^k u, u\rangle, \quad A_V = -\triangle + V(x).$$

The l.h.s is a hamiltonian system with the hamiltonian function $\epsilon^{-1}H(u)$,

$$H(u) = \frac{1}{2}\langle A_V u, u\rangle + \frac{\epsilon}{2q+2}|u|^{2q+2}_{2q+2}.$$

We have $dH(u)(v) = \langle A_V u, v\rangle + \epsilon\gamma_I\langle |u|^{2q}u, v\rangle$, and if v is the vector field in the l.h.s of (3.6.2), then $dH(u)(v) = 0$. So we have

$$\begin{aligned}\frac{d}{d\tau}H(u(\tau)) &= -\gamma_R\langle A_V u, |u|^{2p}u\rangle + \langle A_V u, \triangle u\rangle \\ &\quad - \epsilon\gamma_I\gamma_R|u|^{2p+2q+2}_{2p+2q+2} + \epsilon\gamma_I\langle |u|^{2q}u, \triangle u\rangle,\end{aligned}$$

Denoting $U_q(x) = \frac{1}{q+1}u^{q+1}$ and $U_p = \frac{1}{p+1}u^{p+1}$, we get

$$\langle |u|^{2q}u, \triangle u\rangle \leqslant -\int_{\mathbb{T}^n}|\nabla u|^2|u|^{2q}dx = -||\nabla U_q||_0^2,$$

and a similar relation holds for q replaced by p. Therefore

$$\begin{aligned}\frac{d}{d\tau}H(u(\tau)) \leqslant& -\frac{1}{2}||u||_2^2 - \gamma_R||\nabla U_p||_0^2 - \epsilon\gamma_I||\nabla U_q||_0^2 - \epsilon\gamma_I\gamma_R|u|^{2p+2q+2}_{2p+2q+2} \\ & - \int_{\mathbb{T}^d}V(x)|\nabla u|^2dx + C_1||u||_0^2,\end{aligned}$$

where C_1 depends only on $|V|_{C^2}$. By this relation and (3.6.1), we have

$$H(u(T)) \leqslant H(u(0)) + C_1TB_2^2, \quad \text{for any} \quad T > 0. \tag{3.6.3}$$

So

$$||u(T)||_1^{\smile 2} \leqslant 2H(u(0)) + 2C_1TB_2^2, \quad \text{for any} \quad T > 0. \tag{3.6.4}$$

Simple calculation shows that

$$A_V^2 u = (-\triangle)^2 u - 2V\triangle u - \nabla V\cdot\nabla u + (V^2 - \triangle V)u.$$

We consider
$$\frac{d}{d\tau}\langle A_V^2 u, u\rangle = 2\langle A_V^2 u, \triangle u - \gamma_R |u|^{2p} u - i\gamma_I |u|^{2q} u.\rangle \tag{3.6.5}$$
By the interpolation and Young inequality, we have
$$\begin{aligned}\langle A_V^2 u, \triangle u\rangle &\leqslant -||u||_3^2 + C_1(|V|)||u||_2^2 + C_2(|V|_{C^1})||u||_1^2 + C_3(|V|_{C^2})||u||_0^2 \\ &\leqslant -||u||_3^2 + C_1||u||_3^{\frac{4}{3}}||u||_0^2 + C_2||u||_3^{\frac{2}{3}}||u||_0^2 + C_3||u||_0^2 \\ &\leqslant -\frac{3}{4}||u||_3^2 + C(|V|_{C^2}, ||u||_0).\end{aligned} \tag{3.6.6}$$

We deduce from integration by part and Hölder inequality that
$$-\langle(-\triangle)^2 u, |u|^{2p} u\rangle \leqslant ||u||_3 |(|u|^{2p}\nabla u)|_2 \leqslant ||u||_3 u|_{2pq_1}^{2q} |\nabla u|_{p_1}, \tag{3.6.7}$$
where $p_1, q_1 < \infty$ satisfy $1/p_1 + 1/q_1 = 1/2$. Let p_1 and q_1 have the form
$$p_1 = \frac{2d}{d-2s}, \quad q_1 = \frac{d}{s}.$$
We specify parameter s : For $d \geqslant 3$, choose $s = p(d-2) < \min\{d/2, 2\}$; for $d = 1, 2$, choose $s \in (0, \frac{1}{2})$. Due to condition (3.1.15), we have the Sobolev embeddings
$$H^s(\mathbb{T}^d) \to L^{p_1}(\mathbb{T}^d) \quad \text{and} \quad H^1(\mathbb{T}^d) \to L^{2pq_1}(\mathbb{T}^d),$$
implying that
$$|\nabla u|_{p_1} \leqslant ||u||_{1+s}, \quad |u|_{2pq_1}^{2p} \leqslant ||u||_1^{2p}.$$
Applying again the interpolation and Young inequality we find that for any $\delta > 0$,
$$\begin{aligned}-\langle \triangle^2 u, |u|^{2p} u\rangle &\leqslant ||u||_3 ||u||_{1+s} ||u||_1^{2p} \\ &\leqslant C||u||_3^{1+\frac{1+s}{3}} ||u||_0^{\frac{2-s}{3}} ||u||_1^{2p} \\ &\leqslant \delta ||u||_3^2 + C(\delta)(||u||_0^{\frac{2-s}{3}} ||u||_1^{2p})^{\frac{2-s}{6}},\end{aligned} \tag{3.6.8}$$
We can deal with other terms in (3.6.5) and (3.6.7) similarly. With suitable choice of δ, from the inequality above together with (3.6.6), we can get that for any $T > 0$
$$||u(T)||_2^{\sim 2} + \int_0^T ||u||_3^2 d\tau \leqslant ||u(0)||_2^{\sim 2} + C(2, |V|_{C^4}, T, B_2), \tag{3.6.9}$$
By similar argument, for any $m \geqslant 3$ and $T > 0$ we can obtain
$$||u(T)||_m^{\sim 2} + \int_0^T ||u||_{m+1}^2 d\tau \leqslant ||u(0)||_m^{\sim 2} + C(m, |V|_{C^{4m}}, T, B_2),$$
Then
$$||u(T)||_m \leqslant C(||u(0)||_m, |V|_{C^{4m}}, m, T, B_2), \quad \text{for any } T > 0.$$
This finishes the verification of assumption A.

3.6.2 Verification of Assumption B

We follow [50]. In equation (3.1.13) with $u \in H^2$, we pass to the v-variable, $v = \Psi(u) \in h^2$:

$$\dot{v}_k + i\epsilon^{-1}\lambda_k = P_k(v), \quad k \geqslant 1. \tag{3.6.10}$$

Here

$$P_k = P_k^1 + P_k^2 + P_k^3,$$

where P^1, P^2 and P^3 are, correspondingly, the linear, nonlinear dissipative and nonlinear hamiltonian parts of the perturbation :

$$P^1(v) = \Psi(\triangle u), \quad P^2(v) = -\gamma_R \Psi(|u|^{2p}u), \quad P^3(v) = -i\gamma_I \Psi(|u|^{2q}u),$$

with $u = \Psi^{-1}(v)$. Following the procedure in Section 3, the effective equations for (3.1.13) has the form :

$$\dot{v} = \sum_{i=1}^{3} R^i(v), \tag{3.6.11}$$

where

$$R^i(v) = \int_{\mathbb{T}^\infty} \Phi_{-\theta} P^i(\Phi_\theta)\Big) d\theta, \quad i = 1, 2, 3.$$

Consider the operator

$$\mathcal{L} := \Psi \circ (-\triangle) \circ \Psi^{-1} = \Psi \circ (A_V - V) \circ \Psi^{-1} := \hat{A} - \Psi \circ V \circ \Psi^{-1} := \hat{A} - \mathcal{L}^0.$$

Clearly, $\hat{A}$ is the diagonal operator $\hat{A} = \mathrm{diag}\{\lambda_j \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ j \geqslant 1\}$. By Lemma 1.1, $\mathcal{L}^0 = \Psi \circ V \circ \Psi^{-1}$ defines bounded maps

$$\mathcal{L}^0 : \quad h^m \to h^m, \quad \forall m \leqslant n,$$

and in the space h^0 the operator $\mathcal{L}^0$ is self-adjoint. Since $\hat{A}$ commutes with the rotation Φ_θ, then

$$\begin{aligned} R^1 &= -\int_{\mathbb{T}^\infty} \Phi_{-\theta}\hat{A}\Phi_\theta v d\theta + \int_{\mathbb{T}^\infty} \Phi_{-\theta}\mathcal{L}^0(\Phi_\theta v) d\theta \\ &= -\hat{A}v + R^0(v), \quad R^0(v) = \int_{\mathbb{T}^\infty} \Phi_{-\theta}\mathcal{L}^0(\Phi_\theta v) d\theta. \end{aligned} \tag{3.6.12}$$

Since for $v = (v_1, v_2, \dots)$, we have

$$\mathcal{L}^0(v)_j = \sum_{i=0}^{+\infty} \langle V(x) v_i \varphi_i(x), \varphi_j(x)\rangle, \quad j \geqslant 1,$$

then,

$$R_k^0(v) = \sum_{j=1}^{+\infty} \int_{\mathbb{T}^\infty} \langle V(x) v_j e^{i\theta_j}\varphi_j(x), e^{i\theta_k}\varphi_k(x)\rangle d\theta = v_k \langle V\varphi_k, \varphi_k\rangle.$$

That is,

$$R^1 = \mathrm{diag}\,\{-\lambda_k + M_k, k \geqslant 1\}, \quad M_k = \langle V\varphi_k, \varphi_k\rangle. \tag{3.6.13}$$

The term $R^2(v)$ is defined as an integral with the integrand

$$\Phi_{-\theta}P^2\Phi_\theta(v) = -\gamma_R\Phi_{-\theta}\Psi(f_p(|u|^2)u)|_{u=\Psi^{-1}\Phi_\theta v} := F_\theta(v).$$

Define $\mathcal{H}(u) = \int \mathcal{F}(|u|^2)dx$, where $\mathcal{F}' = \frac{1}{2}f_p$. Then $\nabla\mathcal{H}(u) = f_p(|u|^2)u$. Denoting $\Psi^{-1}\Phi_\theta = L_\theta$, we have

$$F_\theta(v) = -\gamma_R L_\theta^*\nabla\mathcal{H}(u)|_{u=L_\theta(v)} = -\gamma_R\nabla(\mathcal{H}\circ L_\theta(v)).$$

So

$$R^2(v) = -\gamma_R\nabla_v\Big(\int_{\mathbb{T}^\infty}(\mathcal{H}\circ\Psi^{-1})(\Psi_\theta v)d\theta\Big) = -\gamma_R\nabla_v\langle\mathcal{H}\circ\Psi^{-1}\rangle.$$

Similarly, we have $R^3 = -i\gamma_I\nabla_v\langle\mathcal{G}\circ\Psi^{-1}\rangle$ with $\nabla\mathcal{G}(u) = f_q(|u|^2)u$. Since $\langle\mathcal{G}\circ\Psi^{-1}\rangle$ is a function only of the action $(I_1,\dots)$, we have that $\nabla_{v_k}\langle\mathcal{G}\circ\Psi^{-1}\rangle$ is proportional to v_k. Then $v_k\cdot R_k^3(v) = 0$. That is, it contributes a zero term in the averaged equation. Hence we could set the effective equation to be

$$\dot{v} = R^1(v) + R^2(v).$$

It is a quasi-linear heat equation, written in the Fourier coefficients, which is known to be locally well posed. This verifies assumption B.

3.A Whitney's theorem

Consider the l_2-space of sequences $x = (x_1, x_2, \dots)$. The following lemma is a slight modification of the well known theorem of Whitney [73].
Lemma A. For any $n\in\mathbb{N}$, let $f\in C^\infty(l_2)$ be even in n variables, i.e.

$$f(x_1,\dots,x_i,\dots) = f(x_1,\dots,-x_i,\dots), \quad i = 1,2,\dots,n.$$

Then there exists $g_n\in C^\infty(l_2)$ such that

$$g_n(x_1^2,\dots,x_n^2,x_{n+1},\dots) = f(x_1,x_2,\dots).$$

Démonstration. For $n = 1$, we define $g_1(x_1,x_2,\dots) = f(x_1^{\frac{1}{2}},x_2,\dots)$. Since f is even with respect to x_1, for any $s\in\mathbb{N}$, we have

$$f(x_1,x_2,\dots) = f(\hat{x}) + f_1(\hat{x})x_1^2 + \dots + f_{s-1}(\hat{x})x_1^{2s-2} + \phi(x)x_1^{2s},$$

where $\hat{x} = (0,x_2,\dots)$, $f_i = [(2i)!]^{-1}\partial_{x_1}^{2i}f(\hat{x})$ and $\phi(x)$ is smooth when $x_1\neq 0$, even with respect to x_1, and satisfies

$$\lim_{x_1\to 0} x_1^k\partial_{x_1}^k\phi(x) = 0, \quad k = 1,\dots,2s. \tag{A.1}$$

Set $\psi(x) = \phi(x_1^{\frac{1}{2}},x_2,\dots)$, then

$$g_1(x) = f(\hat{x}) + f_1(\hat{x})x_1 + \dots + f_{s-1}(\hat{x})x_1^{s-1} + \psi(x)x_1^s.$$

We wish to check that $g_1(x)$ is C^s-smooth with respect to x_1. It is sufficient to prove that the limits $\lim_{x_1\to 0} x_1^k\partial_{x_1}^k\psi(x)$, $k=1,\dots,s$, exist and are finite. Differentiating $\psi(x_1^2,x_2,\dots)=\phi(x)$ with respect to x_1, we get that there are some constants a_{ki} such that

$$\partial_{x_1}^k\phi(x)=2^kx_1^k\partial_{x_1}^k\psi(x_1^2,x_2,\dots)+\sum_{1\leqslant i\leqslant k/2}a_{ki}x_1^{k-2i}\partial_{x_1}^{k-i}\psi(x_1^2,x_2,\dots),\quad k=1,\dots,s.$$

Solving these equation successively for $x_1^{2k}\partial_{x_1}^k\psi$, $k=1,\dots,s$, we obtain that there are some constant β_{ki} such that

$$x_1^{2k}\partial_{x_1}^k\psi(x_1^2,x_2,\dots)=\sum_{0\leqslant i\leqslant k}\beta_{ki}x_1^{k-i}\partial_{x_1}^{k-i}\phi(x).$$

By $(A.1)$, we know the $\lim_{x_1\to 0} x_1^k\partial_{x_1}^k\psi(x)$, $k=1,\dots,s$, exist and are finite. So $g_1(x)$ is C^s -smooth. Since s is arbitrary and $g_1(x)$ defined in a unique way, we have $g_1\in C^\infty(l^2)$ and $g_1(x_1^2,x_2,\dots)=f(x_1,x_2,\dots)$. This prove the statement of the lemma for $n=1$.

For $n\geqslant 2$, the assertion of the lemma can be prove by induction. Assume we have proved the lemma for $m=n-1$. Then there exists $g_{n-1}\in C^\infty(l_2)$ such that

$$g_{n-1}(x_1^2,\dots,x_{n-1}^2,x_n)=f(x_1,x_2,\dots)$$

and g_{n-1} is even in variable x_n. Applying what we have proved for $m=1$ to g_{n-1} with respect to x_n, we get the assertion for $m=n$. $\square$

Chapitre 4

An averaging theorem for nonlinear Schrödinger equations (resonant case)

Abstract : Consider a weakly nonlinear Schrödinger equation on the torus $\mathbb{T}^d$:

$$-iu_t + \triangle u = \pm\epsilon|u|^{2q}u. \tag{$*$}$$

Here $u = u(t,x)$, $x \in \mathbb{T}^d$, $0 < \epsilon << 1$ and $q \in \mathbb{N}$. Define $I(u) = (I_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d)$, where $I_{\mathbf{k}} = v_{\mathbf{k}}\bar{v}_{\mathbf{k}}/2$ and $v_{\mathbf{k}}$, $\mathbf{k} \in \mathbb{Z}^d$, are the Fourier coefficients of the function u we give. Assume that the equation $(*)$ is well posed on time intervals of order ϵ^{-1} and its solutions have there a-priori bounds, independent of the small parameter. Let $u(t,x)$ solves the equation $(*)$. If ϵ is small enough, then for $t \lesssim \epsilon^{-1}$, the quantity $I(u(t,x))$ can be well described by solutions of a *effective equation* :

$$-iu_t = \epsilon F(u),$$

where the term $F(u)$ is constructed through a resonant averaging of the nonlinearity $\pm|u|^{2q}u$.

4.1 Introduction

We consider a nonlinear Schrödinger equation on d-torus $\mathbb{T}^d = \mathbb{R}^d/2\pi\mathbb{Z}^d$:

$$-\, iu_t + \triangle u = \pm\epsilon|u|^{2q}u, \quad u = u(t,x),\ x \in \mathbb{T}^d, \tag{4.1.1}$$

where $q \in \mathbb{N} \cup \{0\}$ and ϵ is the small parameter. For any $p \in \mathbb{R}$ denote by H^p the Sobolev space of complex-valued periodic functions, provided with the norm $||\cdot||_p$,

$$||u||_p^2 = \Big\langle (-\triangle)^p u, u \Big\rangle + \langle u, u\rangle \quad \text{if} \quad p \in \mathbb{N}$$

where $\langle\cdot,\cdot\rangle$ is the real scalar product in $L^2(\mathbb{T}^d)$,

$$\langle u, v\rangle = Re\int_{\mathbb{T}^d} u\bar{v}dx, \quad u, v \in L^2(\mathbb{T}^d).$$

The equation (4.1.1) is hamiltonian and may be written as

$$-iu_t = \partial_u \mathcal{H}(u), \quad \mathcal{H}(u) = \int_{\mathbb{T}^d} \Big[\frac{1}{2}|\nabla u|^2 \pm \frac{\epsilon}{2q+2}|u|^{2q+2}\Big]dx,$$

where ∂_u stands for the L^2-gradient with respect to u. The sign of the nonlinearity (+ for the defocusing equation and - for the focusing equation) will not play an important role in this work due to the fact that here we mainly study the long-time dynamics of the equation (4.1.1) in the situation $0 < \epsilon \ll 1$. We assume :

Assumption A : *There exists some $p > d/2$ and $T > 0$ such that for every $u_0 \in H^p$, the equation (4.1.1) has a unique solution $u(t,x) \in H^p$ with initial datum u_0 and $||u(t,x)||_p \leqslant C(||u_0||_p, T)$ for $t \leqslant \epsilon^{-1}T$.*

Concerning this assumption the following is true :

Proposition 4.1.1. *(See [17, 13, 29]) The Assumption A holds for equation (4.1.1) if*

$$q \in \mathbb{N},\ q < +\infty,\ \textit{when}\ \ d = 1,2 \quad \textit{and} \quad q = 1,2,\ \textit{when}\ \ d = 3. \tag{4.1.2}$$

For this result see below Section 4.5.

We are mainly interested in the behaviours of solutions for the equation (4.1.1) in the time intervals of order ϵ^{-1}. So it would be convenient to use the slow time $\tau = \epsilon t$. Passing to the slow time τ, we get the rescale equation

$$-i\dot{u} + \epsilon^{-1}\triangle u = \pm|u|^{2q}u, \tag{4.1.3}$$

where $u = u(\tau,x)$, $x \in \mathbb{T}^d$ and the dot $\dot{}$ stands for $\frac{d}{d\tau}$.

For a complex function $u(x)$ on $\mathbb{T}^d$ we define

$$\mathcal{F}(u) = (v_{\mathbf{k}},\ \mathbf{k} \in \mathbb{Z}^d),$$

where the vector $(v_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d)$ is formed by the Fourier coefficients of u :

$$u(x) = \sum_{\mathbf{k}\in\mathbb{Z}^d} v_{\mathbf{k}} e^{i\mathbf{k}\cdot x}, \quad v_{\mathbf{k}} = \int_{\mathbb{T}^d} u(x)e^{-i\mathbf{k}\cdot x}dx.$$

In the space of complex sequence $v = (v_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d)$, we introduce the norm :

$$|v|_p^2 = \sum_{\mathbf{k}\in\mathbb{Z}^d} (|\mathbf{k}|^{2p} \vee 1)|v_{\mathbf{k}}|^2, \quad p \in \mathbb{R},$$

and denote $h^p = \{v :\ |v|_p < \infty\}$. Obviously, for $p \geqslant 0$, $h^p = \mathcal{F}(H^p)$.

The equation (4.1.3) has a rather transparent form in the space h^p. Let $u(\tau,x)$ be its solutions, then the Fourier coefficients $v_{\mathbf{k}}(\tau)$ of $u(\tau,x)$ solves the infinite dimensional ODE :

$$\dot{v}_{\mathbf{k}} - \epsilon^{-1}i\lambda_{\mathbf{k}}v_{\mathbf{k}} = \pm i \sum_{(\mathbf{k}_1,\dots,\mathbf{k}_{2q+1})\in\mathcal{S}(\mathbf{k})} v_{\mathbf{k}_1}\bar{v}_{\mathbf{k}_2}\cdots v_{\mathbf{k}_{2q-1}}\bar{v}_{\mathbf{k}_{2q}}v_{\mathbf{k}_{2q+1}} := P_{\mathbf{k}}(v), \tag{4.1.4}$$

where $\mathbf{k} \in \mathbb{Z}^d$, $\lambda_{\mathbf{k}} = |\mathbf{k}|^2$ and

$$\mathcal{S}(\mathbf{k}) = \{(\mathbf{k}_1, \ldots, \mathbf{k}_{2q+1}) \in (\mathbb{Z}^d)^{2q+1} : \sum_{j=1}^{2q+1} (-1)^{j-1}\mathbf{k}_j = \mathbf{k}\}.$$

Denote $\Lambda = (\lambda_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d)$ and we call it the *frequency vector* of the equation (4.1.3).

For every $\mathbf{k} \in \mathbb{Z}^d$, denote $I_{\mathbf{k}} = \frac{1}{2}v_{\mathbf{k}}\bar{v}_{\mathbf{k}}$ and $\varphi_{\mathbf{k}} = \mathrm{Arg}\, v_{\mathbf{k}}$. Notice that the quantities $I_{\mathbf{k}}$ are conservation laws of the linear equation $(4.1.1)_{\epsilon=0}$. We call them the action variables (correspondingly, call the quantities $\varphi_{\mathbf{k}}$ the angle variables). We introduce the weighted l^1-space h_I^p :

$$h_I^p := \{I = (I_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d) \in \mathbb{R}^\infty : |I|_p^\sim = \sum_{\mathbf{k}\in\mathbb{Z}^d} 2(|\mathbf{k}|^{2p} \vee 1)|I_{\mathbf{k}}| < \infty\}.$$

Using the action-angle variables (I, φ), we can write equation (4.1.4) as a slow-fast system :

$$\dot{I}_{\mathbf{k}} = v_{\mathbf{k}} \cdot P_{\mathbf{k}}(v), \quad \dot{\varphi}_{\mathbf{k}} = \epsilon^{-1}\lambda_{\mathbf{k}} + |v_{\mathbf{k}}|^{-2} \cdots, \quad \mathbf{k} \in \mathbb{Z}^d. \tag{4.1.5}$$

Here the dots stand for a term of order 1 (as $\epsilon \to 0$). Our task is to study the evolution of quantities $I_{\mathbf{k}}$. Following the averaging theory for PDEs (see, e.g. [53, 49, 31, 32]), we consider the averaged system

$$\dot{I}_{\mathbf{k}} = \langle v_{\mathbf{k}} \cdot P_{\mathbf{k}}(v)\rangle_\Lambda, \quad \mathbf{k} \in \mathbb{Z}^d. \tag{4.1.6}$$

Here $\langle\cdot\rangle_\Lambda$ signifies some kind of averaging (related to the frequency vector Λ) in the angles $\varphi \in \mathbb{T}^\infty$. The hope is that the averaged equation (4.1.6) may approximately describe the behaviour of the action variables $I_{\mathbf{k}}$ of equation (4.1.3). However due to the singular nature of the action-angle variables (I, φ) and the resonance of the frequency Λ, the equations (4.1.6) may have singularities at the set $\{I : I_{\mathbf{k}} = 0 \text{ for some } \mathbf{k}\}$ which is dense in the wighted l^1-space h_I^p. Moreover, the vector field in the averaged equation (4.1.6) may not be Lipschitz in the variables I, so its well-posedness is unclear. A way to overcome these difficulties was introduced in [49] by S. Kuksin. Namely, there exists a regular system

$$\dot{v}_{\mathbf{k}} = R_{\mathbf{k}}(v), \quad \mathbf{k} \in \mathbb{Z}^d, \tag{4.1.7}$$

where $R_{\mathbf{k}}(v)$ is defined through a certain averaging of the term $P_{\mathbf{k}}(v)$, such that under the mapping $v_{\mathbf{k}} \mapsto I_{\mathbf{k}} = \frac{1}{2}v_{\mathbf{k}}\bar{v}_{\mathbf{k}}$, solutions of equation (4.1.7) transform to solutions of the averaged equation (4.1.6). This method also was used by the author in [32] to establish an averaging theorem for NLS under some non-resonance conditions (see Chapter 3). The system (4.1.7) is called the *effective equation*.

For the equation (4.1.3), due to the polynomial form of the nonlinearity, there exists another way to derive the effective equation. That is to use the so-called *interaction representation picture*. Let us define

$$a_{\mathbf{k}}(\tau) = e^{-i\epsilon^{-1}\lambda_{\mathbf{k}}\tau}v_{\mathbf{k}}(\tau).$$

Clearly, $|a_{\mathbf{k}}|^2 = |v_{\mathbf{k}}|^2 = I_{\mathbf{k}}/2$. Therefore the limiting behaviour (as $\epsilon \to 0$) of the quantity $|a_{\mathbf{k}}|$ characterizes the limiting behaviour of the action variables $I_{\mathbf{k}}$. Using equation (4.1.4), we obtain the equation satisfied by $a_{\mathbf{k}}(\tau)$:

$$-i\dot{a}_{\mathbf{k}}(\tau) = \pm \sum_{(\mathbf{k}_1,\ldots,\mathbf{k}_{2q+1})\in S(\mathbf{k})} a_{\mathbf{k}_1}(\tau)\overline{a_{\mathbf{k}_2}(\tau)}\cdots a_{\mathbf{k}_{2q-1}}(\tau)\overline{a_{\mathbf{k}_{2q}}(\tau)}a_{\mathbf{k}_{2q+1}}(\tau)$$
$$\times \exp\{i\epsilon^{-1}\tau[-\lambda_{\mathbf{k}} + \sum_{j=1}^{2q+1}(-1)^{j-1}\lambda_{\mathbf{k}_j}]\},$$

where $\mathbf{k} \in \mathbb{Z}^d$. The terms in the right hand side oscillate very fast if ϵ is very small, except the terms that the sum in the exponential equals zero. This leads to the guess that only these terms determine the limiting behavior of $a_{\mathbf{k}}(\tau)$ as $\epsilon \to 0$, and that the effective equation is the following :

$$-i\dot{a}_{\mathbf{k}}(\tau) = \pm \sum_{(\mathbf{k}_1,\ldots,\mathbf{k}_{2q+1})\in \mathcal{R}(\mathbf{k})} a_{\mathbf{k}_1}(\tau)\overline{a_{\mathbf{k}_2}(\tau)}\cdots a_{\mathbf{k}_{2q-1}}(\tau)\overline{a_{\mathbf{k}_{2q}}(\tau)}a_{\mathbf{k}_{2q+1}}(\tau), \tag{4.1.8}$$

where $\mathcal{R}(\mathbf{k}) := \{(\mathbf{k}_1,\ldots,\mathbf{k}_{2q+1}) \in S(\mathbf{k}) : \ -\lambda_{\mathbf{k}} + \sum_{j=1}^{2q+1}(-1)^{j-1}\lambda_{\mathbf{k}_j} = 0\ \}$.

Let us denote

$$\mathcal{RES} := \{(\mathbf{k}_1,\ldots,\mathbf{k}_{2q+2}) \in (\mathbb{Z}^d)^{2q+2} : \sum_{j=1}^{2q+2}(-1)^{j-1}\lambda_{\mathbf{k}_j} = 0\}.$$

Then the equation (4.1.8) is hamiltonian with Hamiltonian function :

$$\mathcal{H}_{res}(v) = \pm\frac{1}{2q+2}\sum_{(\mathbf{k}_1,\ldots,\mathbf{k}_{2q+2})\in\mathcal{RES}} v_{\mathbf{k}_1}\bar{v}_{\mathbf{k}_2}\cdots v_{\mathbf{k}_{2q+1}}\bar{v}_{\mathbf{k}_{2q+2}},$$

We will see in Section 4.2 that the effective equation (4.1.7) for the equation (4.1.3) defined through a resonant averaging process is exactly the equation (4.1.8). It is well posed in the space h^p, $p > d/2$. Besides its Hamiltonian $\mathcal{H}_{res}$, possess two extra integrals :

$$H_1(v) = \sum_{\mathbf{k}\in\mathbb{Z}^d}|v_{\mathbf{k}}|^2, \quad H_2(v) = \sum_{\mathbf{k}\in\mathbb{Z}^d}\lambda_{\mathbf{k}}|v_{\mathbf{k}}|^2.$$

The main result of this work is the following theorem where $u(t,x)$ is a solution of the equation (4.1.1), $v(\tau) = \mathcal{F}(u(\epsilon^{-1}\tau, x))$ and $a'(\tau)$ is a solution of the effective equation (4.1.8).

Theorem 4.1.2. *If Assumption A holds and $|v(0) - a'(0)|_p \leqslant \epsilon^{1/2}$, then the solution $a'(\tau)$ of equation (4.1.8) exists for $0 \leqslant \tau \leqslant T$ and for sufficiently small parameter ϵ, we have*

$$|I(v(\tau)) - I(a'(\tau))|_p^{\sim} \leqslant C\epsilon^{1/2}, \quad \tau \in [0,T],$$

where the constant C depend only on T and the size of the initial datum $|v(0)|_p$.

Remark 4.1.3. *1) In the case that the H^p-norm of the solution $u(x,t)$ for equation (4.1.1) grows as $||u(x,t)||_p \lesssim e^{\epsilon t C(||u_0||_p)}$, the Theorem 4.1.2 can be extended to time intervals of order $\epsilon^{-1}\log\epsilon^{-1}$ with the exponent $1/2$ replaced by certain $\alpha > 0$.*

2) The method of this paper also applies to nonlinear Schrödinger equations with other polynomial nonlinearities, e.g. with the nonlinearities with Hamiltonians of the forms $\mathcal{H}_3' = \int(u^3 + \bar{u}^3)dx$ and $\mathcal{H}_3 = \int|u|^2(u+\bar{u})dx$.

Equations that are similar to the effective equation (4.1.8) recently appear in a number of works. E.g. a stochastic damp-driven version of it is constructed in [51], using the same philosophy of the present paper. In [28], similar equation is used to determine an effective integrable equation for a 1D wave equation. In [24], a 2D version of equation (4.1.8) is evoked as an intermediate equation for understanding the *Large box limit* of the cubic NLS on $\mathbb{T}^2$. The effective equation (4.1.8) also is known in the theories of *wave turbulence*. There, it is called the *equation of discrete turbulence*, see [65], Chapter 12. In this work, the equation (4.1.8) is deduced in the spirit of the averaging theory for PDEs. We believe that our result provides an useful insight on the relevant topics.

The Chapter is planed as follows : In Section 4.2, we introduce the concept of resonant averaging in the Hilbert space. We deduce the effective equation through the resonant averaging process in Section 4.3. The Section 4.4 is devoted to the proof of Theorem 4.1.2. Finally, in Section 4.5, we discuss the validity of Proposition 4.1.1.

4.2 Resonant averaging in the Hilbert space

We first introduce the resonant averaging of smooth functions in finite dimensional space. Let $W \in \mathbb{Z}^n$, $n \geqslant 1$ be a non-zero integer vector. We call the ensemble

$$\mathcal{A}(W) := \{s \in \mathbb{Z}^n : W \cdot s = 0\},$$

the *set of resonance for* W. Notice that if $s \in \mathbb{Z}^n \setminus \mathcal{A}(W)$, then $|W \cdot s| \geqslant 1$. For a continuous function f on $\mathbb{T}^n$, we define its *resonant average* with respect to the integer vector W as the function

$$\langle f \rangle_W(\varphi) = \frac{1}{2\pi} \int_0^{2\pi} f(\varphi + tW)dt. \tag{4.2.1}$$

Lemma 4.2.1. *Let f be a C^∞-function on $\mathbb{T}^n$ and $f = \sum f_s e^{is\cdot\varphi}$. Then*

$$\langle f \rangle_W(\varphi) = \sum_{s \in \mathcal{A}(W)} f_s e^{is\cdot\varphi}. \tag{4.2.2}$$

Now we pass to the corresponding definition in the Hilbert space h^p.

For infinite integer vectors $s = (s_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d) \in \mathbb{Z}^\infty$ we will write the l_1-norm of s as $|s|$,

$$|s| = \sum_{\mathbf{k} \in \mathbb{Z}^d} |s_{\mathbf{k}}|.$$

Let $\mathbb{Z}_0^\infty = \{s \in \mathbb{Z}^\infty : |s| < \infty\}$. Obviously, for each $s = (s_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d) \in \mathbb{Z}_0^\infty$, only finite many $s_{\mathbf{k}}$ are not zero. Fix some $m \in \mathbb{N}$ and define the set of resonant of order m for an integer vector $\Omega = (\omega_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d) \in \mathbb{Z}^\infty$ as

$$\mathcal{A}(\Omega, m) = \{s = (s_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d) \in \mathbb{Z}_0^\infty : \sum_{\mathbf{k} \in \mathbb{Z}^d} \omega_{\mathbf{k}} s_{\mathbf{k}} = 0,\ |s| \leqslant m\}.$$

For any $s \in \mathbb{Z}_0^\infty$ and $v = (v_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d) \in h^p$, $p \geqslant 0$, we denote $v^s = \prod_{\mathbf{k} \in \mathbb{Z}^d} \tilde{v}_j^{|s_{\mathbf{k}}|}$, where $\tilde{v}_{\mathbf{k}} = v_{\mathbf{k}}$ if $s_{\mathbf{k}} \geqslant 0$ and $\tilde{v}_{\mathbf{k}} = \bar{v}_{\mathbf{k}}$ if $s_{\mathbf{k}} < 0$. Consider a series $F(v)$ on h^p,

$$F(v) = \sum_{s \in \mathbb{Z}_0^\infty} C_s v^s.$$

We assume the series converges normally in h^p in the sense that for each $R > 0$ we have

$$\sup_{|v|_p \leqslant R} \sum_{s \in \mathbb{Z}_0^\infty} |C_s||v^s| < \infty. \tag{4.2.3}$$

To define the resonant averaging of $F(v)$ we introduce for each $\theta = (\theta_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d) \in \mathbb{T}^\infty$, the rotation operator Φ_θ, which is a linear operator in h^p :

$$\Phi_\theta(v) = v', \quad v'_{\mathbf{k}} = e^{i\theta_{\mathbf{k}}} v_{\mathbf{k}}.$$

This is a unitary isomorphism of each space h^p. Note that

$$(I, \varphi)(\Phi_\theta v) = (I(v), \varphi(v) + \theta).$$

For any integer vector Ω we define the resonant average of the function $F(v)$ by analogy to definition (4.2.1).

Definition 4.2.2. *Let a function $F(v) \in C(h^p)$ be given by a normally converging series. Then its resonant average with respect to Ω is the function*

$$\langle F \rangle_\Omega(v) = \frac{1}{2\pi} \int_0^{2\pi} F(\Phi_{t\Omega}) dt.$$

Defining a function $F'(I, \varphi)$ by the relation $F(v) = F'(I(v), \varphi(v))$, we see that

$$\langle F \rangle_\Omega(v) = \frac{1}{2\pi} \int_0^{2\pi} F'(I, \varphi + t\Omega) dt.$$

So this definition agree with its finite dimensional counter part. If the series $F(v)$ is *of order* $m \leqslant \infty$ in the sense that $C_s = 0$ unless $|s| \leqslant m$, then

$$\langle F \rangle_\Omega = \sum_{s \in \mathcal{A}(\Omega, m)} C_s v^s. \tag{4.2.4}$$

If the series $F(v)$ is normally converging, so does the series $\langle F \rangle_\Omega(v)$.

4.3 The Effective equation

In this section we will deduce the effective equation for equation (4.1.3) through a resonant averaging process.

Consider the Fourier transform for complex functions on $\mathbb{T}^d$ which we write as the mapping

$$\mathcal{F} : H^p \ni u(x) \mapsto v = (v_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d) \in \mathbb{C}^\infty,$$

defined by the relation $u(x) = \sum_{\mathbf{k} \in \mathbb{Z}^d} v_{\mathbf{k}} e^{i\mathbf{k}x}$. Then $|\mathcal{F}u|_p = ||u||_p$, for every $p \geqslant 0$. For each $\mathbf{k} \in \mathbb{Z}^d$, denote

$$I_{\mathbf{k}} = I(v_{\mathbf{k}}) = \frac{1}{2} v_{\mathbf{k}} \bar{v}_{\mathbf{k}} \quad \text{and} \quad \varphi_{\mathbf{k}} = \varphi(v_{\mathbf{k}}),$$

where $\varphi(v_{\mathbf{k}}) = \operatorname{Arg} v_{\mathbf{k}} \in \mathbb{S}^1$ if $v_{\mathbf{k}} \neq 0$ and $\varphi(0) = 0 \in \mathbb{S}^1$. Let us write equation (4.1.3) in the v-variables :

$$\dot{v}_{\mathbf{k}} - \epsilon^{-1} i\lambda_{\mathbf{k}} v_{\mathbf{k}} = P_{\mathbf{k}}(v), \quad \mathbf{k} \in \mathbb{Z}^d. \tag{4.3.1}$$

Here $P_{\mathbf{k}}$ is the coordinate component of the mapping $P(v)$ defined by

$$P(v) = \mathcal{F}(\pm i|u|^{2q}u), \quad u = \mathcal{F}^{-1}(v).$$

We have $P_k(v) = \sum_{s \in \mathbb{Z}_0^\infty} C_s v^s$, where $C_s = 0$ if $|s| \neq 2q+1$. It is of order $2q+1$. The mapping is analytic of polynomial growth :

Lemma 4.3.1. *The mapping $P(v)$ is an analytic transform of the space h^p with $p > d/2$. Moreover the norm of $P(v)$ and its derivative $dP(v)$ have polynomial growth with respect to $|v|_p$.*

Now we write equation for the quantities $I_{\mathbf{k}}$, $\mathbf{k} \in \mathbb{Z}^d$:

$$\dot{I}_{\mathbf{k}} = v_{\mathbf{k}} \cdot P_{\mathbf{k}}(v), \quad \mathbf{k} \in \mathbb{Z}^d. \tag{4.3.2}$$

We consider the following resonant averaged system for equation (4.3.2) :

$$\dot{I}_{\mathbf{k}}(\tau) = \Big\langle v_{\mathbf{k}}(\tau) \cdot P_{\mathbf{k}}(v(\tau)) \Big\rangle_{\Lambda}, \quad \mathbf{k} \in \mathbb{Z}^d. \tag{4.3.3}$$

However, as have mentioned in the introduction, the vector field on the right hand side of the equation (4.3.3) may have singularities on the dense subset

$$\{I = (I_{\mathbf{k}}, \mathbf{k} \in \mathbb{Z}^d) \in h_I^p : I_{\mathbf{k}} = 0 \text{ for some } \mathbf{k} \in \mathbb{Z}^d\}.$$

One efficient way to overcome this obstacle is to introduce a regular effective equation. Notice that

$$\begin{aligned} \Big\langle v_{\mathbf{k}} \cdot P_{\mathbf{k}}(v) \Big\rangle_{\Lambda} &= \frac{1}{2\pi} \int_0^{2\pi} e^{i\lambda_{\mathbf{k}} t} v_{\mathbf{k}} \cdot P_{\mathbf{k}}(\Phi_{t\Lambda} v) dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} v_{\mathbf{k}} \cdot e^{-i\lambda_{\mathbf{k}} t} P_{\mathbf{k}}(\Phi_{t\Lambda} v) dt = v_{\mathbf{k}} \cdot R_{\mathbf{k}}(v), \end{aligned} \tag{4.3.4}$$

where $R_{\mathbf{k}}(v) = \frac{1}{2\pi} \int_0^{2\pi} e^{-i\lambda_{\mathbf{k}} t} P_{\mathbf{k}}(\Phi_{t\Lambda}(v)) dt$.

Let $R(v) = (R_{\mathbf{k}}(v),\ \mathbf{k} \in \mathbb{Z}^d)$, then

$$R(v) = \frac{1}{2\pi} \int_0^{2\pi} \Phi_{-t\Lambda} P(\Phi_{t\Lambda} v) dt. \tag{4.3.5}$$

Lemma 4.3.2. *The vector field $R(v)$ is locally Lipschitz in the Hilbert space h^p, $p > d/2$.*

Démonstration. Let $v_1, v_2 \in h^p$ and $|v_1|_p$, $|v_2|_p \leqslant M$. Then using Lemma 4.3.1 and the fact that the operators $\Phi_{t\Lambda}$, $t \in \mathbb{R}$ define isometries in h^p, we have

$$\begin{aligned} |R(v_1) - R(v_2)|_p &\leqslant \frac{1}{2\pi} \int_0^{2\pi} |\Phi_{-t\Lambda}[P(\Phi_{t\Lambda} v_1) - P(\Phi_{t\Lambda} v_2)]|_p dt \\ &\leqslant \frac{1}{2\pi} \int_0^{2\pi} C(M) |\Phi_{t\Lambda}(v_1 - v_2)|_p dt \leqslant C(M) |v_1 - v_2|_p. \end{aligned}$$

□

Consider the following equation

$$\dot{v} = R(v). \tag{4.3.6}$$

By Lemma 4.3.2 we know that this equation is well posed, at least locally, in the space h^p, $p > d/2$. From the relation 4.3.4, we have that if $v(\tau)$ solves the equation (4.3.6), then $I(v(\tau))$ satisfies the relations (4.3.3). We call equation (4.3.6) the *effective equation* for equation (4.1.3).

Proposition 4.3.3. *1) The effective equation (4.3.6) is a hamiltonian equation with the Hamiltonian function $\mathcal{H}_{res}(v) = \langle \mathcal{H}(\mathcal{F}^{-1}(v))\rangle_\Lambda$.*

2) The Hamiltonian $\mathcal{H}_{res}$ is invariant under the operators $\Phi_{t\Lambda}$ and $\Phi_{t\mathbf{1}}$, where $t \in \mathbb{R}$ and $\mathbf{1} = (1, 1, 1, \dots)$.

Démonstration. Since $P(v) = i\nabla\mathcal{H}(\mathcal{F}^{-1}(v))$, then using $\Phi^*_{t\Lambda} = \Phi_{-t\Lambda}$, we have

$$\begin{aligned}
R(v) &= \frac{1}{2\pi}\int_0^{2\pi} \Phi_{-t\Lambda}\Big(i\nabla\mathcal{H}(\mathcal{F}^{-1}(\Phi_{t\Lambda}v))\Big)dt = i\frac{1}{2\pi}\nabla\int_0^{2\pi}\mathcal{H}(\mathcal{F}^{-1}(\Phi_{t\Lambda}v))dt \\
&= i\nabla\mathcal{H}_{res}(v).
\end{aligned}$$

This prove the assertion 1). The assertion 2) follows from the assertion 1) and the Definition 4.2.2. □

Let give an explicit formula for the quantity $R_{\mathbf{k}}(v)$, $\mathbf{k} \in \mathbb{Z}^d$. Since

$$P_{\mathbf{k}}(v) = \pm i \sum_{(\mathbf{k}_1,\dots,\mathbf{k}_{2q+1})\in\mathcal{S}(\mathbf{k})} v_{\mathbf{k}_1}\bar{v}_{\mathbf{k}_2}\cdots v_{\mathbf{k}_{2q-1}}\bar{v}_{\mathbf{k}_{2q}}v_{\mathbf{k}_{2q+1}},$$

then

$$\begin{aligned}
R_{\mathbf{k}}(v) &= \frac{1}{2\pi}\int_0^{2\pi} e^{-i\lambda_{\mathbf{k}}t}P_{\mathbf{k}}(\Phi_{t\Lambda}(v))dt \\
&= \frac{\pm i}{2\pi}\int_0^{2\pi} \sum_{(\mathbf{k}_1,\dots,\mathbf{k}_{2q+1})\in\mathcal{S}(\mathbf{k})} \Big[v_{\mathbf{k}_1}\bar{v}_{\mathbf{k}_2}\cdots v_{\mathbf{k}_{2q-1}}\bar{v}_{\mathbf{k}_{2q}}v_{\mathbf{k}_{2q+1}} \\
&\qquad\qquad \times \exp[-it(-\lambda_{\mathbf{k}} + \sum_{j=1}^{2q+1}(-1)^{j-1}\lambda_{\mathbf{k}_j})\Big]dt \\
&= \pm i \sum_{(\mathbf{k}_1,\dots,\mathbf{k}_{2q+1})\in\mathcal{R}(\mathbf{k})} v_{\mathbf{k}_1}\bar{v}_{\mathbf{k}_2}\cdots v_{\mathbf{k}_{2q-1}}\bar{v}_{\mathbf{k}_{2q}}v_{\mathbf{k}_{2q+1}}.
\end{aligned}$$

Therefore the effective equation (4.3.6) is exact by the same form as the equation (4.1.8).

4.4 The averaging theorem

In this section we will prove the Theorem 4.1.2 by using an averaging process. We denote

$$B(M) = \{v \in h^p : |v|_p \leqslant M\}, \quad \forall M > 0.$$

Fix a $M_0 > 0$. Let $u(\tau, x)$ be a solution of equation (4.1.3) such that

$$||u(0,x)||_p \leqslant M_0,$$

and $v(\tau) = \mathcal{F}(u(\tau, x))$. Without loss of generality, suppose the Assumption A hold with $T = 1$. Then we have that there exists $M_1 \geqslant M_0$ such that

$$v(\tau) \in B(M_1), \quad \tau \in [0,1].$$

Let $a(\tau) = \Phi_{-\tau\epsilon^{-1}\Lambda}(v(\tau))$. Then $a(\tau)$ is the interaction representation picture of $v(\tau)$. We have

$$\dot{a}(\tau) = \Phi_{-\tau\epsilon^{-1}\Lambda}\Big(P\Big(\Phi_{\tau\epsilon^{-1}\Lambda}(a(\tau))\Big)\Big) := Y\Big(a(\tau), \tau\Big). \tag{4.4.1}$$

Using Lemma 4.3.1 and the fact the the operators $\Phi_{t\Lambda}$, $t \in \mathbb{R}$ define isometries on h^p, we have for any $v,\ v' \in B(M_1)$ and $\tau \in \mathbb{R}$,

$$|Y(v,\tau)|_p \leqslant C(M_1), \quad |Y(v,\tau) - Y(v',\tau)|_p \leqslant C(M_1)|v - v'|_p. \tag{4.4.2}$$

Denote $\mathcal{Y}(v,\tau) = Y(v,\tau) - R(v)$. Then by Lemma 4.3.2, the relation (4.4.2) also holds for the map $\mathcal{Y}(v,\tau)$.

The following lemma is the main step of our proof.

Lemma 4.4.1. *For $\tau \in [0,1]$,*

$$|a(\tau) - a(0) - \int_0^\tau R(a(s))ds|_p \leqslant C(M_1)\epsilon^{1/2}.$$

Démonstration. Denote by $Y_{\mathbf{k}}(v,\tau)$, $\mathbf{k} \in \mathbb{Z}^d$ the coordinate components of the map $Y(v,\tau)$, similarly, $\mathcal{Y}_{\mathbf{k}}$ of $\mathcal{Y}$.

We first fix some $T_0 \in [0,1]$ and divide the time interval $[0,1]$ into subintervals $[b_l, b_{l-1}]$, $l = 1, \cdots, m$ such that :

$$b_0 = 0, b_l - b_{l-1} = T_0, \quad \text{for} \quad l = 1, \ldots, m-1, b_m - b_{m-1} \leqslant T_0,\ a_m = 1,$$

where $m \leqslant 1/T_0 + 1$. For $\tau \in [b_l, b_{l+1}]$, we have

$$a_{\mathbf{k}}(\tau) = a_{\mathbf{k}}(b_l) + \int_{b_l}^\tau [R_{\mathbf{k}}(a(s)) + \mathcal{Y}_{\mathbf{k}}(a(s), s)]ds.$$

The first term of the integrand is the vector field of the effective equation. Our task is to estimate the second term.

Let us denote $\mathbf{Y}_{\mathbf{k}}(\tau) = \int_{b_l}^\tau \mathcal{Y}_{\mathbf{k}}(a(s), s)ds$. Then

$$\mathbf{Y}_{\mathbf{k}}(\tau) = \mathbf{Y}_{\mathbf{k}}(\tau) - \int_{b_l}^\tau \mathcal{Y}_{\mathbf{k}}(a(b_l), s)ds + \int_{b_l}^\tau \mathcal{Y}_{\mathbf{k}}(a(b_l), s)ds.$$

The last term equals

$$\begin{aligned}\mathcal{I}_{\mathbf{k}}(b_l, \tau) = &\sum_{(\mathbf{k}_1, \ldots, \mathbf{k}_{2q+1}) \in S(\mathbf{k}) \setminus \mathcal{R}(\mathbf{k})} a_{\mathbf{k}_1}(b_l)\overline{a_{\mathbf{k}_2}(b_l)} \ldots a_{\mathbf{k}_{2q-1}}(b_l)\overline{a_{\mathbf{k}_{2q}}(b_l)}a_{\mathbf{k}_{2q+1}}(b_l) \\ &\times \frac{\epsilon}{i[-\lambda_{\mathbf{k}} + \sum_{j=1}^{2q+1}(-1)^{j-1}\lambda_{\mathbf{k}_j}]} \exp\{\epsilon^{-1} i[-\lambda_{\mathbf{k}} + \sum_{j=1}^{2q+1}(-1)^{j-1}\lambda_{\mathbf{k}_j}]s\}\Big|_{b_l}^\tau.\end{aligned}$$

Let $\mathcal{I}(b_l,\tau)=\Big(\mathcal{I}_{\mathbf{k}}(b_l,\tau),\mathbf{k}\in\mathbb{Z}^d\Big)$. Since the quantities $|-\lambda_{\mathbf{k}}+\sum_{j=1}^{2q+1}(-1)^{j-1}\lambda_{\mathbf{k}_j}|$, if do not equal to zero, are alway bigger than 1, hence we have

$$\max_{\tau\in[b_l,b_{l+1}]}|\mathcal{I}(b_l,\tau)|_p\leqslant 2\epsilon\max_{v\in B(M_1)}|P(v)|_p\leqslant\epsilon 2C(M_1).$$

Then choosing $T_0=\epsilon^{1/2}$, using (4.4.2), we obtain

$$\begin{aligned}&|a(\tau)-a(0)-\int_0^\tau R(a(s))ds|_p\\&\leqslant\sum_{l=0}^{m-1}\Big\{\int_{b_l}^{b_{l+1}}|\mathcal{Y}(a(s),s)-\mathcal{Y}(a(b_l),s)|_pds+|\mathcal{I}(b_l,b_{l+1})|_p\Big\}\\&\leqslant\sum_{l=0}^{m-1}\Big[\int_{b_l}^{b_{l+1}}C(M_1)|a(s)-a(b_l)|_pds+\epsilon 2C(M_1)]\\&\leqslant[\frac{1}{2}T_0^2C(M_1)+\epsilon 2C(M_1)](\frac{1}{T_0}+1)\leqslant C'(M_1)\epsilon^{1/2}.\end{aligned}$$

This proof the assertion of the Lemma. □

Let $a'(\tau)$ be a solution of the effective equation (4.1.8) with initial data $a'(0)\in B(M_1)$. Denote

$$T'=\min\{\tau:\ |a'(\tau)|_p\geqslant M_1+1\}.$$

By Lemmata 4.3.2 and 4.4.1, we have for $\tau\in[0,\min\{1,T'\}]$,

$$|a'(\tau)-a(\tau)|_p\leqslant|a'(0)-a(0)|_p+\int_0^\tau C(M_1+1)|a'(s)-a(s)|_pds+C(M_1)\epsilon^{1/2}.$$

By Gronwall's lemma, we have that if $|a'(0)-a(0)|_p\leqslant\epsilon^{1/2}$, then

$$|a'(\tau)-a(\tau)|_p\leqslant C\epsilon^{1/2},\quad \tau\in[0,\min\{1,T'\}].$$

Assuming ϵ small enough and using the bootstrap argument we get that $T'\geqslant 1$.

Since $I(a(\tau))=I(v(\tau))$, we have

$$|I(v(\tau))-I(a'(\tau))|_p^{\sim}\leqslant C\epsilon^{1/2},\quad \tau\in[0,1].$$

This finishes the proof of Theorem 4.1.2.

4.5 Discussion of Proposition 4.1.1

Briefly speaking, the Proposition 4.1.1 directly follows from the global existence theory of the nonlinear Schrödinger equation (4.1.1). The equation (4.1.1) has two conservative quantities :

$$||u(t)||_0=||u(0)||_0, \tag{4.5.1}$$

and

$$E_q(u(t))=\int_{\mathbb{T}^d}\frac{1}{2}|\nabla u(x,t)|^2dx\pm\frac{\epsilon}{2q+2}\int_{\mathbb{T}^d}|u(x,t)|^{2q+2}dx=E(u(0)).$$

We claim the H^1-norm $||u(t)||_1$ remains bounded if the parameter ϵ is small enough. Indeed, the defocusing case is clear. In the focusing case, we have

$$\int_{\mathbb{T}^d} |\nabla u(x,t)|^2 dx = \frac{\epsilon}{q+1} \int_{\mathbb{T}^d} |u(x,t)|^{2q+2} dx + 2E(u(0)).$$

Using the L^2-conservation law and the Sobolev embedding :

$$H^1(\mathbb{T}^d) \to L^r, \quad r < \infty \quad \text{and} \quad r \leqslant \frac{2d}{d-2},$$

we obtain for d and q satisfying condition (4.1.2),

$$||u(t)||_1^2 \leqslant ||u(0)||_1^2 + \epsilon C(q,d)||u(t)||_1^{2q+2}.$$

So

$$||u(t)||_1^2 \leqslant \frac{||u(0)||_1^2}{1-\epsilon C(q,d)||u(t)||_1^{2q}}.$$

If $\epsilon \leqslant C(q,d)^{-1}2^{4q-1}||u(0)||_1^{-2q}$, we have

$$||u(t)||_1 \leqslant C(||u(0)||_1). \tag{4.5.2}$$

Now we give a direct proof of the Proposition 4.1.1 in the case $d=2$ and $q=1$, following [17]. Similar proof works for the cases $d=1$ and $q \in \mathbb{N}$.

Lemma 4.5.1. *For every $u \in H^2(\mathbb{T}^2)$ with $||u||_1 \leqslant 1$, we have*

$$||u||_{L^\infty} \leqslant C(1+\sqrt{\log(1+||u||_2)}).$$

For a proof of this lemma, see See Lemma 2 in [17].

Lemma 4.5.2. *For $u \in H^2(\mathbb{T}^2)$, we have*

$$|||u|^2u||_2 \leqslant C||u||_{L^\infty}^2||u||_2.$$

Démonstration. For $u \in H^2(\mathbb{T}^2)$ we have

$$|\triangle(|u|^2u)| \leqslant C(|u|^2|\triangle u| + |u||\nabla u|^2),$$

and so

$$|||u|^2u||_2 \leqslant C||u||_{L^\infty}^2||u||_2 + C||u||_{L^\infty}(\int_{\mathbb{T}^2}|\nabla u|^4dx)^{1/2}. \tag{4.5.3}$$

Using the Gagliardo-Nirenberg inequality (see [67]), we have

$$(\int_{\mathbb{T}^2}|\nabla u|^4dx)^{1/2} \leqslant C||u||_{L^\infty}||u||_2. \tag{4.5.4}$$

Combining (4.5.3) and (4.5.4) we obtain the statement of the lemma. □

Let us denote by $S(t)$ the L^2 isometry $S(t) = e^{-it\triangle}$. Then we have

$$u(t) = S(t)u(0) + i\epsilon \int_0^t S(t-s)|u(s)|^2 u(s) ds.$$

Using Lemmata 4.5.1 and 4.5.2 and the boundness of H^1-norm we have

$$||u(t)||_2 \leqslant ||u(0)||_2 + C\epsilon \int_0^t ||u(s)||_2[1 + \log(1 + ||u(s)||_2)]ds.$$

So

$$||u(t)||_2 \leqslant ||u(0)||_2 e^{C_1 e^{C_2 \epsilon t}}.$$

This verifies the statement of Proposition 4.1.1 in this case.

Remark 4.5.3. *The same proof also applies to nonlinear Schrödinger equations on $\mathbb{T}^2$ with other cubic nonlinearities, e.g. the nonlinearities with Hamiltonians of the forms $\mathcal{H}_3 = \int |u|^2(u + \bar{u})dx$ and $\mathcal{H}_3' = \int u^3 + \bar{u}^3 dx$.*

For the other cases, more sophisticated theory is needed. we refer the readers to the theories of the Cauchy problem for NLS equations in [13, 29]. From there we know that for a solution $u(t)$ of the equation (4.1.1) with $u(0) \in H^2$, there exist $T_1 > 0$ and $C_1 > 0$ that depend only on the bound of the H^1-norm $||u(t)||_1$ (inequality (4.5.2)) such that for every $t_0 \in [0, \infty)$, we have

$$||u(t)||_2^2 \leqslant ||u(t_0)||_2^2 + \epsilon T_1 C_1 ||u(t_0)||_2^2, \quad t \in [t_0, t_0 + T_1].$$

Therefore

$$||u(t)||_2 \leqslant C(||u(0)||_2) e^{\epsilon t C'(||u(0)||_2)}.$$

This confirms the assertion of Proposition 4.1.1.

Bibliographie

[1] V. I. Arnold. *Mathematical methods in classical mechanics.* Springer, Berlin, 3rd edition, 1989.

[2] V. I. Arnold, V. V. Kozlov, and A. I. Neistadt. *Mathematical Aspects of Classical and Celestial Mechanics.* Springer, Berlin, third edition, 2006.

[3] R. Aron and J. Cima. A theorem on holomorphic mappings into Banach spaces with basis. *Proceeding of the American Mathematical Society*, 36(1), 1972.

[4] D. Bambusi. Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equation. *Math. Z.*, 130 :345–387, 1999.

[5] D. Bambusi. An averaging theorem for quasilinear hamiltonian PDEs. *Ann. Henri. Poincaré*, 4 :685–712, 2003.

[6] D. Bambusi. Galerkin averaging method and Poincaré normal form for some quasilnear pdes. *Ann. Scuola Norm. Sup. Pisa C1. Sci.*, pages 669–702, 2005.

[7] D. Bambusi and B. Grebert. Forme normale pour NLS en dimension quelconque. *Comptes Rendus Mathematique*, 337 :409–414, 2003.

[8] D. Bambusi and B. Grébert. Birkhoff normal form for partial differential equations with tame modulus. *Duke Mathematical Journal*, 135(3) :507–567, 2006.

[9] E. D. Belokolos, A. I. Bobenko, V. Z. Enolskii, A. R. Its, and V. B. Matveev. *Algebro-geometric approach to nonlinear integrable equations.* Springer, 1994.

[10] A. Bobenko and S. Kuksin. Finite-gap periodic solutions of the KdV equation are non-degenerate. *Physics Letters A*, 161(3) :274–276, 1991.

[11] V. Bogachev. *Differentiable measures and the Malliavin calculus.* American Mathematical Society, 2010.

[12] V. Bogachev and I. Malofeev. On the absolute continuity of the distributions of smooth functions on infinite-dimensional spaces with measures. *preprint*, 2013.

[13] J. Bourgain. Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. *Geometric and Functional Analysis*, 3(3) :209–262, 1993.

[14] J. Bourgain. *Global solutions of nonlinear Schrödinger equations*, volume 46. American Mathematical Society, 1999.

[15] J. Bourgain. On diffusion in high-dimensional Hamiltonian systems and PDEs. *Journal d'Analyse Mathématique*, 80(1) :1–35, 2000.

[16] J. Boussinesq. Théorie de l'intumescence liquid appelée onde solitaire ou de translation, se propageant dans un canal rectangulaire. *Comptes Rend. Acad. Sci (Paris)*, 72 :755–759, 1871.

[17] H. Brézis and T. Gallouet. Nonlinear Schrödinger evolution equations. *Nonlinear Analysis : Theory, Methods & Applications*, 4(4) :677–681, 1980.

[18] N Burq and N Tzvetkov. Random data cauchy theory for supercritical wave equations ii : a global existence result. *Inventiones mathematicae*, 173(3) :477–496, 2008.

[19] J. Colliander, M.Keel, G.Staffiliani, and T. Tao. Symplectic nonsqueezing of the Korteweg-de Vries flow. *Acta Math.*, 195 :197–252, 2005.

[20] B. Dubrovin. Theta functions and nonlinear equations. *Russ. Math. Surv*, 36 :11–92, 1981.

[21] B. Dubrovin, V. Matveev, and S. Novikov. Nonlinear equations of Korteweg-de Vries type, finite zone linear operators, and Abelian varieties. *Russ. Math. Surv*, 31(1) :55–136, 1976.

[22] R. M. Dudley. *Real Analysis and Probability.* Cambridge University Press, 2002.

[23] H. Eliasson. Normal forms for Hamiltonian systems with Poisson commuting integrals — elliptic case. *Commentarii Mathematici Helvetici*, 65(1) :4–35, 1990.

[24] E. Faou, P. Germain, and Z. Hani. The weakly nonlinear large box limit of the 2d cubic nonlinear Schrödinger equation. *preprint*, 2013. arXiv :1308.6267.

[25] H. Flaschka and D. McLaughlin. Canonically conjugate variables for the Koreweg-de Vries equation and the Toda lattice with periodic boundary conditions. *Progress of Theoretical Physics*, 55(2) :438–457, 1976.

[26] C. Gardner. Korteweg-de Vries equation and generalizations. i.v. the Korteweg-de Vries equation as a Hamiltonian system. *J. Math. Phys*, 12 :1548–1551, 1971.

[27] J. Garnett and E. Trubowitz. Gaps and bands of one dimensional periodic Schrödinger operators. *Comment. Math. Helv.*, 59 :258–312, 1984.

[28] P. Gérard and S. Grellier. Effective integrable dynamics for a certain nonlinear wave equation. *Analysis and PDE*, 5 :1139–1154, 2012.

[29] S. Herr, D. Tataru, and N. Tzvetkov. Global well-posedness of the energy-critical nonlinear Schrödinger equation with small initial data in $H^1(\mathbb{T}^3)$. *Duke Mathematical Journal*, 159(2) :329–349, 2011.

[30] G. Huang. An averaging theorem for nonlinear Schrödinger equations with small nonlinearities. *to appear in DCDS-A*. arXiv :1312.0759.

[31] G. Huang. An averaging theorem for a perturbed KdV equation. *Nonlinearity*, 26 :1599–1621, 2013.

[32] G. Huang. On long time dynamics of perturbed KdV equations. *preprint*, 2013.

[33] G. Huang and S. Kuksin. KdV equation under periodic boundary conditions and its perturbations. *preprint*, 2013.

[34] H. Ito. Convergence of Birkhoff normal forms for integrable systems. *Commentarii Mathematici Helvetici*, 64(1) :412–461, 1989.

[35] A. Its and V. Matveev. Schrödinger operators with the finite-band spectrum and the N-soliton solutions of the Korteweg-de Vries equation. *Teoret. Mat. Fiz.*, 23 :51–68, 1975.

[36] T. Kappeler and S. Kuksin. Strong non-resonance of Schrödinger operators and an averaging theorem. *Physica D*, 86 :349–362, 1995.

[37] T. Kappeler, C. Möhr, and P. Topalov. Birkhoff coordinates for KdV on phase spaces of distributions. *Selecta Mathematica, New Series*, 11(1) :37–98, 2005.

[38] T. Kappeler and J. Pöschel. *KdV & KAM.* Springer, 2003.

[39] T. Kappeler, B. Schaad, and P. Topalov. Asymptotics of spectral quantities of Schrödinger operators. In *Spectral geometry*, Proc. Sympos. Pure Math. 84, pages 243–284. American Mathematical Society, 2012.

[40] T. Kappeler, B. Schaad, and P. Topalov. Qualitative features of periodic solutions of KdV. *Communications in Partial Differential Equations*, 2013. arXiv :1110.0455.

[41] T. Kappeler and P. Topalov. Global wellposedness of KdV in $H^{-1}(T, R)$. *Duke Math. J.*, 135 :327–360, 2006.

[42] E. Korotyaev. Estimates for the Hill operator, II. *Journal of Differential Equations*, 223(2) :229–260, 2006.

[43] D. Korteweg and G. de Vries. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Philos. Mag.*, 539 :422–443, 1895.

[44] I. Krichever. Perturbation theory in periodic problems for two-dimensional integrable systems. *Sov. Sci. Rev. C. Math. Phys.*, 9 :1–101, 1991.

[45] M. Kruskal and N. Zabusky. Interaction of “solitons" in a collision-less plasma and the recurrence of initial states. *Phys. Rev. Lett.*, 15 :240–243, 1965.

[46] S. Kuksin. Perturbation theory for quasi-periodic solutions of infinite dimensional Hamiltonian systems, and its application to the Koreteweg-de Vries equation. *Math. USSR Subornik*, 67 :397–413, 1989.

[47] S. Kuksin. *Analysis of Hamiltonian PDEs.* Oxford University Press, Oxford, 2000.

[48] S. Kuksin. Hamiltonian PDEs. In *Handbook of dynamical systems*, volume 1, pages 1087–1133. Elsevier, 2006.

[49] S. Kuksin. Damped driven KdV and effective equations for long-time behaviour of its solutions. *GAFA*, 20 :1431–1463, 2010.

[50] S. Kuksin. Weakly nonlinear stochastic CGL equations. *Annales de l’insitut Henri Poincaré-Probabilité et Statistiques*, 49(4) :1033–1056, 2013.

[51] S. Kuksin and A. Maiocchi. Resonant averaging for weakly nonlinear stochastic Schrödinger equations. *preprint*, 2013. arXiv :1309.5022.

[52] S. Kuksin and G. Perelman. Vey theorem in infinite dimensions and its application to KdV. *DCDS-A*, 27 :2010, 1-24.

[53] S. Kuksin and A. Piatnitski. Khasminskii-Whitham averaging for randomly perturbed KdV equation. *J. Math. Pures Appl.*, 89 :400–428, 2008.

[54] S. Kuksin and J. Poschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Annals of Mathematics*, pages 149–179, 1996.

[55] P. Lax. Integrals of nonlinear equations of evolution and solitary waves. *Comm. Pure Appl. Math.*, 21 :467–490, 1968.

[56] P. Lax. Outline of a theory of the KdV equation. In *Recent mathematical methods in nonlinear wave propagation*, Lecture Notes in Mathematics 1640, pages 70–102. Springer, 1996.

[57] P. Lochak and C. Meunier. *Multiphase averaging for classical systems : with applications to adiabatic theorems.* Springer, 1988.

[58] V. Marcênko. *Sturm-Liouville operators and applications.* Birkhäuser,Basel, 1986.

[59] H. McKean and E. Trubowitz. Hill's operator and hyper-elliptic function theory in the presence of infinitely many branch points. *Comm. Pure Appl. Math*, 29 :143–226, 1976.

[60] H. McKean and E. Trubowitz. Hill's surfaces and their theta functions. *Bull. Am. Math. Soc*, 84 :1042–1085, 1978.

[61] R. Miura. Korteweg-de Vries equation and generalizations. I. A remarkable explicit nonlinear transformation. *J. Math. Phys.*, 9 :1202–1204, 1968.

[62] J. Moser. Lectures on Hamiltonian systems. *Mem. Amer. Math. Soc.*, 81 :1–60, 1968.

[63] J. Moser and C. L. Siegel. *Lectures on celestial mechanics.* Springer, Berlin, 1971.

[64] R. Muira, C. Gardner, and M. Kruskal. Korteweg-de Vries equation and generalizations. II. Existence of conservation laws and constants of motion. *J. Math. Phys*, 9 :1204–1209, 1968.

[65] S. Nazarenko. *Wave Turbulence.* Springer, 2011.

[66] A. I. Neishtadt. Averaging in multi-frequency systems, I and II. *Soviet Phys, Doklady*, 20(7) and 21(2) :492–494 and 80–82, 1975.

[67] L. Nirenberg. On elliptic partial differential equations. *Ann. Sci. Norm. Sup. Pisa*, 13 :115–162, 1959.

[68] J. Pöschel. On Nekhoroshev estimates for a nonlinear Schrödinger equation and a theorem by Bambusi. *Nonlinearity*, 12 :1587–1600, 1999.

[69] T. Runst and W. Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3. de Gruyter, 1996.

[70] J. Saut and R. Teman. Remarks on the Kortewed-de Vries equation. *Israel J. Math*, 24 :1976, 78-87.

[71] N. Temirgaliev. A connection between inclusion theorems and the uniform convergence of multiple fourier series. *Matematicheskie Zametki*, 12(2) :139–148, 1972.

[72] J. Vey. Sur certain systèmus dynamiques séparables. *Am. J. Math*, 100 :591–614, 1978.

[73] H. Whitney. Differentiable even functions. *Duke Math. Journal*, 10 :159–160, 1942.

[74] V. E. Zakharov, S. V. Manakov, S. P. Novikov, and L. P. Pitaevskij. *Theory of Solitons : The Inverse Scattering Method*. Plenum Press, 1984.

[75] P. E. Zhidkov. *Korteweg-de Vries and Nonlinear Schrödinger equations : qualitative theory*. Lecture Notes in Mathematics 1756. Springer-Verlag, Berlin, 2001.