SAPIENZA
UNIVERSITÀ DI ROMA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

PhD Thesis in Mathematics

by

Roberto Feola

# Quasi-periodic solutions for fully nonlinear NLS

Supervisor

Michela Procesi

The Candidate

The Supervisor

# Abstract

In the thesis we consider a class of fully nonlinear (i.e. *strongly nonlinear*) Schrödinger equations and we prove the existence and the stability of Cantor families of quasi-periodic, small amplitude solutions. We firstly deal with forced reversible or Hamiltonian nonlinearities depending quasi-periodically on time, and secondly we discuss the existence of *analytic* solutions for reversible *autonomous* nonlinearities. Note that this is the first result on analytic quasi-periodic solutions for fully nonlinear PDEs.

In the forced cases we use a Nash-Moser scheme on scales of Sobolev spaces combined with a reducibility argument of the linearized operator in a neighborhood of zero. In the reducibility step we exploit the pseudo-differential structure of the linearized operator by using changes of variables, induced by torus diffeomorphisms and pseudo-differential operators, which conjugate it to a constant coefficients differential operator plus a bounded remainder. Then we use a KAM-like scheme which diagonalizes the linearized operator and provides a sufficiently accurate asymptotic expansion of the eigenvalues. This procedure produces a change of variables which diagonalizes the operator linearized at the solution. In the Hamiltonian case the linearized operator has multiple eigenvalues so we are able to obtain only a block-diagonalization: in any case we deduce informations about the linear stability of the quasi-periodic solution.

In the autonomous case we first perform a "weak" Birkhoff normal form step in order to find an invariant manifold of the third order approximate NLS on which the dynamics is integrable. Then, in order to deal with reversible and analytic nonlinearities, we introduce a suitable generalization of a KAM nonlinear iteration for "tame" vector fields which works both in Sobolev and analytic regularity. In particular such scheme preserves the "pseudo-differential structure" of the vector field and allows us to use the techniques developed in forced cases in order to invert the linearized operator in the "normal" directions.

# Introduction

Nonlinear partial differential equations (nonlinear PDEs in the following) play a very important rôle in various physical phenomena. Nonlinear models appear in non-equilibrium statistical mechanics, fluids dynamics, quantum mechanics and many other fields. Naturally the qualitative description of solutions strongly depends on the spatial domain in which one is studying the problem. Let us consider for instance the case of dispersive PDEs, i.e. equations for which different frequencies propagate at different group velocities. When the spatial domain is the real line, on long time scales one expect to see the effect of the dispersion as a $L^{\infty}$ decay of the solution. On the other hand on compact domains one might expect the presence of recursive behaviors. In this thesis we consider the latter type of phenomena in a neighborhood of zero. In general the methods used to study nonlinear PDEs are problem depending, and the behavior of the system could differ drastically from its linear approximation, even for solutions in a neighborhood of zero. Many evolution partial differential equations on compact domains, can be written as infinite dimensional dynamical systems in some separable Hilbert space. Precisely they can be written as

$$\dot{w} = Lw + f(w), \tag{1}$$

where $L$ is a linear operator, $F$ is a sufficiently smooth non linearity and $w$ belongs to some Sobolev space $H^s$.

Typically $L : H^s \to H^{s-n}$ is an unbounded differential operator of order $n$. For instance in the case of NLS equation $L$ is $i\Delta$ where $\Delta$ is the Laplace-Beltrami operator. In general one can consider nonlinearities that map some neighborhood of $H^s$ to $H^{s-\delta}$ for some $\delta \geq 0$. In other words one can consider *bounded* ($\delta = 0$) or *unbounded* ($\delta > 0$) nonlinearities. Most of the literature concerns model

PDEs deriving from a suitable approximation of physical models, such as the wave equation, water waves equation and many models deriving from it, like the non linear Schrödinger (NLS), the Korteweg-de Vries (KdV), the Boussinesq and so on. See for instance [24] for an extensive exposition of the topic. We mention the most famous models: the cubic wave equation (NLW):

$$u_{tt} - \Delta u + \mathtt{m} u + u^3 = 0, \tag{2}$$

the cubic NLS

$$\mathrm{i} u_t = \Delta u + |u|^2 u, \tag{3}$$

as examples of equation with a bounded nonlinearity and the KdV equation

$$u_t - u u_x + u_{xxx} = 0, \tag{4}$$

which present an unbounded nonlinearity ($\delta = 1$).

In this thesis we study *fully nonlinear* ( i.e. $\delta = n$ ) NLS on the circle with the purpose of proving the existence of very special global solutions. The NLS equation is certainly one of the most studied physical models. Such equation appears in different context like Bose-Einstein condensate theory or quantum mechanics. In fluids dynamics the NLS appears in the study of small-amplitude gravity waves. In general NLS appears in the description of various phenomena involving evolution of waves packets in media which have dispersion, by performing asymptotic expansions and scaling limits, see for instance [60]. It may also happen that, in some regime of approximation, appears with unbounded nonlinearities. This is why we considered the fully nonlinear model.

**Dynamical system approach.** When considering equations like (1) on tori and more generally on compact manifolds, the spectrum of the linear operator is purely pointwise, hence the so called "dynamical system approach" turned out to be a very fruitful approach to the study of these problems. In other words some ideas and techniques coming from the theory of dynamical systems have been extended to the infinite dimensional setting. One of the most important ideas one borrows from the finite dimensional theory is to look for invariant manifold on which the dynamics is particularly simple. Then one would like to deduce general properties of the whole system by studying the behavior of a typical initial datum. The simplest invariant manifold to look for are equilibria, periodic orbits, and then *quasi*-periodic solutions. It is known that the measure of the initial data corresponding to periodic or quasi-periodic motions is zero. However one can hope to understand the dynamics at least for initial data in a neighborhood of such particular solutions. This is actually one of the most famous conjectures of Poincaré. Historically such conjecture was an important motivation for a systematic study of such particular orbits in the finite dimensional setting, which lead in $50s$ and $60s$ to the celebrated KAM

Theorem proved by Kolmogorov [40], Arnold [1] and Moser [49]. The theory, developed for finite-dimensional Hamiltonian systems, guarantees that for *nearly* integrable systems a "large" measure set of the phase space corresponds to quasi-periodic solutions (KAM tori). The key problem arising in the search of such quasi-periodic solutions is the presence of the so called "small divisors", which are arbitrary small quantities appearing in the denominators when one computes explicitly the approximate perturbative series. Even in finite dimension is clear that such small denominators are main source of chaotic behaviors.

In the PDEs context the problems concerning the small denominators are much harder with respect to the finite dimensional case. As an example one can think that, for Hamiltonian PDEs, the problem of small divisors appears also in the search of *periodic* solutions, while, in the finite-dimensional setting, the periodic solutions are not plagued by small divisors.

In order to pass from finite-dimensional systems to the infinite-dimensional ones, one of the first questions analyzed has been the "integrability" of a system. This matter has been widely studied since '60*s* and it is connected to the "solvability" of the equations of motions by exploiting possible symmetries of the system. A "strong" notion of integrability in the case of Hamiltonian systems is the "Liouville integrability" which relies on the existence of a foliation of the phase space in manifolds which are invariant for the flow of the systems. In this case there exist a canonical coordinate systems, called *action-angle* variables.

Proving that a dynamical system is integrable is clearly a non trivial task. However integrability has been proved for various nonlinear PDEs by explicitly constructing the so-called Lax pair. Famous examples are the 1-dimensional cubic NLS (proved by Zakharov and Shabat (1972)), the KdV equation, the Benjamin-Ono equation, the Degasperis-Procesi and Camassa-Holm equations. Then, as in the finite-dimensional case, a perturbation theory has been developed for *nearly*-integrable systems, see for instance paper by Kuksin [41]. A more recent approach for nearly-integrable equations is to consider directly systems which are non-integrable perturbations of equations already written in action-angle variables. This approach is close to the finite dimensional case and the key idea is to start the perturbation theory by bifurcating from a quasi-periodic solution of the integrable system. However finding local or global action-angle variables is a delicate question; see for instance [44], [38], [36], [47], [39]. Therefore in practice this latter approach may not be possible. Another approach is to implement the perturbation theory near zero, by bifurcating form a solution of the linear equation. This is in some sense simpler because linear equations are integrable and one does not need the complex analysis required to find action-angle variables for nonlinear systems. On the other hand KAM theory was first developed for anisochronous systems because one needs parameters to modulate (for instance a frequency to amplitude modulation). It could happen that the quasi-periodic solutions appear only due to the nonlinearity (this happens for *completely*-resonant PDEs). Hence bifurcating from a *periodic* solution can make the perturbation theory very difficult due to the possible *resonances* of the frequencies. In any case the perturbation theory developed up to now does not provide the existence of *almost*-periodic solutions, i.e. embeddings of maximal tori, the so called *Lagrangian* tori. In finite dimension the classical KAM theory provides precisely the existence of such kind of embeddings. In infinite dimension this matter seems to be out of reach at the present time, except for some ad hoc models, (see [18], [21], [51]). On the contrary a lot of efforts are focused on the search of embeddings of finite-dimensional tori, that is quasi-periodic solutions. This is actually the equivalent in infinite dimension of the search for lower dimensional tori in finite-dimensional dynamical systems; see for instance [49], [35], [33], for classical references and the more recent papers [20], [25], [26] where degenerate situations are considered.

In general, due to the infinite-dimensional setting and to the presence of resonances, one expects the dynamics to exhibit a very complex behavior, with coexistence of both stable and unstable phenomena. On the one hand people started to investigate some invariant structure, such as periodic and quasi-periodic orbits; on the other hand one exploit the structure of resonances and diffusive orbits in order to prove the existence of instability and phenomena such as the growth of Sobolev norms. As references on the latter topic we mention the results of Bourgain in [15] and Staffilani [56] who proved that the Sobolev norms grow at most polynomially in time, and the breakthrough results by Colliander, Keel, Staffilani, Takaoka and Tao, who proved in [23] the existence of solutions of cubic NLS evolving from a small initial data and undergoing an arbitrarily large growth. A dual point of view of such theory are long time existence and stability results obtained by using techniques of Birkhoff normal form and Nekhoroshev theory. Such theories are quite well-established in the finite-dimensional setting while results in the infinite dimensional setting are more recent. In a neighborhood of an elliptic equilibrium point the Birkhoff normal form gives a quite good description of the dynamics. Roughly speaking a Birkhoff's like theorem provides the existence of a canonical map which puts the Hamiltonian in *normal* form up to a remainder of arbitrary large order. Hence, while KAM results provides the existence of *special global* solutions, through Birkhoff theory one describes the evolution of *any* initial data in a neighborhood of the equilibrium point, for large but finite time. We mention for instance Delort [28], Bambusi-Grèbert [6] which adapted these techniques to the infinite-dimensional setting.

**Some Literature.** In this thesis we are interested in studying phenomena of stability, and in particular in studying existence of quasi-periodic solutions for PDEs. The existence of small-amplitude periodic quasi-periodic solutions was one of the first successes of KAM theory for PDEs, obtained by Kuksin [41] and Wayne [58]. Such results were restricted to the case in which the spatial variable ranges in a finite interval with Dirichlet boundary conditions. In order to consider the case of periodic boundary conditions, Craig-Wayne used a Lyapunov-Schmidt reduction method in [27] later generalized by Bourgain in [19], [16]. Other developments of KAM Theory for PDEs can be found in [50], [22], [43], [45]. All the papers quoted above deal with "non-resonant" PDEs, i.e. equations for which the linear frequencies satisfies some non degeneracy conditions. In this way the bifurcation analysis is much simpler. For example we mention the model of NLS considered in [29] having the form

$$\mathrm{i}u_t = \Delta u + V(x) * u + O(u^2) \tag{5}$$

where $V$ is a regular convolution potential. In this case the linear frequencies have the form $|j|^2 + V_j$, (with $V_j$ the Fourier coefficients of the function $V$) . In the "completely resonant" case, i.e. $V \equiv 0$, in addition to the small divisors problem, appears also the issue of infinite-dimensional bifurcation problem. See for instance [55], [34], [53]. Let us put aside the bifurcation problem for the moment, and focus on the small divisors matter. Typically one has that the equation linearized about zero is not invertible, since the spectrum of the linearized operator accumulate to zero. In other words the inverse of the linearized operator produces a "loss" of regularity. Hence the classical implicit function theorem does not hold. To overcome this problem one typicaly uses use an iterative scheme in order to find a sequence of approximate solutions rapidly converging to the true solution. Such fast convergence is used to control the loss of regularity due to the small divisors. Two different approaches have been mainly developed so far: the Nash-Moser scheme and the KAM algorithm.

A Nash-Moser scheme is essentially a Newton method to find zeros of polynomial extended to functional spaces. The idea is to reduce the search of quasi-periodic solutions to the search of zeros of a suitable functional. In order to run an algorithm of this type one must be able to control the linearized operator in a neighborhood of the expected solution see Figure 1. Due to the presence of small divisors it is not possible to invert such operator as a functional from a Sobolev space to itself (not even the operator linearized about zero). However, since the Newton scheme is quadratic, one may accept that $d_uF^{-1}$ is well defined as "tame" operator from $H^s$ to $H^{s-\mu}$ for some appropriate $\mu$.

The KAM scheme is to find a converging sequence of changes of coordinates such that in the final variables the solution is trivially at the origin of the phase space. The key idea is to remove from the system the non homogeneous terms through a translation: in this way the solution is obviously at the origin of the new coordinate system

However it turns out that the equation for the translation is essentially the same as the one for the approximate solution in the Nash-Moser scheme. Hence the solvability conditions are quite the same and involves lower bounds on the eigenvalues (the so called "First Mel'nikov" conditions) of the linearized operator. However one sees immediately that there is no hope to have uniform lower bounds for the eigenvalues. Besides this, in classical KAM approach one asks for "stronger" non-degeneracy conditions

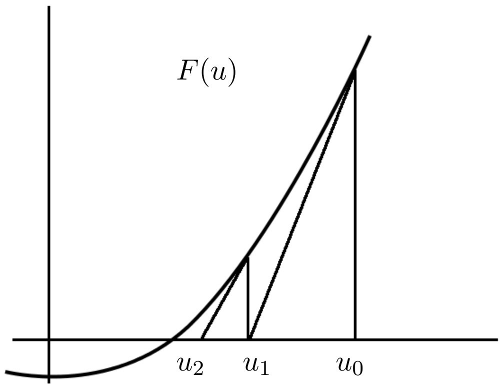


Figure 1: Three steps of the Newton algorithm $u_{n+1} := u_n - (d_u F(\lambda, \varepsilon, u_n))^{-1}[F(\lambda, \varepsilon, u_n)]$

of the eigenvalues in order to simplify the equation for the translation. The idea is to use a reducibility argument, i.e. to diagonalize the linearized operator. It is well known that it is possible to diagonalize a finite dimensional matrix, with a regular transformation, if it has distinct eigenvalues. Hence in order to diagonalize the infinite dimensional matrix appearing in the KAM schemes one asks for lower bounds on the differences of the eigenvalues (the so called "Second Mel'nikov" conditions) in addiction to the First Mel'nikov conditions. This requirement together with some structural hypothesis on the system (Hamiltonianity, reversibility, ....) provides the linear stability of the possible solution. Naturally the reducibility is a sufficient condition for the invertibility and typically in a Nash-Moser scheme it is not required. Historically this is considered to be the main difference between the two approaches: in general when one refers to *KAM* theorem, implicitly one means "reducible" solutions. Notice moreover that the Nash-Moser scheme is, in principle, coordinate independent while the KAM scheme by its nature relies on the existence of privileged coordinate systems. This last point implies that one loses track of the fact that the original Banach space is a space of functions.

We have to underline that, imposing conditions like the First or the Second Mel'nikov is not trivial at all. First one needs some parameters to modulate in order to avoid resonances (in (5) the $V_j$ play the rôle of such parameters). In other words one exclude parameters associated to "bad" denominators. Secondly one has to satisfy *infinite* conditions, hence it may be possible to exclude a full measure set of such parameters, and in principle one may have to introduce some weaker property such as a block diagonalization. Moreover in certain cases it is very difficult to verify second Mel'nikov conditions, in particular in presence of multiple eigenvalues. For instance the NLS in (3) on $\mathbb{T}^n$ with $n \geq 1$. This is one of the reason why the first KAM results were for PDEs with Dirichelet boundary conditions.

Indeed the first existence results for periodic and quasi-periodic solutions on tori have been obtained by using a Nash-Moser approach by Bourgain in [16] for the nonlinear Schrödinger equation on $\mathbb{T}^2$ with a convolution potential. Here the author used subtle multiscale argument to estimate the decay of the inverse of the linearized operators developed in the following in [17] to obtain the existence result on $\mathbb{T}^d$. We mention also the remarkable results by Berti and Bolle [13], [10] which study equations in presence of a more natural multiplicative potential; in particular the authors never exploit properties of "localization" of the eigenfunctions of $-\Delta + V(x)$ with respect to the exponentials that actually might be *not* true. This is why their approach applies also to equations with multiplicative potential and not only with convolution potential. The latter approach, based on a multi-scale analysis, has been very fruitfully exploited in the study of PDEs also on manifolds different from tori. In [14] Berti, Corsi and Procesi studied NLW and NLS on compact Lie groups and homogeneous manifolds. Again we remark that these papers rely on the so called "multi-scale" analysis based on the first Mel'nikov condition and geometric properties of "separation of singular sites", and that this methods do not imply reducibility and linear stability of the solutions. There are very few and recent results on reducibility on tori. We mention Geng-You in [32] for the smoothing NLS, Eliasson-Kuksin in [29] for the non resonant NLS (see (5)) and Procesi-Procesi [53] for the completely resonant NLS which involves deep arguments of normal forms developed in [52], [54]. All the aforementioned papers, both using KAM or multi-scale, are on semi linear PDEs with no derivatives in the non linearity.

More recently KAM theory has been developed also for dispersive semilinear PDEs on the one dimensional torus when the nonlinearity contains derivatives of order $\delta \leq n-1$, here $n$ is the order of the highest derivative appearing in the linear constant coefficients term. The additional difficulty in this case is that, due to the presence of derivatives in the nonlinearity, the KAM transformations used to diagonalize the linearized operator might be *unbounded*. The key idea to overcome such problem has been introduced by Kuksin in [42] in order to deal with non-critical unbounded perturbations, i.e. $\delta < n-1$, with the purpose of studying KdV type equations (4), see also [39]. The general strategy can be explained as follows. Roughly speaking the aim of a reducibility scheme is to iteratively conjugate an operator $D + \varepsilon M$ w.r.t. the exponentials basis, where $D$ is diagonal, to $D_+ + \varepsilon^2 M_+$ where $D_+$ is again diagonal in the exponentials basis. Let us consider the KdV case on $\mathbb{T}$ where the diagonal operator $D$ in the Fourier space (in time and space) has the form

$$D_j^k(l) = \mathrm{i}\omega \cdot l + \mathrm{i}j^3, \quad l \in \mathbb{Z}^d, \quad j \in \mathbb{Z}, \quad \omega \in \mathbb{R}^d \;\; d \geq 1. \tag{6}$$

The equation which defines the change of variables is called the *homological equation* while the operators $D, D_+$ are called the *normal form*. In one defines $A$ as the generator of the *quasi*-identically transformation, the homological equation has the form

$$\mathrm{ad}(D)[A] := [D, A] = \varepsilon M - \varepsilon[M], \tag{7}$$

where $[M] \in \mathrm{Ker}\big(\mathrm{ad}(D)\big)$. It is clear that the eigenvalues of the adjoint operator $\mathrm{ad}(D)$ involves the differences of the eigenvalues of $D$: here one uses conditions the Second Mel'nikov conditions in order to have a lower bounds on such differences. If $M$ is an unbounded operator then one may exploit the properties of the dispersion law of the equation to prove a stronger bound like

$$|\omega \cdot l + j^3 - k^3| \geq \frac{\gamma |j^3 - k^3|}{1 + |l|^\tau}. \tag{8}$$

Of course for $j \neq k$ one has $|j^3 - k^3| \geq |j^2 + k^2|$ and this good separation property allows to control derivatives in the nonlinearities. However it turns out that $D_+$ is diagonal in the space variable (with coefficients depending on time). The purpose of the so called *Kuksin-Lemma* is to show that such an algorithm can be run, namely that one can solve the homological equation also when the normal form is diagonal only in the space variable (as $D_+$ is). This approach, developed for the $KdV$ that has a *strong* dispersion law, has been further exploited by the Chinese school to cover the "less" dispersive case of NLS in presence of one derivative in the non linearity, i.e. the *critical* case when $\delta = n - 1$. Here the eigenvalues of the spatial differential operator have an asymphtotics like $|j|^2$, hence the separation properties are worse. In particular we mention Zhang, Gao and Yuan [61] which studied the reversible NLS and Liu and Yuan which in [46] deal with the Hamiltonian case. The previously mentioned results require that the equation is semi-linear and dispersive; in the "weakly dispersive" case of the derivative Klein-Gordon equation we mention the results [7]-[8], also based on KAM theory. In the latter papers the key idea is in the explicit computation of the first order asymptotic expansion of the perturbed normal frequencies, obtained using the notion of quasi-Töplitz function. This concept was introduced by Procesi-Xu [59] and it is connected to the Töplitz-Lipschitz property in Eliasson-Kuksin [29].

The KAM approach described above seems to fail in the fully nonlinear case and one has to develop different strategies. The point is that in solving (7) one is not able to prove that $A$ is a bounded operator. This is due the fact that along the iterative procedure one loses information on the PDEs structure. This is one of the motivation for developing the idea of quasi-Töplitz vector fields. The first breakthrough result for fully nonlinear PDEs is due to Iooss-Plotnikov-Toland who studied in [37] the existence of *periodic* solutions for water-waves; we mention also the papers by Baldi [2], [3] on periodic solutions for the Kirckhoff and Benjamin-Ono equations. These papers strongly rely on the PDEs structure. They are based on Nash-Moser methods and the key point is to apply appropriate diffeomorphisms of the torus and pseudo-differential operators in order to invert the operator linearized at an approximate solution.

However these results do not imply the linear stability of the solutions and they do not work in the quasi-periodic case. Here the idea, borrowed from pseudo-differential calculus, is to conjugate $D+\varepsilon M$ to an operator $D_+ + \varepsilon M_+$ where $D_+$ is again diagonal while $M_+$ is of lower order w.r.t. $M$. After a finite number of such steps one obtains an operator of the form $D_F + \varepsilon M_F$ where $D_F$ is diagonal and $M_F$ is as smoothing (in space) operator of order $k$, with $k$ arbitrarily large. In the search of *periodic* solutions this is sufficient to get the invertibility of $D_F + \varepsilon M_F$. Indeed the "vector" of frequency $\omega$ is actually one-dimensional, hence the small divisor in (6) becomes $\omega l + j^3$ with $l \in \mathbb{Z}$. Hence such quantity is a true small divisor only if $|l| \approx |j|^3$. In such a case one prove lower bounds of the form

$$|\omega l + j^3| \geq \frac{\gamma}{1+|j|^{\tau_1}}, \quad \text{for some} \quad \tau_1 > 0. \tag{9}$$

This implies that, if $M_F$ maps $H^s$ to $H^{s+\tau_2}$ with $\tau_2 \geq \tau_1$, i.e. it is sufficiently smoothing, then one can invert $D_F + \varepsilon M_F$ by Neumann series. Unfortunately, this arguments does not hold in the case of *quasi*-periodic solutions. In this case $\omega \in \mathbb{R}^d$ hence $l$ is not controlled by $|j|$ to some power and the bounds (9) does not hold.

Quite recently this problem has been overcome by Berti, Baldi, Montalto who studied fully nonlinear perturbations of the KdV equation first in [5], for the forced case, then in [4] for the autonomous. This was the first result for quasi-periodic solutions for quasi linear PDEs. As in the periodic case the main point is conjugate $D+\varepsilon M$ to an operator of the form $D_F$ diagonal and $M_F$ a bounded operator. This is done by exploiting the PDEs structure and using conjugation by flows of pseudo-differential vector fields. Note that such transformations preserves the PDEs structure. This is the main difference with the KAM reducibility scheme. Once one has obtained the structure $D_F + \varepsilon M_F$ one can apply a KAM reducibility scheme in order to diagonalize. One needs only to invert such operator and this could be done by a multi-scale argument, however since they are working in one space dimension they show that the second Mel'nikov conditions can be imposed. This gives the stronger stability result. This scheme, i.e. Nash-Moser plus reducibility of the linearized operator, is very reminiscent of the classical KAM scheme. The main difference with the classical KAM approach is that, following the ideas introduced in [5] one does not apply the changes of variables that diagonalize the linearized operator. Even if this scheme has been developed for the particular case of the KdV, it is quite clear that it can be generalized to larger class of dispersive PDEs on the circle, see for instance the application to capillary water waves (Beti-Montalto in preparation). An important point is that their scheme runs on Sobolev spaces, i.e. spaces with only finite regularity. This means that even if one starts with an analytic PDE (as in the case of most natural models) still the quasi-periodic solution has only finite regularity, both in time and space. This is not a technical question, but is related to the loss of regularity in the reducibility scheme. In other words in order to get the estimates on the inverse of the linearized operator one needs to control the norm of the operator which diagonalizes it. In the case of Sobolev regularity such operator is bounded from $H^s$ to itself. In the analytic case this is false. Consider as an example the torus diffeomorphism $\Gamma$ of the form $\mathbb{T} \ni x \to x+\varepsilon f(x)$. It is clear that if $x \in \mathbb{T}_a$, namely $x \in \mathbb{C}$ with $\mathrm{Re}(x) \in \mathbb{T}$ and $|\mathrm{Im}(x)| < a$, as in analytic cases, there is no reasons why $\Gamma$ should map $\mathbb{T}_a$ into itself. Another important remark is that the techniques to study the autonomous KdV, see [4], are deeply connected to the Hamiltonian structure.

**Plan of the Thesis.** The main purpose of the present thesis is to provide existence and linear stability of quasi-periodic solutions for a class of equations of the form

$$\mathrm{i}u_t = u_{xx} + \mathtt{f}(u, u_x, u_{xx}) \,, \quad x \in \mathbb{T} \,, \tag{10}$$

where $\mathtt{f}$ is a suitably smooth fully nonlinear nonlinearity. The general approach is similar to the one developed in [5], [4] for the KdV. We deal both with forced and autonomous cases, both in reversible and Hamiltonian setting. Finally we discuss both analytic and finite regularity results. Our results are stated in Theorem 1.1.1 for the forced reversible case, 1.1.2 for forced Hamiltonian case. Regarding the autonomous case we consider a reversible NLS and we discuss analytic solutions in Theorem 1.2.5 and Sobolev regularity in 1.2.6.

The main novelties with respect to the strategy of [4] are the following.

- we deal with analytic nonlinearities and prove the existence of analytic solutions. This is the first results of this type for fully nonlinear equations. Note that even though we have done this only in the case of the autonomous reversible NLS, recovering the same result for forced and/or Hamiltonian setting is completely straightforward at this point. See paragraph Analytic solutions on page `xxi`.

- we use a unified procedure valid both for finite regularity and analytic cases, since in our abstract theorem we rely only on properties of "tame" vector fields on scales of weighted Hilbert spaces. If one has exponential weights then the setting is analytic, otherwise it is Sobolev. Thus our algorithm is very flexible. As a drawback, in the analytic case, we do not use Cauchy estimates which may simplify some technical points; instead we use the fact that our functions are analytic on the complex domain $\mathbb{T}_a$ and Sobolev on the boundary.

- in the autonomous case we deal with a wide class of NLS equations where the leading term of the nonlinearity is cubic but contains derivatives up to the second order. Things would be significantly easier if we studied perturbations of the integrable cubic NLS, as it has been done for the autonomous KdV. In the construction of frequency-amplitude modulation we must deeply use a suitable argument of "genericity" of the nonlinearity and of the tangential frequencies. Such problem does not appear in perturbations of the integrable cubic NLS. See paragraph Weak Birkhoff normal form on page xvi.

- we consider a very general cubic term for our autonomous equation. It depends on several free parameters. By our techniques, we not only prove that for "generic" choices of such parameters then one has "infinitely" many choices of tangential sites, but we also give an explicit set of "resonant" parameters (see Definition 4.2.65) for which our method does not apply. In other words for such bad parameters we are not able to prove that there are choices of tangential sites for which the frequency-amplitude map is a diffeomorphism. It would be interesting to understand what happens in the case of such "null-forms". Anyway our class of fully nonlinear cubic terms cover the cases studied in the literature of semi-linear PDEs as well as many other quasi linear and fully non linear models.

- since the dispersion law of the NLS is even, in the Hamiltonian cases we have to deal with double eigenvalues. This is not trivial even is the simplest cases (bounded or semilinear nonlinearities) since the Kuksin Lemma in this context does not apply. See paragraph Inversion of the linearized operator on page xiv for a brief description of these first two points.

- in the autonomous case we consider reversible equations. This requires developing a more general KAM iteration algorithm. See paragraph Abstract KAM on page xix for an informal description.

- we deal with complex equation; this essentially brings only technical problems.

In the following we explain how the thesis is organized. In particular we focus on the novelty introduced and the differences we have with respect the works [5] and [4].

In Chapter 1 we state our main Theorems. In particular in Section 1.1 we present the first results we obtained on the *forced* NLS. In forced cases the non linearity $\mathtt{f}$ depends explicitly on time in a quasi-periodic way, i.e. $f(u, u_x, u_{xx}) = \mathtt{f}(\omega t, x, u, u_x, u_{xx})$ where $\omega \in \mathbb{R}^d$ for some $d \geq 1$ is the frequency vector of the forcing. We started by studying the forced case because here one does not have to handle the bifurcation equation. Hence we focused on the small divisors problems and extended the methods used in [5] to the NLS case. As main differences with respect to the KdV equations we mention the "weaker" dispersion law of the linear operator ($|j|^2$ instead of $|j|^3$) and that, due to its complex nature, the NLS is a system and not a scalar equation. We studied two cases: in Theorem 1.1.1 we analyze the NLS under some "reversibility" assumption, then in Theorem 1.1.2 we study NLS with Hamiltonian nonlinearities.

Regarding our reversibility condition (actually a very natural condition appearing in various works, starting from Moser [48]) some comments are in order. First of all some symmetry conditions are needed in order to have existence, because this excludes the presence of dissipative terms. Also such conditions guarantee that the eigenvalues of the linearized operator are all imaginary. All this properties could be imposed by using a Hamiltonian structure, however preserving the symplectic structure during the Nash-Moser iteration is not straightforward. Another property which follows by the reversibility is that the spectrum of the operator linearized at zero is simple, this is not true in the Hamiltonian case. Actually the reversibility conditions is given in terms of the parity of the functional (see Hypothesis 1). In other words we assume that some subspace of functions is invariant for the system. We selected a subspace of *odd* functions of the variable $x$. In some sense it is equivalent of working with Dirichelet boundary conditions because the linear operator in the Fourier space with the "sine" basis has single eigenvalues. It turns out that in this case there are no deep differences with the problems tackled in [5]. On the contrary many problems arise from the Hamiltonian structure and not only of technical nature. We briefly describe the general strategy we used to tackle the forced cases. First we note that in the forced cases the only unknown of the problem is the embedding of the torus $\mathbb{T}^d \ni \varphi \mapsto u(\varphi, x) \in H^s(\mathbb{T}^{d+1}; \mathbb{C})$ such that $u(\omega t, x)$ solves (10). Assuming that the forcing is small, the equation for the embedding reads

$$\mathrm{i}\omega \cdot \partial_\varphi u = u_{xx} + \varepsilon \mathtt{f}(\varphi, x, u, u_x, u_{xx}). \tag{11}$$

The proof of the existence of the embedding is based on an iterative scheme that produces "approximate" solutions at each step. Due to the presence of the quasi-periodic forcing with frequency $\omega$, the starting point (at $\varepsilon = 0$) is $u = 0$.

**Nash-Moser scheme.** The first ingredient is a generalized implicit function theorem with parameters (in our case the frequency $\omega$). In Section 3.1 of Chapter 3 we state such abstract Theorem. This is a well-established iterative scheme which allows to find zeros of a functional provided that one is able to prove the invertibility of its linearization in a neighborhood of the origin. This is fairly standard material. For instance in [11] the authors uses a similar algorithm. Anyway our abstract formulation allows us to apply the theorem to a wide class of operators. This formulation is based on a formal definition of *good parameters* where the algorithm runs through. In other words the Theorem (see Theorem 3.1.18) provides a "possibly empty" set of good parameters for which there is convergence of the sequence of approximate solutions to a true solution of the equation. We use this result in order to prove Proposition 4.0.46 in Section 4.

**Inversion of the linearized operator.** The second step is to study the invertibility of the linearized operator in order to give a more explicit formulation of the set of good parameters obtained at the previous step. An efficient way to prove bounds on the inverse of a linear operator is to diagonalize it: this is the so called reducibility. In the case of fully nonlinear NLS the linearized operator at some point

$u$ has the form

$$\mathcal{L}(u) = \omega \cdot \partial_\varphi \mathbb{1} + i(\mathbb{1} + A_2(\varphi, x))\partial_{xx} + iA_1(\varphi, x)\partial_x + iA_0(\varphi, x) \tag{12}$$

where $A_i$ are multiplication operators on the space $H^s \times H^s$ (see (4.1.14)) and $\mathbb{1}$ is the $2 \times 2$ identity matrix.

The key point is to control the differences of eigenvalues. This is the content of Proposition 4.1.51 in Section 4.1 for the reversible case. In the Hamiltonian case, where the eigenvalues of the linear operator are double, we cannot obtain the diagonalization. We only we obtain a reduction to a $2 \times 2$ block-diagonal matrix; this is the content of Proposition 5.0.85 in Chapter 5. In both cases the proof is divided in two steps:

1. Since we are dealing with unbounded non-linearities, before performing the diagonalization, we need to apply some changes of variables in order to reduce the operator to a constant coefficients unbounded operator plus a bounded remainder. This is a common feature of the above-mentioned literature and the reduction to constant coefficients of the differential operators of highest order can be iterated obtaining a constant coefficients unbounded operator plus a remainder which is regularizing of degree $k$. In the reversible case we set $k = 0$ as one can see in Lemmata 4.2.53 and 4.2.58 in Section 4.2. Concerning the Hamiltonian case we need to set $k = 1$. Indeed, due to the multiplicity of the unperturbed eigenvalues, we need a more precise control of their asymptotic. This analysis is performed in Lemmata 5.1.87 and 5.1.95. Some minor difficulties appear related to preserving the Hamiltonian structure.

2. The previous step gives a precise understanding of the eigenvalues of the matrix which we are diagonalizing. This allow us to impose the Second Mel'nikov conditions and to diagonalize by a linear KAM scheme. The conditions which we require are explicitly stated in Proposition 4.1.51 of Section 4.1 in the reversible case and in Theorem 5.0.85 in Chapter 5. In the latter case we obtain a block diagonal reduction to a $2 \times 2$ block diagonal time independent matrix. Note that the arguments of Step 1. allows us to obtain a complete diagonalization, that is a better result with respect to the one obtained by using the Kuksin Lemma. We remark also that, due to the multiplicity of the eigenvalues we need to require a weaker condition (see the definition of $\mathcal{O}_\infty^{2\gamma}$ in (5.0.4)). This is needed in order to perform the measure estimates and it is a matter exclusively related to the Hamiltonian structure and the periodic boundary conditions.

Once we have diagonalized the matrix, the bounds on the inverse follow from bounds on the eigenvalues, see Proposition 4.4.77 in Section 4.4 and Proposition 5.3.110. Now the sets of good parameters are rather explicit. They are formulated only in terms of the eigenvalues of the linearized operator $\mathcal{L}$

(see for instance formulæ(4.1.23) and 4.4.208). Before discussing about the measure estimates of the set of parameters we want to remark which is the core of the steps we just analyzed.

The key idea to deal with fully nonlinear equations stands in Step 1. In such procedure one uses techniques different from the standard KAM idea (such as the Kuksin Lemma). The important issue is that our methods deeply rely on the structure of pseudo-differential operator of $\mathcal{L}$ in (4.1.29). Such particular structure comes form the fact that $\mathcal{L}$ is the linearized operator of a PDE in which the non-linearity $\mathtt{f}$ is a composition operator. The transformation of coordinates used in Step 2 destroys the particular structure of $\mathcal{L}$. This is the reason why we follow a Nash-Moser scheme. Indeed we just use the fact that in a special coordinate system we are able to give estimates on the inverse of $\mathcal{L}$, but we never changes the coordinates of the whole system. This is the advantages of the Nash-Moser with respect to the KAM approach. It is automatic to preserve any structure of the initial system at each step of the iteration. This issue has to be taken into account also when studying the autonomous case.

**Measure estimates.** As last step of our strategy we perform measure estimates on the sets of good parameters in order to ensure that for "most" parameters $\omega$ we are able to find a quasi-periodic solution with such frequency vector. There are two principal issues concerning such measure estimates. Essentially, following our scheme one construct at the $n$-th step a set $\mathcal{G}_n$ of parameters which depends on the approximate solution $u_n$. The first matter is to show that each $\mathcal{G}_n$ has measure $1-O(\varepsilon)$. The first basic requirement is to prove that we may impose each *single* non-resonance condition by only removing a small set of parameters. This is relatively simple in the reversible case, while in the Hamiltonian case this is a non trivial problem which we overcome by imposing a non-degeneracy condition (see Hypothesis 3) and by considering vectors $\omega$ as in (1.1.7) instead of (1.1.2).

Secondly one needs to study the dependence of the Cantor sets on the function $u_n$. Indeed in principle as $n$ varies this sets are unrelated and then the intersection might be empty. Indeed $\mathcal{G}_n$ is constructed by imposing **infinitely** many Mel'nikov conditions. We show that such infinitely many conditions imply **finitely** many second Mel'nikov conditions on a neighborhood of $u_n$. This implies a sort of "summability" condition and allows us to show that the union of the resonant sets is still small. This analysis is performed in Section 4.5 for reversible NLS an in Section 5.4 for the Hamiltonian case. This is the most delicate part where substantial new ideas with respect to reversible case, are needed.

Since the $u_n$ are a rapidly converging Cauchy sequence this proposition allows us to prove that $\mathcal{G}_\infty = \cap_{n\geq 0}\mathcal{G}_n$ has asymptotically full measure. The strategy described above is similar to that followed in [5] and [31]. It is quite general and can be applied to various case. The main differences are in the proof of Proposition 4.1.51 (or 5.0.85). Clearly it depends on the unperturbed eigenvalues and on the symmetries one ask for on the system.

Chapter 6 is devoted to the proof of Theorems 1.2.5 and 1.2.6 which are about the more interesting case of autonomous NLS i.e. equation (10) with `f` which does not depend on time. On one hand one can use the result obtained in Chapters 4 on the forced case. On the other hand in the autonomous case there are no external parameters in the nonlinearity that fixed the frequency vector $\omega$. So that, by following straightforward the same scheme used for forced cases, one cannot exclude that the solution of the problem is $u = 0$. In other words cannot choose $u = 0$ as starting point. Hence some preliminary steps are required. In [4] the authors studied the autonomous KdV, and exploited the Hamiltonian structure of the equation in order to reduce the autonomous case to the forced one, using arguments developed by Berti-Bolle in [9].

Here we study a "reversible" equation and moreover we are interested in looking for *analytic* solutions for the NLS in (10) with an analytic `f`. Actually Theorem 1.2.5 is the first results about the existence of analytic quasi-periodic solutions for fully nonlinear PDEs. Indeed using the techniques used in [4] one cannot find analytic solutions and in moreover one cannot deal with reversible systems. The core of our proof relies on the abstract Theorem (3.2.39) stated, in Section 3.2, for general "tame" vector fields and not on the symplectic geometry arguments developed in [9].

Anyway in the autonomous case the proof is sophisticated and involves many different arguments. The structure of Chapter 6 follows essentially the structure if the proof. Here we give an overview of the proof by referring the reader to the main Propositions proved in the Sections of Chapter 6. Here we want to remark the differences between the autonomous case and the forced one and in the following the differences between our approach and the one followed in [4].

An important feature of the autonomous equation (10) is that the NLS is completely resonant near $u = 0$. Note that this is not true for equation (5). In other words here we have that all the linear solutions of (10) are periodic. This means that we are looking for a quasi-periodic solution close to some *periodic* solution

$$u(t,x) = \sum_{i=1}^{d} u_{\mathtt{v}_i} e^{\mathrm{i}\mathtt{v}_i^2 t} e^{\mathrm{i}\mathtt{v}_i x} \tag{13}$$

of the linear equation which involves $d$ frequencies in the set $S = \{\mathtt{v}_1, \ldots, \mathtt{v}_d\} \subset \mathbb{Z}$. This means also that the existence of quasi-periodic solution is purely a non linear phenomenon. Due to the presence *resonances* it can happen that, for specific choices of the sites $S$, the behavior of the solutions of the non linear equation differ drastically from the one of the linear equation (see for instance [23]). In order to avoid such phenomena one has to restrict to "generic" choices of $S$, in the sense of Definition 1.2.4.

**Weak Birkhoff normal form.** As preliminary step one looks for an approximate solution for the NLS 1.2.16 which will be the starting point for an iterative algorithm. We said that a linear approximation is not enough. Hence in Section 6.1 we find an solution of the third order approximate

NLS. This is done in the following way. One first rewrites the NLS as a dynamical system

$$\dot{u} = \chi(u) = \sum_{j\in\mathbb{Z}} \chi_j(u)\partial_{u_j} \tag{14}$$

where $\chi$ is vector field defined on the space of sequences $u = \{u_j\}_{j\in\mathbb{Z}} \in h^{a,p}$. In our case $h^{a,p} \leftrightarrow H^p(\mathbb{T}_a;\mathbb{C})$ the functions analytic on the toroidal domain $x \in \mathbb{C}$ such that $\text{Re}(x) \in \mathbb{T}$ and $|\text{Im}(x)| \leq a$, for some $a > 0$. For a precise definition see formula (2.1.16) in Section 2.1. Once selected, the sites $S \subset \mathbb{Z}$ decompose the space of sequences $h^{a,p}$ into two orthogonal subspaces $u = (v, z) := \big(\{u_k\}_{k\in S}, \{u_k\}_{k\in S^c}\big)$. Then one looks for a coordinates system such that $\{z = 0\}$ is an invariant manifold of solutions for the approximate NLS. More precisely one splits the vector field in (14) as $\chi = \Pi_S\chi\partial_v + \Pi_{S^c}\chi\partial_z$, hence, through a step of Birkhoff normal form, one removes all the cubic terms $O(v^3), O(v^2 z)$ from $\Pi_S\chi$ and all the term $O(v^3)$ form $\Pi_{S^c}\chi$ that do not commute with the linear part. This implies that the dynamics on $z = 0$ is integrable for the third order approximation of the new vector field. We perform this step in Proposition 6.1.124 of Section 6.1. It turns out that such transformation is close to the identity up to a *finite* dimensional operator.

In principle one could remove also the term $O(v^2 z)$ from $\Pi_{S^c}\chi$ by performing a stronger normal form. In classical KAM Theory for semi-linear NLS this is actually done. See for instance [52]. With a complete normal form one can completely diagonalize the third order term of the vector field. Here we do not do this due to the following question.

**Structure.** We see in (2.1.19) (se also (4.1.29)) that the linearized operator of an unbounded vector field as $\chi$ in (2.1.16) is a non constant coefficients pseudo-differential operator. In the forced cases we analyzed the spectrum of similar linear operator in Proposition 4.1.51. To prove the invertibility of the linearized operator in the autonomous case we need to use similar arguments. The proof of Theorem 4.1.51 strongly relies on the particular differential structure of $\mathcal{L}$. In order to perform a step of Birkhoff normal form as in [52] one has that the map is close to the identity up to a bounded operator. The advantage of using the map defined in Proposition 6.1.124 is that the linearized operator of the field is modified only up to a finite rank operator. Hence the spectral analysis is essentially the same as the analysis in the original coordinates. Hence we can use the same arguments used for the forced case.

**Action-angles variables.** We have underlined that in autonomous cases there are no external parameters to modulate in order to fulfill non-degeneracy conditions. Now thanks to the step of weak Birckhoff norml form we selected an approximatively invariant manifold where the dynamics is integrable and non-isocronous. On this manifolds we introduce action-angle variables, on the tangential sites (i.e. the sites in $S$), in such a way one can use the initial data as parameters which will be denoted by $\xi$. This is done in Section 6.2 using the change of coordinates in (6.2.9) one obtains (see (6.2.14)) a system like

$$\begin{cases} \dot{\theta} = \omega(\xi) + F^{(\theta)}(\theta, y, w) \\ \dot{y} = F^{(y)}(\theta, y, w) \\ \dot{z} = \Omega(\theta; \xi) z + F^{(w)}(\theta, y, z), \end{cases} \tag{15}$$

with $F$ small in some suitable sense. The system in (15) are of the form considered in the abstract Theorem in (3.2.39). Some comments are required.

The tangential frequencies $\omega(\xi)$ are given by the frequency-amplitude modulation in (6.2.16). Here one can see one the most important issue of autonomous cases with respect to the forced one. Indeed one has that the frequency vector $\omega$ is $O(|\xi|)-$close to integer vectors as $\xi \to 0$ hence the perturbation theory is more difficult. Indeed we need a diophantine frequency, i.e. which satisfies a bounds like

$$|\omega(\xi) \cdot l| \geq \frac{\gamma}{1 + |l|^{\tau}}, \quad \text{for some} \quad \tau > 0, \tag{16}$$

and for $\gamma \sim O(|\xi|)$.

Moreover in usual KAM Theorems the whole normal form $\mathcal{N} := \omega(\xi) \cdot \partial_\theta + \Omega(\theta, \xi) z \partial_z$ is reduced to constant coefficients. Here the operator $\Omega(\theta, \xi)$ depends on the angles $\theta$ since we did not remove the terms $O(v^2 z)$ from the vector field in the normal directions. This means that in $\Omega(\theta, \xi)$ there are some terms which, are not "perturbative" with respect to the size of the small divisor $\gamma$. These term are dealt in Lemma 6.4.142 of Section 6.4.2 where we study the invertibility of the linearized operator in the normal directions. Such terms are essentially the non-resonant terms we did not cancel with the initial Birkhoff step.

A crucial point is the so called "twist" condition with respect to the parameters $\xi$. What we need to check is that if one "moves" the initial data $\xi$ then the frequencies move in a non "trivial" way. We firstly prove that for the tangential frequency the map $\xi \to \omega(\xi)$ is a diffeomorphism. Then we prove that also the normal frequencies satisfy an appropriate twist condition. These terms are given by the average in $\theta$ of $\Omega(\theta, \xi)$. The analysis of the last issue is performed in Lemmata 6.6.157 and 6.6.158 in Section 6.6. Note that this is a delicate question, since we are requiring a modulation of infinitely many normal frequencies by only finitely many parameters. The analysis would be much simpler if one considers a fully nonlinear perturbation, of order at least four, of the cubic integrable NLS. In such a case, for *any* choices of the tangential site in $S$, one would obtain that the map $\xi \to \omega(\xi)$ is a diffeomorphism by exploiting the integrability properties of the system. Here we need to introduce a notion of "genericity" (see Definition 1.2.4) which implies that for "most" choices of the cubic terms and "most" choices of the tangential sites the frequencies satisfy a "twist" condition. Interestingly we can produce explicitly non generic choices of cubic non linearities such that for *any* choice of tangential sites the twist condition is false. In particular it turns out that the Jacobian of the map $\xi \mapsto \omega(\xi)$ has at most rank 2. It would be interesting to investigate whether quasi periodic solutions exist for such "degenerate" cases.

The previous steps are not trivial at all, and depends deeply on the equation and on the domain on which one works. Now we briefly comment the strategy we follow in order to find a torus embedding for a general vector field of the form (15), which is one of the most important differences with respect to the approach followed in [4].

**Abstract KAM.** The existence of an invariant torus means that it is possible to describe the system in coordinates $(\theta, y, w)$ adapted to such torus, i.e. such that the system has the form

$$\begin{cases} \dot{\theta} = \omega + g^{(\theta)}(\theta, y, w) \\ \dot{y} = g^{(y)}(\theta, y, w) \\ \dot{w} = g^{(w)}(\theta, y, w) \end{cases} \tag{17}$$

with $g^{(\mathtt{v})}(\theta, 0, 0) \equiv 0$ for $\mathtt{v} = \theta, y, w$. Of course at this level of course there is no particular reason for separating the variables $y, w$; however in all applications the variables $y$ naturally appear as variables "conjugated" to $\theta$: in the Hamiltonian setting they come from the symplectic structure and in the reversible setting they are characterized by the fact that $\langle g^{(y)}(\cdot, y, w)\rangle \equiv 0$.

As said, possibly after the Birkhoff procedure one typically obtain a system like (15) which has only an "approximate" invariant torus since $F^{(\mathtt{v})}(\theta, 0, 0)$ are "small".

Thus the idea, which goes back to Moser [48], is to find a change of coordinates such that in the new coordinates the system (15) takes the form (17). Precisely one has to consider a sequence of changes of variables approximating better and better (17).

We have already said that an important feature of Moser's scheme is that the operator $\Omega(\theta; \xi)$ needs to have distinct eigenvalues with suitable lower bound on the distance between two distinct eigenvalues in order to be diagonalized. Under such condition it is possible to show that in the final coordinates the system assume the form

$$\begin{cases} \dot{\theta} = \widetilde{\omega}(\xi) + \widetilde{F}^{(\theta)}(\theta, y, w) \\ \dot{y} = \widetilde{F}^{(y)}(\theta, y, w) \\ \dot{w} = \widetilde{\Omega}(\xi) w + \widetilde{F}^{(w)}(\theta, y, w) \end{cases} \tag{18}$$

not only with $\widetilde{F}^{(\mathtt{v})}(\theta, 0, 0) \equiv 0$ for $\mathtt{v} = \theta, y, w$ but also quadratic in $y, w$: in turn this gives informations about the linear stability of the invariant torus $(\theta, 0, 0)$.

However in the fully nonlinear case we need to preserve some properties of the initial vector field, for instance the fact that the vector field comes from a Nemytskii operator, and it is well known that the change of variables which diagonalizes $\Omega(\theta; \xi)$ does not preserve such structure.

The aim of the Theorem 3.2.39 is Section 3.1 is to show that in principle there is in fact no need to require the second order Mel'nikov conditions.

Given a dynamical system it is very natural to look for a set of variables in which the dynamics is "easier" to study and describe: of course also being "easier" heavily depend on the aspect of the dynamic one is willing to study. When looking for invariant tori, one has to perform an iterative scheme: at each step, the goal is to find a change of variables close to the identity such that the term constant in $y, w$ is "reduced drastically" (in some suitable sense) so that after infinitely many steps it vanishes. If one simply wants to reduce (15) to (17) the natural thing to consider are the translations, i.e. one may start looking for a change of coordinates $(\theta, y, w) \mapsto (\theta + h^{(\theta)}(\theta), y + h^{(y)}(\theta), w + h^{(w)}(\theta))$ with $h^{(\mathtt{v})}$ suitably small, in order to simply eliminate the dependence on the angular variable; however in order to find such a change of coordinates one needs to invert the whole linearized operator about $(\theta, 0, 0)$, which may be a too hard task. Another possibility is to consider a rototraslation in such a way that the rotation makes the linearized operator diagonal and hence "easier" to invert. However, a part from the fact that the rotation is the one involving the second order Mel'nikov condition and there are case in which one cannot impose them, as said also when one is able to impose the second Mel'nikov conditions, after rotating the $w$-variables in principle one loses the Töplitz structure. The compromise is to "rotate" the operator only in the $y$-variables; precisely one consider changes of variables which are merely translations in the $\theta$,$w$-components and contain a term linear in $y, w$ in the $y$-component. Such changes of coordinates are obtained as time-1 flow associated with vector fields in a suitable subspace of "almost-identical" vector fields that we shall denote $\mathcal{A}$; see (3.2.53) below. In general, under any change of coordinates $\Phi$ generated by a vector field $G$, a vector field $f$ is changed into

$$\Phi_*(f) = f + [G, f] + O(G^2)$$

so if we look for $\Phi$ such that $\Pi_{\mathcal{A}}(\Phi_* f) = 0$ up to a quadratic remainder (which in the scheme will converge to zero very fast), we need to find $G$ such that

$$\Pi_{\mathcal{A}}(f + [G, f]) = 0\,;$$

in other words we need to invert the operator $\Pi_{\mathcal{A}}[f, \cdot]$. If, as said, we confine to the case $G \in \mathcal{A}$ it is natural to decompose $f$ as

$$f = X + X^{\perp}\,, \qquad X \in \mathcal{A},\ X^{\perp} \in \mathcal{A}^{\perp}.$$

since $[X, G] = 0$ for $G \in \mathcal{A}$, thus we need to invert the operator $\mathcal{L} := \Pi_{\mathcal{A}}([X^{\perp}, \cdot])$. Then it is convenient to consider a "degree decomposition" of $X^{\perp}$, namely writing

$$X^{\perp} = N + R$$

with $N, R$ being the ($\leq 0$)-degree and the ($\geq 2$)-degree vector fields respectively. The second order Mel'nikov conditions allow to diagonalize the operator $\mathfrak{D} = \mathfrak{D}(\xi) := \Pi_{\mathcal{A}}([N, \cdot])$ and deduce the invertibility of $\mathcal{L}$ by Neumann series. We show that the invertibility of $\mathfrak{D}$ implies the invertibility of $\mathcal{L}$, because $\mathfrak{R} := \Pi_{\mathcal{A}}([R, \cdot])$ is nihilpotent and it commutes with $\mathfrak{D}$, thus the minimal condition to require on the parameters seems to be the invertibility of $\mathfrak{D}$. The invertibility of $\mathfrak{D}$ rely on the following two important properties which has to be satisfied at each step of the iteration:

- the frequency vector $\omega$ must be diophantine;
- the operator $\omega \cdot \partial_\theta + \Omega(\theta, \xi)$ must be "approximatively invertible".

In classical KAM scheme in the second item one has that the operator $\Omega$ is approximatively diagonal. In our approach we do not have such property, but on the other hand we just use transformation which are close to the identity up a a *finite* rank operator. This implies that, at each step, the linearized operator in the normal directions has the form (4.1.29) up to a finite rank operator. In this way, in order to invert it, we are allowed to use all the techniques we used in studying the forced cases. Actually the abstract theorem 3.2.39 provides a general methods that allows us to pass from the forced case to the autonomous one. This is done without exploiting any Hamiltonian structure, as done in [4]. We remark again that, in applications, one must be able to perform the preliminary step of Birkhoff normal form, otherwise the result provided by our abstract algorithm should be empty.

**Analytic solutions.** An important point is that our algorithm is in some sense "stable" under slight modifications of the scheme we have just described. We have already said that the invertibility of a matrix does not depend on the coordinates one choose. On the other hand lots of estimates we perform actually depend on the coordinate system. Hence it can be convenient to have the possibility of choosing appropriate coordinates. The iterative scheme can be changed, without affecting the convergence, by applying at each step linear transformations on the normal variables. This is actually the rôle of the transformations $\mathcal{T}$ introduced in (3.2.66) in Section 3.2. The key point is that such map $\mathcal{T}$ *must preserve* any kind of structure one needs. This clearly depends on the particular problem that one studies, in our case the pseudo-differential structure of the linearized operator. Just to fix the ideas one can think that, in the case of bounded nonlinearities, one can choose the maps $\mathcal{T}$ as the maps which diagonalize $\Omega$. Note that in this way one recovers the results of classical KAM theory for semilinear PDEs. This question is analyzed in Section 6.4.1 where we study what kind of linear maps $\mathcal{T}$ are allowed in our case. Thanks to good choices of such maps $\mathcal{T}$, we are able to show that the loss of analyticity at each step of the iteration is exponentially small, and hence to obtain a solution that is analytic in slightly smaller domain.

**Analytic and differentiable cases.** In Theorems 1.2.5 and 1.2.6 we deal with analytic and differentiable nonlinearities respectively. We remark that we work on spaces of functions which are analytic in the complex toroidal domain $\mathcal{T}_a$ and Sobolev on the boundary. In the abstract Theorem 3.2.39 we never exploit any analytic properties, but we just use properties of "tameness" of the vector fields, exactly as done in differentiable cases like [4], or in Chapter 4 and 5. Hence one obtain differentiable the result in Theorem 1.2.6 just following word by word the proof of Theorem 1.2.5.

# Contents

**Abstract** **i**

**Introduction** **iii**

**1 Main Results** **1**

1.1 The forced equation . . . . . . . . 1

1.2 The autonomous equation . . . . . . . . 4

**2 Functional Setting** **9**

2.1 Scales of Sobolev spaces . . . . . . . . 9

2.2 Reversible structure. . . . . . . . . 13

2.3 Hamiltonian structure . . . . . . . . 14

**3 Abstract Nash-Moser Theorems** **19**

3.1 An Abstract Existence Theorem for forced equations . . . . . . . . 19

3.1.1 Nash-Moser scheme . . . . . . . . 19

3.2 An Existence Theorem for torus embeddings . . . . . . . . 26

3.2.1 The Phase Space . . . . . . . . 26

3.2.2 The iterative scheme . . . . . . . . 39

**4 Forced NLS: reversible case** **51**

4.1 Reversible Operators . . . . . . . . 52

4.1.1 Proof of Proposition 4.0.46 . . . . . . . . 53

4.2 The diagonalization algorithm: regularization . . . . . . . . 57

4.3 The diagonalization algorithm: KAM reduction . . . 67
4.3.1 Functional setting and notations . . . 69
4.3.2 Reduction Algorithm . . . 72
4.4 Conclusion of the diagonalization algorithm . . . 85
4.4.1 Proof of Proposition 4.1.51 . . . 85
4.4.2 Inversion of $\mathcal{L}$ . . . 87
4.5 Measure estimates and conclusions . . . 89
4.5.1 Proof of Theorem 1.1.1 . . . 93

**5 Forced NLS: Hamiltonian case** **95**
5.1 Regularization of the linearized operator . . . 97
5.1.1 Non-degeneracy Condition . . . 111
5.2 Reduction to constant coefficients . . . 113
5.2.1 Reduction algorithm . . . 116
5.2.2 Conclusions and Proof of Theorem 5.2.96 . . . 127
5.3 Inversion of the linearized operator . . . 130
5.4 Measure estimates . . . 132
5.4.1 Proof of Theorem 1.1.2 . . . 140

**6 Autonomous NLS** **143**
6.1 Weak Birkhoff Normal Form . . . 143
6.2 Action-angles variables . . . 147
6.3 Inizialization . . . 150
6.4 Pseudo-differential Vector Fields . . . 154
6.4.1 Regularization . . . 154
6.4.2 Inversion of the linearized operator in the normal directions . . . 169
6.5 The sets of “good” parameters . . . 182
6.6 Measure estimates . . . 191
6.6.1 The “twist” condition . . . 192
6.6.2 The estimates of “good” parameters . . . 197

**A General Tame and Lipschitz estimates** **203**
A.1 Smooth functions and vector fields on the torus . . . 206

**B Properties of tame vector fields** **211**

**C Proofs** **223**
C.1 Proof of Lemmata 4.5.84 and 4.5.83 . . . . . 223
C.2 Proof of Lemma 2.3.15 . . . . . 225
C.3 Proof of Lemma 3.2.43 . . . . . 225

**Bibliography** **229**

# 1. Main Results

In this introductive Section we state our main results. We prove the existence and the stability of quasi-periodic solutions for the NLS in different contexts. First on *forced* equations, both in the *reversible* and *Hamiltonian* cases. Hence in Section 1.2 we present the results obtained on the *autonomous* reversible NLS.

## 1.1 The forced equation

In the first part of the thesis we study a class of forced fully non linear Schrödinger equations of the form

$$iu_t = u_{xx} + \varepsilon \mathtt{f}(\omega t, x, u, u_x, u_{xx}), \quad x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}, \tag{1.1.1}$$

where $\varepsilon > 0$ is a small parameter, the nonlinearity is quasi-periodic in time with diophantine frequency vector

$$\omega := \lambda\bar{\omega}, \quad \lambda \in \Lambda := \left[\frac{1}{2}, \frac{3}{2}\right] \subset \mathbb{R}, \;\; |\bar{\omega} \cdot \ell| \geq \frac{\gamma_0}{|\ell|^{\tau_0}}, \; \forall \, \ell \in \mathbb{Z}^d \backslash \{0\}. \tag{1.1.2}$$

For instance one can fix $\tau_0 = d+1.$ and $\mathtt{f}(\varphi, x, z)$, with $\varphi \in \mathbb{T}^d$, $z = (z_0, z_1, z_2) \in \mathbb{C}^3$ is in $C^q(\mathbb{T}^{d+1} \times \mathbb{C}^3; \mathbb{C})$ in the real sense (i.e. as function of $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$). For this class we prove existence and stability of quasi-periodic solutions with Sobolev regularity for any given diophantine vector $\bar{\omega}$ and for all $\lambda$ in an appropriate positive measure Cantor-like set.

A quasi-periodic solution, with frequency $\omega \in \mathbb{R}^d$, for an equation such as (1.1.1) is a function of the form $\mathtt{u}(t, x) = u(\omega t, x)$ where

$$u(\varphi, x) : \mathbb{T}^d \times \mathbb{T} \to \mathbb{C}.$$

In other words we look for non-trivial $(2\pi)^{d+1}-$periodic solutions $u(\varphi, x)$ of

$$i\omega \cdot \partial_\varphi u = u_{xx} + \varepsilon \mathtt{f}(\omega t, x, u, u_x, u_{xx}) \tag{1.1.3}$$

in the Sobolev space

$$H^s := H^s(\mathbb{T}^d \times \mathbb{T}; \mathbb{C}) := \{u(\varphi, x) = \sum_{(\ell,k) \in \mathbb{Z}^d \times \mathbb{Z}} u_{\ell,k} e^{i(\ell \cdot \varphi + k \cdot x)} : ||u||_s^2 := \sum_{i \in \mathbb{Z}^{d+1}} |u_i|^2 \langle i \rangle^{2s} < +\infty\}, \tag{1.1.4}$$

where $s > \mathfrak{s}_0 := (d+2)/2 > (d+1)/2$, $i = (\ell, k)$ and $\langle i \rangle := \max(|\ell|, |k|, 1)$, $|\ell| := \max\{|\ell_1|, \dots, |\ell_d|\}$. For $s \geq \mathfrak{s}_0$ $H^s$ is a Banach Algebra and $H^s(\mathbb{T}^{d+1}) \hookrightarrow C(\mathbb{T}^{d+1})$ continuously. We impose the *reversibility condition*

**Hypothesis 1.** *Assume that* $\mathtt{f}$ *is such that*

*(i)* $\mathtt{f}(\varphi, -x, -z_0, z_1, -z_2) = -\mathtt{f}(\varphi, x, z_0, z_1, z_2)$.

*(ii)* $\mathtt{f}(-\varphi, x, z_0, z_1, z_2) = \overline{\mathtt{f}(\varphi, x, \bar{z}_0, \bar{z}_1, \bar{z}_2)}$,

*(iii)* $\mathtt{f}(\varphi, x, 0) \neq 0\,, \quad \partial_{z_2}\mathtt{f} \in \mathbb{R}\backslash\{0\}$,

*where* $\partial_z = \partial_{\mathrm{Re(z)}} - i\,\partial_{\mathrm{Im(z)}}$.

Chapter 4 is devoted to the proof of the following Theorem.

**Theorem 1.1.1.** *There exist* $s := s(d) > 0$, $q = q(d) \in \mathbb{N}$ *such that for every nonlinearity* $\mathtt{f} \in C^q(\mathbb{T}^{d+1} \times \mathbb{C}^3; \mathbb{C})$ *that satisfies Hypothesis 1 and for all* $\varepsilon \in (0, \varepsilon_0)$, *with* $\varepsilon_0 = \varepsilon_0(\mathtt{f}, d)$ *small enough, there exists a Cantor set* $\mathcal{C}_\varepsilon \subset \Lambda$ *of asymptotically full Lebesgue measure, i.e.*

$$|\mathcal{C}_\varepsilon| \to 1 \quad as \quad \varepsilon \to 0, \tag{1.1.5}$$

*such that for all* $\lambda \in \mathcal{C}_\varepsilon$ *the perturbed NLS equation (1.1.3) has solution* $u(\varepsilon, \lambda) \in H^s$ *such that* $u(t,x) = \bar{u}(-t, x)$ *and* $u(t,x) = -u(t,-x)$ *with* $||u(\varepsilon,\lambda)||_s \to 0$ *as* $\varepsilon \to 0$. *In addition,* $u(\varepsilon, \lambda)$ *is* linearly stable.

Regarding our reversibility condition (actually a very natural condition appearing in various works, starting from Moser [48]) some comments are in order. First of all some symmetry conditions are needed in order to have existence, in order to exclude the presence of dissipative terms. Also such conditions guarantee that the eigenvalues of the linearized operator are all imaginary. All this properties could be imposed by using a Hamiltonian structure, however preserving the symplectic structure during our Nash-Moser iteration is not straightforward. Another property which follows by the reversibility is that the spectrum of the operator linearized at zero is simple, this is not true in the Hamiltonian case, see [30]. On the contrary, the spectrum the linearized Schrödinger equation at zero, in the Hamiltonian case, has double eigenvalues, hence there are a number of difficulties not only of a technical nature.

Here we have also considered the equation

$$iu_t = u_{xx} + \mathtt{m}u + \varepsilon\mathtt{f}(\omega t, x, u, u_x, u_{xx}), \quad x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}, \tag{1.1.6}$$

$m > 0$ and the nonlinearity is quasi-periodic in time with the frequency

$$\omega \in \Lambda := \left[\frac{1}{2}, \frac{3}{2}\right]^d \subset \mathbb{R}^d, \ \ |\omega \cdot \ell| \geq \frac{\gamma_0}{|\ell|^{\tau_0}}, \ \forall\, \ell \in \mathbb{Z}^d \backslash \{0\}. \tag{1.1.7}$$

The difference with respect to definition 1.1.2 is that here we have $d$ parameters, while for the reversible case we fixed a diophantine direction $\bar{\omega}$ and we use only a one-dimensional parameters. Again we work in differentiable class and we assume that $\mathtt{f}(\varphi, x, z)$, with $\varphi \in \mathbb{T}^d$, $z = (z_0, z_1, z_2) \in \mathbb{C}^3$ is such that

$$\mathtt{f}(\varphi, x, u, u_x, u_{xx}) = f_1(\varphi, x, \xi, \eta, \xi_x, \eta_x, \xi_{xx}, \eta_{xx}) + i f_2(\varphi, x, \xi, \eta, \xi_x, \eta_x, \xi_{xx}, \eta_{xx}),$$

where we set $u = \xi + i\eta$, with $\xi(\varphi, x), \eta(\varphi, x) \in H^s(\mathbb{T}^{d+1}; \mathbb{R})$ for some $s \geq 0$, and where

$$f_i(\varphi, x, \xi_0, \eta_0, \xi_1, \eta_1, \xi_2, \eta_2) : \mathbb{T}^{d+1} \times \mathbb{R}^6 \to \mathbb{R}, \quad i = 1, 2. \tag{1.1.8}$$

is $C^q$ for some $q \in \mathbb{N}$ large enough. In this case the equation for $u(\varphi, x) : \mathbb{T}^d \times \mathbb{T} \to \mathbb{C}$ reads

$$i\omega \cdot \partial_\varphi u = u_{xx} + \mathtt{m} u + \varepsilon \mathtt{f}(\varphi, x, u, u_x, u_{xx}) \tag{1.1.9}$$

On the non linearity $\mathtt{f}$ we assume the following:

**Hypothesis 2.** *Assume that* $\mathtt{f}$ *is such that*

$$\mathtt{f}(\omega t, x, u, u_x, u_{xx}) = \partial_{\bar{z}_0} G(\omega t, x, u, u_x) - \frac{d}{dx}[\partial_{\bar{z}_1} G(\omega t, x, u, u_x)] \tag{1.1.10}$$

*with* $\partial_{\bar{z}_i} = \partial_{\xi_i} + i\partial_{\eta_i}$, $i = 0, 1$, *and*

$$G(\omega t, x, u, u_x) := F(\omega t, x, \xi, \eta, \xi_x, \eta_x) : \mathbb{T}^{d+1} \times \mathbb{R}^4 \to \mathbb{R}, \tag{1.1.11}$$

*of class* $C^{q+1}$.

**Hypothesis 3.** *Assume that* $\mathtt{f}$ *is such that*

$$\frac{1}{(2\pi)^{d+1}} \int_{\mathbb{T}^{d+1}} (\partial_{\bar{z}_1} \mathtt{f})(\varphi, x, 0, 0, 0) dx d\varphi = \mathfrak{e} \neq 0. \tag{1.1.12}$$

Hypothesis 3 si quite technical and we will see in the following where we need it. On the contrary Hypothesis 2 si quite natural and it implies that the equation (1.1.6) can be rewritten as an Hamiltonian partial differential equation

$$u_t = i\partial_{\bar{u}} \mathcal{H}(u), \quad \mathcal{H}(u) = \int_{\mathbb{T}} |u_x|^2 + \mathtt{m}|u|^2 + \varepsilon G(\omega t, x, u, u_x) \tag{1.1.13}$$

with respect to the non-degenerate symplectic form

$$\Omega(u, v) := \mathrm{Re} \int_{\mathbb{T}} i u \bar{v} dx, \quad u, v \in H^s(\mathbb{T}^{d+1}; \mathbb{C}), \tag{1.1.14}$$

where $\partial_{\bar{u}}$ is the $L^2-$gradient with respect the complex scalar product. The main result proved in Chapter 5 is the following.

**Theorem 1.1.2.** *There exist $s := s(d, \tau_0) > 0$, $q = q(d) \in \mathbb{N}$ such that for every nonlinearity $\mathtt{f} \in C^q(\mathbb{T}^{d+1} \times \mathbb{R}^6; \mathbb{C})$ that satisfies Hypotheses 2 and 3 if $\varepsilon \le \varepsilon_0(s, d)$ small enough, then there exists a Lipschitz map*

$$u(\varepsilon, \lambda) : [0, \varepsilon_0] \times \Lambda \to H^s(\mathbb{T}^{d+1}; \mathbb{C})$$

*such that, if $\lambda \in \mathcal{C}_\varepsilon \subset \Lambda$, $u(\varepsilon, \lambda)$ is a solution of (1.1.3). Moreover, the set $\mathcal{C}_\varepsilon \subset \Lambda$ is a Cantor set of asymptotically full Lebesgue measure, i.e.*

$$|\mathcal{C}_\varepsilon| \to 1 \quad as \quad \varepsilon \to 0, \tag{1.1.15}$$

*and $||u(\varepsilon, \lambda)||_s \to 0$ as $\varepsilon \to 0$. In addiction, $u(\varepsilon, \lambda)$ is* linearly stable.

Theorem 1.1.2 is the equivalent result of Theorem (1.1.1) in the Hamiltonian case. Here we just changing the equation by introducing a "mass" $\mathtt{m} > 0$. This slightly simplify the problem and allows us to concentrate on the problems given by the multiplicity of the eigenvalues.

## 1.2 The autonomous equation

We started to study fully non linear equations in the forced case. The presence of the forcing simplify the bifurcation problem, and hence we focused on the small divisor problems, which, due to the presence of derivatives in the non linearity, is more difficult with respect to classical bounded cases. Then we considered autonomous equation and we studied it by using the results obtained on the forced case. In particular we consider the equation

$$u_t = -i(u_{xx} + \mathtt{f}(u, u_x, u_{xx}))\,, \quad x \in \mathbb{T}\,. \tag{1.2.16}$$

The non linearity $\mathtt{f}$ is a gauge preserving and $x$-independent function of the form

$$\mathtt{f}(u, u_x, u_{xx}) := f(u, u_x, u_{xx}) + g(u, u_x, u_{xx}) \tag{1.2.17}$$

where $f$ is the homogeneous components of degree 3 and $g$ contains all terms of higher order.

We will consider two cases

1. $g$ is analytic as function $\mathbb{C}^3 \to \mathbb{C}$ in the ball of radius $\mathtt{r}_0$. Then we fix $\mathfrak{a} > 0$ and extend (1.2.16) to $x \in \mathbb{T}_{\mathfrak{a}}$. Here $\mathbb{T}_a$ is the compact subset of the complex torus $\mathbb{T}_{\mathbb{C}} := \mathbb{C}/2\pi\mathbb{Z}$ with $x \in \mathbb{C}$ and $|\mathrm{Im}(x)| \le a$.

2. $g \in C^q(U_{r_0}, \mathbb{R}^2)$, where $U_{\mathtt{r}_0}$ is the ball of radius $\mathtt{r}_0$ in $\mathbb{R}^6$, for some large $q$ in the real sense.

**Hypothesis 4.** *Assume that* $\mathtt{f}$ *is such that*

*(i)* $\mathtt{f}(-\eta_0, \eta_1, -\eta_2) = -\mathtt{f}(\eta_0, \eta_1, \eta_2)$.

*(ii)* $\mathtt{f}(\eta_0, \eta_1, \eta_2) = \overline{\mathtt{f}(\bar{\eta}_0, \bar{\eta}_1, \bar{\eta}_2)}$,

*(iii)* *we require that* $\int_{\mathbb{T}} |u|^2 dx$ *and* $\mathrm{Im} \int_{\mathbb{T}} \bar{u} u_x \, dx$ *are constants of motion for* (1.2.16),

*(iv)* $\partial_{\eta_2} \mathtt{f} \in \mathbb{R} \backslash \{0\}$, *where* $\partial_\eta = \partial_{\mathrm{Re}(\eta)} - i\, \partial_{\mathrm{Im}(\eta)}$.

*(v)* *the cubic term in the non linearity has the form*

$$\begin{aligned} f(u, u_x, u_{xx}) &= \mathtt{a}_1 |u|^2 u + \mathtt{a}_2 |u|^2 u_{xx} + \mathtt{a}_3 |u_x|^2 u + \mathtt{a}_4 |u_x|^2 u_{xx} + \mathtt{a}_6 |u_{xx}|^2 u_{xx} \\ &\quad + \mathtt{b}_2 u^2 \bar{u}_{xx} + \mathtt{b}_3 (u_x)^2 \bar{u} + \mathtt{b}_4 (u_x)^2 \bar{u}_{xx} \end{aligned} \tag{1.2.18}$$

*with* $\mathtt{a}_i \in \mathbb{R}$ *for* $i = 1, 2, 3, 4, 6$ *and* $\mathtt{b}_i \in \mathbb{R}$ *for* $i = 2, 3, 4$.

Items $(i), (ii), (iv)$ in Hypotheses 4 are the same reversibility assumptions we did in Hypotheses 1. Item $(v)$ is nothing but the request that the leading term of $f$ is cubic. The form in (1.2.18) comes from items $(i) - (iv)$.

**Definition 1.2.3.** *We say that* $(\mathtt{a}_1, \mathtt{a}_2, \mathtt{a}_3, \mathtt{a}_4, \mathtt{a}_6, \mathtt{b}_2, \mathtt{b}_3, \mathtt{b}_4) \neq 0$ *is resonant if either:*

1. $\mathtt{a}_6 = \mathtt{a}_1 = 0$ *and* $\mathtt{a}_4 - \mathtt{b}_4 = 0$ *and* $\mathtt{a}_3 - \mathtt{a}_2 - \mathtt{b}_2 - \mathtt{b}_3 = 0$.

2. $\mathtt{a}_6 = \mathtt{a}_1 = 0$ *and* $\mathtt{a}_4 - \mathtt{b}_4 = 0$ *and* $\mathtt{a}_3 - \mathtt{a}_2 - \mathtt{b}_2 - \mathtt{b}_3 \neq 0$ *but one has either* $\mathtt{a}_2 = 0$ *and* $\mathtt{b}_2 \neq 0$ *and* $\mathtt{a}_3 - \mathtt{b}_3 = (6d+1)/(2d+1)\mathtt{b}_2$ *or* $a_2 \neq 0$ *and* $\mathtt{a}_3 - (1+3d)\mathtt{a}_2 - \mathtt{b}_3 = 0$ *and* $d\mathtt{a}_2 = \mathtt{b}_2$.

3. $\mathtt{a}_6 = \mathtt{a}_1 = 0$ *and* $\mathtt{a}_3 - \mathtt{a}_2 - \mathtt{b}_2 - \mathtt{b}_3 = 0$ *and* $\mathtt{a}_4 - \mathtt{b}_4 \neq 0$ *but one has* $\mathtt{a}_4 \neq 0$ *and* $(2d-1)\mathtt{a}_4 = \mathtt{b}_4$.

We are now ready to state our main Theorem on the existence of quasi-periodic solutions of $d$ frequencies which is based on the following "genericity" condition.

**Definition 1.2.4 (Genericity).** *Given a finite number of variables* $x \in \mathbb{N}^d$ *and a non-trivial polynomial in* $x$, *we say that a point* $x_0$ *is* "generic" *if it is not a zero of the polynomial.*

Equation (1.2.16) can be seen as an infinte dimensional dynamical system with phase space a scale of complex Hilbert spaces $u \in h^{a,p}$ with

$$h^{a,p} := \{u = \{u_k\}_{k \in \mathbb{Z}} \ : \|u\|_{a,p}^2 := \sum_{k \in \mathbb{Z}} |u_k|^2 e^{2a|k|} \langle k \rangle^{2p} < \infty\} \tag{1.2.19}$$

where $0 \leq a \leq \mathfrak{a}/2$ and $p \geq 1/2$. Note that there is an isometric one-to-one correspondance between a sequence $\{u_k\}$ and a function $u = \sum_k u_k e^{ik\cdot x}$ in $H^p(\mathbb{T}_a)$, i.e. the analytic function on the complex strip $\mathbb{T}_a$ that are p-Sobolev on the boundary. We will use the same symbol $u \in h^{a,p}$ to indicate both the sequence and the function.

Note that in case 1 $\mathtt{f}$ is an analytic map from the ball $B_{\mathtt{r}_0} \subset h^{a,p}$ to $h^{a,p-2}$ for any $a \leq \mathfrak{a}/2$. In case 2 we have $a = \mathfrak{a} = 0$ and $\mathtt{f}$ is a $C^q$ map from $B_{\mathtt{r}_0} \subset h^{0,p}$ to $h^{0,p-2}$.

A *quasi*-periodic solutions of (1.2.16),is an embedding

$$\mathbb{T}^d \ni \varphi \mapsto v(\varphi, x) \in h^{a,p}\,, \quad d \geq 1, \tag{1.2.20}$$

and a frequency vector $\omega_\infty \in \mathbb{R}^d$ such that $u(t,x) = v(\omega_\infty t, x)$ is a solution of the equation and $v(\varphi, x) \in H^p(\mathbb{T}_a^{d+1})$. Note that, in the autonomous case, both the embedding $v$ and the frequency vector $\omega_\infty$ are a unknown of the problems. Again these are the analytic functions on the complex strip $\mathbb{T}_a^{d+1}$ that are p-Sobolev on the boundary.

**Theorem 1.2.5.** *Consider the equation* (1.2.16) *in case 1, namely when* $\mathtt{f}$ *as in* (1.2.17) *is an analytic function. Assume the Hypothesis 4 and moreover that* $(\mathtt{a}_1, \mathtt{a}_2, \mathtt{a}_3, \mathtt{a}_4, \mathtt{a}_6, \mathtt{b}_2, \mathtt{b}_3, \mathtt{b}_4)$ *is not resonant. There exists a non trivial polynomial such that for any* $d \in \mathbb{N}$ *with* $d > 2$ *and for any choice of* $\mathtt{v}_1, \dots, \mathtt{v}_d \in \mathbb{N}$ generic *with respect to the polynomial the following holds.*

*There exists* $a = a(d, \mathtt{f})$ *and* $\varepsilon_0 = \varepsilon_0(d, \mathtt{f})$ *such that for all* $\varepsilon \in (0, \varepsilon_0)$*, there exists a Cantor set*

$$\mathcal{C}_\varepsilon \subset \varepsilon^2 \left[\frac{1}{2}, \frac{3}{2}\right]^d, \quad |\mathcal{C}_\varepsilon| \to 1 \quad as \quad \varepsilon \to 0, \tag{1.2.21}$$

*of asymptotically full Lebesgue measure, such that for all* $\xi \in \mathcal{C}_\varepsilon$ *the NLS equation (1.2.16) has a quasi-periodic solution with frequency* $\omega^\infty$ *given by the embedding* $v(\xi) \in H^1(\mathbb{T}_a^{d+1})$*:*

$$v = \sum_{i=1}^d \sqrt{\xi_i} e^{\mathrm{i}\varphi_i} \sin(\mathtt{v}_i x) + o(\sqrt{\xi})\,, \quad \omega_i^\infty(\xi) = \mathtt{v}_i^2 + \sum_j \mathcal{M}_i^j \xi_j + o(\xi) \tag{1.2.22}$$

*with* $\mathcal{M}$ *an invertible matrix. Moreover one has* $v(\varphi, x) = -v(\varphi, -x)$ *and* $v(\varphi, x) = \bar{v}(-\varphi, x)$*, and the solution is* linearly stable.

Note that in particular we are able to separate the condition on $(\mathtt{a}_1, \mathtt{a}_2, \mathtt{a}_3, \mathtt{a}_4, \mathtt{a}_6, \mathtt{b}_2, \mathtt{b}_3, \mathtt{b}_4)$ from the one on $\mathtt{v}_1, \dots, \mathtt{v}_d$. For example given any choice of $(\mathtt{a}_1, \mathtt{a}_2, \mathtt{a}_3, \mathtt{a}_4, \mathtt{a}_6, \mathtt{b}_2, \mathtt{b}_3, \mathtt{b}_4)$ such that $a_1 \neq 0$, then the genericity condition can be verified by removing only a co-dimension one algebraic manifold in the variables $\mathtt{v}_1, \dots, \mathtt{v}_d$.

In the case of finite regularity we have a similar result.

**Theorem 1.2.6.** *Consider the equation* (1.2.16) *in case 2. There exists* $q = q(d)$ *such that for any non linearity* $\mathtt{f} \in C^q$ *that satisfies Hypothesis 4 and moreover such that* $(\mathtt{a}_1, \mathtt{a}_2, \mathtt{a}_3, \mathtt{a}_4, \mathtt{a}_6, \mathtt{b}_2, \mathtt{b}_3, \mathtt{b}_4)$ *is not resonant, there exists a non trivial polynomial such that for any* $d \in \mathbb{N}$ *with* $d > 2$ *and for any choice of* $\mathtt{v}_1, \dots, \mathtt{v}_d \in \mathbb{N}$ generic *with respect to the polynomial the following holds.*

*There exist* $p = p(d, \mathtt{f})$, $\varepsilon_0 = \varepsilon_0(d, \mathtt{f})$ *such that for all* $\varepsilon \in (0, \varepsilon_0)$, *there exists a Cantor set*

$$\mathcal{C}_\varepsilon \subset \varepsilon^2 \left[\frac{1}{2}, \frac{3}{2}\right]^d, \quad |\mathcal{C}_\varepsilon| \to 1 \quad as \quad \varepsilon \to 0, \tag{1.2.23}$$

*of asymptotically full Lebesgue measure, such that for all* $\xi \in \mathcal{C}_\varepsilon$ *the NLS equation (1.2.16) has a quasi-periodic solution with frequency* $\omega^\infty$ *given by the embedding* $v(\xi) \in H^s(\mathbb{T}^{d+1})$*:*

$$v = \sum_{i=1}^{d} \sqrt{\xi_i} e^{\mathrm{i}\varphi_i} \sin(\mathtt{v}_i x) + o(\sqrt{\xi}), \quad \omega_i^\infty(\xi) = \mathtt{v}_i^2 + \sum_j \mathcal{M}_i^j \xi_j + o(\xi)$$

*with* $\mathcal{M}$ *an invertible matrix. Moreover one has* $v(\varphi, x) = -v(\varphi, -x)$ *and* $v(\varphi, x) = \bar{v}(-\varphi, x)$, *and the solution is* linearly stable.

In the autonomous case we provides two existence results. One in analytic class and one in Sobolev regularity. Theorem 1.2.5 is the first result of analytic solutions for quasi-linear partial differential equations.

# 2. Functional Setting

In this Section we introduce the functional spaces on which we work. Moreover we analyze in a specific way the role of the "reversibility" condition and how we use it in Theorems 1.1.1, 1.2.5 and 1.2.6. The Hamiltonian structure of NLS will be analyzed in Section 2.3.

## 2.1 Scales of Sobolev spaces

We have already seen that we will work on Sobolev spaces $H^s$ defined in (1.1.4). For the analytic contest we introduce the space of analytic functions that are Sobolev on the boundary

$$H^s(\mathbb{T}^b_a;\mathbb{C}) := \big\{u = \sum_{l\in\mathbb{Z}^b} u_l e^{\mathrm{i}l\cdot\theta} : \|u\|^2_{s,p} := \sum_{l\in\mathbb{Z}^b} \langle l\rangle^{2p}|u_l|^2 e^{2s|l|} < \infty\big\}\,. \tag{2.1.1}$$

for $a > 0$ and for some $b \geq 1$. Clearly the space $H^s(\mathbb{T}^b_a)$ is in one-to-one correspondence with the sequences space. We denote the space of sequences by $h^{a,s}$ (see (1.2.19)).

In the forced cases of Theorems 1.1.1 and 1.1.2, we directly study the equations for the embeddings (see (1.1.3) and (1.1.9)) and we look for a solutions $v(\varphi, x) \in H^s(\mathbb{T}^d\times\mathbb{T};\mathbb{C})$ for some $s$. On the other hand in the autonomous case of Theorems 1.2.5 and 1.2.6 it is convenient to study the equation as dynamical system on the phase space $H^1(\mathbb{T}_a;\mathbb{C})$ (or $H^1(\mathbb{T};\mathbb{C})$ in the Sobolev case), i.e. look for $u(t) \in H^1(\mathbb{T}_a;\mathbb{C})$ quasi-periodic in $t$. In order to distinguish these two cases, for the autonomous system we will use the equivalent notation $h^{a,p}$ to denote the functions in $H^p(\mathbb{T}_a;\mathbb{C})$. We will always use the capitol letter $H$ to denote the space of functions of $d+1$ variables.

Due to the complex nature of the NLS we need to work on product spaces. We will usually denote

$$\begin{aligned} \mathrm{H}^s &:= \mathrm{H}^s(\mathbb{T}^{d+1};\mathbb{R}) = H^s(\mathbb{T}^{d+1};\mathbb{R}) \times H^s(\mathbb{T}^{d+1};\mathbb{R}),\\ \mathbf{H}^s &:= \mathbf{H}^s(\mathbb{T}^{d+1};\mathbb{C}) = H^s(\mathbb{T}^{d+1};\mathbb{C}) \times H^s(\mathbb{T}^{d+1};\mathbb{C}) \cap \mathcal{U}, \end{aligned} \tag{2.1.2}$$

where

$$\mathcal{U} = \{(h^+, h^-) \in H^s(\mathbb{T}^{d+1};\mathbb{C}) \times H^s(\mathbb{T}^{d+1};\mathbb{C}) \;:\; h^+ = \overline{h^-}\}. \tag{2.1.3}$$

There is a one-to-one correspondence between these two spaces given by $\mathrm{H}^s \ni v = (v^{(1)}, v^{(2)}) \mapsto w = (u, \bar{u}) \in \mathbf{H}^s$ with $u = v^{(1)} + iv^{(2)}$. To simplify the notation, in the thesis we should use the same symbol $v$ to indicate a function $v \in \mathrm{H}^s$ or $v \in \mathbf{H}^s$. We will use different symbols in some cases only to avoid confusion.

We also write $\mathrm{H}^s_x$ and $\mathbf{H}^s_x$ to denote the phase space of functions in $\mathrm{H}^s(\mathbb{T};\mathbb{R}) = H^s(\mathbb{T}^1;\mathbb{R}) \times H^s(\mathbb{T}^1;\mathbb{R})$ and $\mathbf{H}^s_x(\mathbb{T};\mathbb{C}) = H^s(\mathbb{T}^1;\mathbb{C}) \times H^s(\mathbb{T}^1;\mathbb{C}) \cap \mathcal{U}$, On the product spaces $\mathrm{H}^s$ and $\mathbf{H}^s$ we define, with abuse of notation, the norms

$$\begin{aligned} &||z||_{\mathrm{H}^s} := \max\{||z^{(i)}||_s\}_{i=1,2}, \quad z = (z^{(1)}, z^{(2)}) \in \mathrm{H}^s, \\ &||w||_{\mathbf{H}^s} := ||z||_{H^s(\mathbb{T}^{d+1};\mathbb{C})} = ||z||_s, \quad w = (z, \bar{z}) \in \mathbf{H}^s, \quad z = z^{(1)} + iz^{(2)}. \end{aligned} \tag{2.1.4}$$

For a function $f : \Lambda \to E$ where $\Lambda \subset \mathbb{R}^n$ and $(E, \|\cdot\|_E)$ is a Banach space we define

$$sup\ \ norm : ||f||_E^{sup} := ||f||_{E,\Lambda}^{sup} := \sup_{\lambda \in \Lambda} ||f(\omega)||_E, \tag{2.1.5}$$

$$Lipschitz\ \ semi-norm : ||f||_E^{lip} := ||f||_{E,\Lambda}^{lip} := \sup_{\substack{\omega_1,\omega_2 \in \Lambda \\ \omega_1 \neq \omega_2}} \frac{||f(\omega_1) - f(\omega_2)||_E}{|\lambda_1 - \lambda_2|}$$

and for $\gamma > 0$ the weighted Lipschitz norm

$$||f||_{E,\gamma} := ||f||_{E,\Lambda,\gamma} := ||f||_E^{sup} + \gamma ||f||_E^{lip}. \tag{2.1.6}$$

In the paper we will work with parameter families of functions in $\mathcal{H}_s$, If one deal with parameters family $u = u(\lambda) \in \mathrm{Lip}(\Lambda, \mathcal{H}_s)$ where $\mathcal{H}_s = \mathrm{H}^s, \mathbf{H}^s$ and $\Lambda \subset \mathbb{R}^d$ we simply write $\|f\|_{\mathcal{H}_s,\gamma} := \|f\|_{s,\gamma}$, or $\|u\|_{s,p,\gamma}$ in the analytic contest. All the discussion above holds for the product space $\mathbf{h}^{a,p} := h^{a,p} \times h^{a,p}$. Along the Thesis we shall write also

$$a \le_s b \quad \Leftrightarrow \quad a \le C(s) b \ \ \text{for some constant } \ C(s) > 0.$$

Moreover to indicate unbounded or regularizing spatial differential operator we shall write $O(\partial_x^p)$ for some $p \in \mathbb{Z}$. More precisely we say that an operator $A$ is $O(\partial_x^p)$ if

$$A : H^s_x \to H^{s-p}_x, \quad \forall s \ge 0. \tag{2.1.7}$$

Clearly if $p < 0$ the operator is regularizing.
Now we define the subspaces of trigonometric polynomials

$$H_n = H_{N_n} := \{u \in L^2(\mathbb{T}^{d+1}) : u(\varphi, x) := \sum_{|(\ell,j)| \le N_n} u_j(\ell) e^{i(\ell\cdot\varphi + jx)}\} \tag{2.1.8}$$

where $N_n := N_0^{(\frac{3}{2})^n}$, and the orthogonal projection

$$\Pi_n := \Pi_{N_n} : L^2(\mathbb{T}^{d+1}) \to H_n, \quad \Pi_n^{\perp} := \mathbb{1} - \Pi_n.$$

This definitions can be extended to the product spaces in (2.1.2) in the obvious way. We have the following classical result.

**Lemma 2.1.7.** *Fo any $s \geq 0$ and $\nu \geq 0$ there exists a constant $C := C(s, \nu)$ such that*

$$\begin{aligned} &\|\Pi_n u\|_{s+\nu,\gamma} \leq C N_n^{\nu} \|u\|_{s,\gamma}, \quad \forall u \in H^s, \\ &\|\Pi_n^{\perp} u\|_s \leq C N_n^{-\nu} \|u\|_{s+\nu}, \quad \forall u \in H^{s+\nu}. \end{aligned} \tag{2.1.9}$$

We omit the proof of the Lemma since bounds (2.1.9) are classical estimates for truncated Fourier series which hold also for the norm in (3.1.1) and in the analytic case.

We have introduced the space $\mathrm{H}^s$ and $\mathbf{H}^s$ because we want to rewrite the problem of the existence of quasi-periodic solutions for equations (1.1.3) and (1.1.9) as into the research of zeros of some functionals on the functional spaces $\mathrm{H}^s$ and $\mathbf{H}^s$.

In the Hamiltonian case of Theorem (1.1.2) we rewrite equation (1.1.9) into the form

$$\mathcal{F}(\omega t, x, w) = 0, \tag{2.1.10}$$

where $w := (\xi, \eta) \in \mathrm{H}^s$ and we defined the functional $\mathcal{F}$ on the space $\mathrm{H}^s$ as

$$\mathcal{F}(\omega t, x, w) := D_\omega w + \varepsilon g(\omega t, x, w), \quad D_\omega = \begin{pmatrix} \omega \cdot \partial_\varphi & -\partial_{xx} - m \\ \partial_{xx} + m & \omega \cdot \partial_\varphi \end{pmatrix}, \tag{2.1.11}$$

where

$$g(\omega t, x, w) := \begin{pmatrix} -f_2(\varphi, x, \xi, \eta, \xi_x, \eta_x, \xi_{xx}, \eta_{xx}) \\ f_1(\varphi, x, \xi, \eta, \xi_x, \eta_x, \xi_{xx}, \eta_{xx}) \end{pmatrix}. \tag{2.1.12}$$

and $f_i$, for $i = 1, 2$, defined in (1.1.8). This will be explained better in Section 2.3. this approach as the advantages that one uses real coordinates to study the equation.

Another possible choice to study the non linear Schrödinger equation is to introduce complex independent coordinates

$$\mathbf{u} := (u^+, u^-) \in H^s(\mathbb{T}^{d+1}; \mathbb{C}) \times H^s(\mathbb{T}^{d+1}; \mathbb{C}), \tag{2.1.13}$$

and then to study the system on the "real" subspace $\mathcal{U}$ in (2.1.3) in which one looks for the solution. The advantages of using the veriables $u$ and $\bar{u}$ is that the linear operator diagonal as we will see in the following. In the reversible case of Theorem 1.1.1 we introduce the following functional.

**Definition 2.1.8.** *Given $\mathtt{f} \in C^q$, we define the "vector" NLS as*

$$\begin{aligned} F(\mathbf{u}) &:= \omega \cdot \partial_\varphi \mathbf{u} + i(\partial_{xx}\mathbf{u} + \varepsilon f(\varphi, x, \mathbf{u})) = 0, \\ f(\varphi, x, \mathbf{u}) &:= \begin{pmatrix} f_1(\varphi, x, u^+, u^-, u_x^+, u_x^-, u_{xx}^+, u_{xx}^-) \\ f_2(\varphi, x, u^+, u^-, u_x^+, u_x^-, u_{xx}^+, u_{xx}^-) \end{pmatrix} \end{aligned} \tag{2.1.14}$$

*where the functions $f = (f_1, f_2)$ extend $(\mathtt{f}, \bar{\mathtt{f}})$ in the following sense. The $f_j$ are in $C^q(\mathbb{T}^{d+1} \times \mathbb{R}^6 \times \mathbb{R}^6; \mathbb{R}^2)$, and moreover on the subspace $\mathcal{U}$ they satisfy $f = (\mathtt{f}, \bar{\mathtt{f}})$ and*

$$\begin{gathered} \partial_{z_2^+} f_1 = \partial_{z_2^-} f_2, \quad \partial_{z_i^+} f_1 = \overline{\partial_{z_i^-} f_2}, \ \ i = 0, 1, \quad \partial_{z_i^-} f_1 = \overline{\partial_{z_i^+} f_2}, \quad i = 0, 1, 2, \\ \partial_{\overline{z_i^+}} f_1 = \partial_{\overline{z_i^+}} f_2 = \partial_{\overline{z_i^-}} f_1 = \partial_{\overline{z_i^-}} f_2 = 0 \end{gathered} \tag{2.1.15}$$

*where $\partial_{\overline{z_j^\sigma}} = \partial_{\mathrm{Re}\, z_j^\sigma} + i\partial_{\mathrm{Im}\, z_j^\sigma}$, $\sigma = \pm$. Note that this extension is trivial in the analytic case.*

By Definition 2.1.8 the (2.1.14) reduces to (1.1.3) on the subspace $\mathcal{U}$. The advantage of working on (2.1.14) is that the linearized operator $dF(\mathbf{u}) := \mathcal{L}(\mathbf{u})$ for $\mathbf{u} \in \mathcal{U}$ is self-adjoint. Note that the linearized operator of (1.1.3) is actually self-adjoint, but even at $\varepsilon = 0$ is not diagonal. To diagonalize one needs to complexify and then to give meaning to $f \in C^q$, thus we introduce the extension.

The proofs af Theorems 1.1.1 and 1.1.2 are based on an generalized Implicit function Theorem that we prove in Section 3.1. As we will see such theorem is based on very mild hypotheses. Hence we analyze the two cases into two different way just to stress the strength of our approach and to underline the fact that the two formulations in (2.1.11) and (2.1.14) present just some little technical differences.

Concerning the autonomous case of Theorems 1.2.5 and 1.2.6 we define

$$\mathbf{u} := (u^+, u^-) \in \mathbf{h}^{a,p} := h^{a,p} \times h^{a,p}.$$

and we consider the dynamical system

$$\dot{\mathbf{u}} := -\mathrm{i}E \left[ \mathbf{u}_{xx} + \begin{pmatrix} \mathtt{f}^+(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \\ \mathtt{f}^-(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \end{pmatrix} \right] = \chi(\mathbf{u}) = \begin{pmatrix} \chi^+(\mathbf{u}) \\ \chi^-(\mathbf{u}) \end{pmatrix}, \qquad E = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.1.16}$$

where $\mathtt{f}^\pm$ are defined in such a way that, on the subspace $\mathcal{U} := \{u^+ = \overline{u^-}\}$, the system (2.1.16) is equivalent to (1.2.16). Essentially one uses the same extension defined in Definition 2.1.8.

If $\mathtt{f}$ is analytic this extension is completely standard, indeed one may Taylor expand $\mathtt{f}$ as totally convergent series in $u, \bar{u}$ (and their derivatives). In the $C^q$ case this requires some care, see for instance [31]. Here the notation of a vector field is the following:

$$\chi(\mathbf{u}) = \sum_{\sigma=\pm} \chi^\sigma(u)\partial_{u^\sigma} = \sum_{\sigma=\pm} \sum_{j\in\mathbb{Z}} \chi_j^\sigma \partial_{u_j^\sigma}, \tag{2.1.17}$$

Note that the map $\mathcal{F} : \mathbf{h}^{a,p} \to \mathbf{h}^{a,p-2}$ defined by

$$\mathcal{F} : \mathbf{u} \mapsto \begin{pmatrix} \mathtt{f}^+(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \\ \mathtt{f}^-(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \end{pmatrix}, \tag{2.1.18}$$

is a composition operator. This implies that the linearized operator at some $\mathbf{u}$ is of the form

$$d_{\mathbf{u}}\mathcal{F}(\mathbf{u}) = d_{\eta_2^\pm} \begin{pmatrix} \mathtt{f}^+(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \\ \mathtt{f}^-(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \end{pmatrix} \partial_{xx} + d_{\eta_1^\pm} \begin{pmatrix} \mathtt{f}^+(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \\ \mathtt{f}^-(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \end{pmatrix} \partial_x + d_{\eta_0^\pm} \begin{pmatrix} \mathtt{f}^+(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \\ \mathtt{f}^-(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}) \end{pmatrix}.$$

Thus $\chi$ linearized at any $\mathbf{u}$ has a very special multiplicative structure, namely on $\mathcal{U}$ is acts on functions $\mathbf{h}(x) = (h^+(x), h^-(x))$ as

$$d_{\mathbf{u}}\chi(\mathbf{u})[\mathbf{h}] = -\mathrm{i}E \left[ \begin{pmatrix} 1 + a_2(x) & b_2(x) \\ \bar{b}_2(x) & 1 + \bar{a}_2(x) \end{pmatrix} \partial_{xx} + \begin{pmatrix} a_1(x) & b_1(x) \\ \bar{b}_1(x) & \bar{a}_1(x) \end{pmatrix} \partial_x + \begin{pmatrix} a_0(x) & b_0(x) \\ \bar{b}_0(x) & \bar{a}_0(x) \end{pmatrix} \right] \mathbf{h}(x). \tag{2.1.19}$$

## 2.2 Reversible structure.

By Hypothesis 1 one has that (2.1.14), restricted to $\mathcal{U}$, is reversible with respect to the involution

$$S : u(t, x) \to -\bar{u}(t, -x), \quad S^2 = \mathbb{1}, \tag{2.2.20}$$

namely, setting $V(t, u) := -i(u_{xx} + \varepsilon \mathtt{f}(\omega t, x, u, u_x, u_{xx}))$ we have

$$-SV(-t, u) = V(t, Su).$$

In the same way using Hypothesis 4 it turns out that equation (2.1.16) is reversible with respect the involution (2.2.20) and hence we have

$$-S \circ \chi(\mathbf{u}) = \chi \circ S(\mathbf{u}).$$

Hence the subspace of “reversible” solutions

$$u(t, x) = -\bar{u}(-t, -x). \tag{2.2.21}$$

is invariant. Actually we look for *odd reversible* solutions i.e. $u$ which satisfy (2.2.21) and $u(t, x) = -u(t, -x)$. Hence we choose as phase space of (2.1.16)

$$\mathbf{h}^{a,p}_{\rm odd} := \left\{ (u^+, u^-) \in \mathbf{h}^{a,p} \,:\, u^\sigma_k = -u^\sigma_{-k} \right\}, \tag{2.2.22}$$

essentially the couples of odd functions in $H^p(\mathbb{T}_a)$. Then (2.2.21) reads $u(t, x) = \bar{u}(-t, x)$.

To formalize this condition also for (2.1.14) we introduce spaces of odd or even functions in $x \in \mathbb{T}$. For all $s \geq 0$, we set

$$\begin{aligned} X^s &:= \left\{ u \in H^s(\mathbb{T}^d \times \mathbb{T}) : \quad u(\varphi, -x) = -u(\varphi, x), \;\; u(-\varphi, x) = \bar{u}(\varphi, x) \right\}, \\ Y^s &:= \left\{ u \in H^s(\mathbb{T}^d \times \mathbb{T}) : \quad u(\varphi, -x) = u(\varphi, x), \;\; u(-\varphi, x) = \bar{u}(\varphi, x) \right\}, \\ Z^s &:= \left\{ u \in H^s(\mathbb{T}^d \times \mathbb{T}) : \quad u(\varphi, -x) = -u(\varphi, x), \;\; u(-\varphi, x) = -\bar{u}(\varphi, x) \right\}, \end{aligned} \tag{2.2.23}$$

Note that odd reversible solutions means $u \in X^s$, moreover an operator reversible w.r.t. the involution $S$ maps $X^s$ to $Z^s$.

**Definition 2.2.9.** *We denote with bold symbols the spaces $\mathbf{G}^s := G^s \times G^s \cap \mathcal{U}$ where $G^s$ is $H^s$, $X^s, Y^s$ or $Z^s$.*
*We denote by $H^s_x := H^s(\mathbb{T})$ the Sobolev spaces of functions of $x \in \mathbb{T}$ only, same for all the subspaces $G^s_x$ and $\mathbf{G}^s_x$.*

**Remark 2.2.10.** *Given a family of linear operators $A(\varphi) : H^s_x \to H^s_x$ for $\varphi \in \mathbb{T}^d$, we can associate it to an operator $A : H^s(\mathbb{T}^{d+1}) \to H^s(\mathbb{T}^{d+1})$ by considering each matrix element of $A(\varphi)$ as a multiplication operator. This identifies a subalgebra of linear operators on $H^s(\mathbb{T}^{d+1})$. An operator $A$ in the sub-algebra identifies uniquely its corresponding "phase space" operator $A(\varphi)$. With reference to the Fourier basis this sub algebra is called "Töpliz-in-time" matrices (see formulæ* (4.3.98), (4.3.99)*).*

**Remark 2.2.11.** *Part of the proof of Theorem 1.1.1 is to control that, along the algorithm, the operator $d_u F(\lambda, \varepsilon, u)$ maps the subspace $\mathbf{X}^0$ into $\mathbf{Z}^0$. In order to do this, we will introduce the notions of "reversible" and "reversibility-preserving" operator in Section 4.1.*

## 2.3 Hamiltonian structure

Given a function $u \in \mathrm{H}^s$ if we write $u = \xi + i\eta$ one has that the equation (1.1.9) reads

$$\begin{cases} \omega \cdot \partial_\varphi \xi = \eta_{xx} + \mathtt{m}\eta + \varepsilon f_2(\varphi, x, \xi, \eta, \xi_x, \eta_x, \xi_{xx}, \eta_{xx}), \\ -\omega \cdot \partial_\varphi \eta = \xi_{xx} + \mathtt{m}\xi + \varepsilon f_1(\varphi, x, \xi, \eta, \xi_x, \eta_x, \xi_{xx}, \eta_{xx}), \end{cases} \tag{2.3.24}$$

where $f_i$ for $i = 1, 2$ are defined in (1.1.8). Equation (2.3.24) is nothing but equation (2.1.10) written in an explicit way. Now we analyze its Hamiltonian structure. Thanks to Hypotesis 2 we can write

$$\dot{w} = \chi_H(w) := J\nabla H(w), \quad w = (\xi, \eta) \in \mathrm{H}^s, \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{2.3.25}$$

If we consider the space $\mathrm{H}^s$ endowed with the symplectic form

$$\widetilde{\Omega}(w,v) := \int_{\mathbb{T}} w \cdot Jv dx = (w, Jv)_{L^2\times L^2}, \quad \forall\, w, v \in \mathrm{H}^s \tag{2.3.26}$$

where $\cdot$ is the usual $\mathbb{R}^2$ scalar product, then $\chi_H$ is the Hamiltonian vector field generator by the hamiltonian function

$$H : \mathrm{H}^s \to \mathbb{R}, \qquad H(w) = \frac{1}{2}\int_{\mathbb{T}} |w_x|^2 + \mathtt{m}|w|^2 + \varepsilon F(\omega t, x, w, w_x). \tag{2.3.27}$$

Indeed, for any $w, v \in \mathrm{H}^s$ one has

$$dH(w)[h] = (\nabla H(w), h)_{L^2(\mathbb{T})\times L^2(\mathbb{T})} = \widetilde{\Omega}(\chi_H(u), h),$$

With this notation one has

$$\begin{aligned} f_1 &:= -\partial_\xi F + \partial_{\xi\xi_x} F\xi_x + \partial_{\eta\xi_x} F\eta_x + \partial_{\xi_x\xi_x} F\xi_{xx} + \partial_{\xi_x\eta_x} F\eta_{xx}, \\ f_2 &:= -\partial_\eta F + \partial_{\xi\eta_x} F\xi_x + \partial_{\eta\eta_x} F\eta_x + \partial_{\xi_x\eta_x} F\xi_{xx} + \partial_{\eta_x\eta_x} F\eta_{xx}, \end{aligned} \tag{2.3.28}$$

where all the functions are evaluated in $(\varphi, x, \xi, \eta, \xi_x, \eta_x, \xi_{xx}, \eta_{xx})$. One can check that the (2.3.24) is equivalent to (1.1.9). It is sufficient to multiply by the constant $i$ the first equation and to add or subtract the second one, one obtains

$$\begin{aligned} i\omega\cdot\partial_\varphi u &= i\omega\cdot\partial_\varphi\xi - \omega\cdot\partial_\varphi\eta = u_{xx} + \mathtt{m}u + \varepsilon\mathtt{f}, \\ i\omega\cdot\partial_\varphi\bar{u} &= i\omega\cdot\partial_\varphi\xi + \omega\cdot\partial_\varphi\eta = -\bar{u}_{xx} - \mathtt{m}\bar{u} - \varepsilon\bar{\mathtt{f}} \end{aligned} \tag{2.3.29}$$

The classical approach is to consider the "double" the NLS in the product space $H^s(\mathbb{T}^{d+1};\mathbb{C})\times H^s(\mathbb{T}^{d+1};\mathbb{C})$ in the complex independent variables $(u^+, u^-)$. One recovers the equation (1.1.9) by studying the system in the subspace $\mathcal{U} = \{u^+ = \overline{u^-}\}$ (see the (2.3.29)).

On the contrary we prefer to use the real coordinates, because we are working in a differentiable structure. To define a differentiable structure on complex variables is more less natural. Anyway, one can see in [31] how to deal with this problem. There, the authors find an extension of the vector fields on the complex plane that is merely differentiable. The advantage of that approach, is to deal with a diagonal linear operator. How we will see in the following of this paper, it is not necessary to apply the abstract Nash-Moser Theorem proved in Section Section 3.1.

The phase space for the NLS is $\mathrm{H}^1 := H^1(\mathbb{T};\mathbb{R}) \times H^1(\mathbb{T};\mathbb{R})$. In general we have the following definitions:

**Definition 2.3.12.** *We say that a time dependent linear vector field $\chi(t) : \mathrm{H}^s \to \mathrm{H}^s$ is* Hamiltonian *if $\chi(t) = J\mathcal{A}(t)$, where $J$ is defined in (2.3.25) and $\mathcal{A}(t)$ is a real linear operator that is self-adjoint with respect the real scalar product on $L^2 \times L^2$. The corresponding Hamiltonian has the form*

$$H(u) := \frac{1}{2}(\mathcal{A}(t)u, u)_{L^2\times L^2} = \int_{\mathbb{T}} \mathcal{A}(t)u \cdot u dx$$

*Moreover, if $\mathcal{A}(t) = \mathcal{A}(\omega t)$ is quasi-periodic in time, then the associated operator $\omega \cdot \partial_\varphi \mathbb{1} - J\mathcal{A}(\varphi)$ is called Hamiltonian.*

**Definition 2.3.13.** *We say that a map $A : \mathrm{H}^1 \to \mathrm{H}^1$ is* symplectic *if the symplectic form $\widetilde{\Omega}$ in (2.3.26) is preserved, i.e.*

$$\widetilde{\Omega}(Au, Av) = \widetilde{\Omega}(u, v), \quad \forall\ u, v \in \mathrm{H}^1. \tag{2.3.30}$$

*If one has a family of symplectic maps $A(\varphi)$, $\forall\ \varphi \in \mathbb{T}^d$ then we say that the corresponding operator acting on quasi-periodic functions $u(\varphi, x)$*

$$(Au)(\varphi, x) := A(\varphi)u(\varphi, x),$$

*is symplectic.*

**Remark 2.3.14.** *Note that in complex coordinates the phase space is $\mathbf{H}^1 := H^1(\mathbb{T}; \mathbb{C}) \times H^1(\mathbb{T}; \mathbb{C})$. The definitions above are the same by using the symplectic form defined in (1.1.14) and the complex scalar product on $L^2$.*

Now it is more convenient to pass to the complex coordinates. In other words we identify an element $V := (v^{(1)}, v^{(2)}) \in \mathrm{H}^s$ with a function $v := v^{(1)} + iv^{(2)} \in H^s(\mathbb{T}^{d+1}; \mathbb{C})$. Consider the linear operator $d_z\mathcal{F}(\omega t, x, z)$ linearized in some function $z$, and consider the system

$$D_\omega V + \varepsilon d_z g(\omega t, x, z)V = 0, \quad V \in \mathrm{H}^s. \tag{2.3.31}$$

We introduce an invertible linear change of coordinate of the form

$$\begin{aligned} &T : \mathrm{H}^s \to \mathrm{H}^s, \\ &TV := \begin{pmatrix} \frac{i}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix} = \begin{pmatrix} \frac{i}{\sqrt{2}} v \\ \frac{1}{\sqrt{2}} \bar{v} \end{pmatrix}, \qquad T^{-1} := \begin{pmatrix} -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix}. \end{aligned} \tag{2.3.32}$$

We postponed the proof of the following Lemma in the Appendix:

**Lemma 2.3.15.** *The transformation of coordinates $T$ defined in (2.3.32) is symplectic. Moreover, a function $V := (v^{(1)}, v^{(2)}) \in \mathrm{H}^s$ is a solution of the system*

$$d_z\mathcal{F}(\omega t, x, z)V = 0, \tag{2.3.33}$$

*if and only if the function*

$$\begin{pmatrix} v \\ \bar{v} \end{pmatrix} := T_1^{-1}TV, \quad v \in H^s(\mathbb{T}^{d+1};\mathbb{C}), \quad T_1^{-1} := \begin{pmatrix} -i\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \tag{2.3.34}$$

*solves the system*

$$\mathcal{L}(z)\begin{pmatrix} v \\ \bar{v} \end{pmatrix} := T_1^{-1}Td_z\mathcal{F}(\omega t, x, z)T^{-1}T_1\begin{pmatrix} v \\ \bar{v} \end{pmatrix} = 0 \tag{2.3.35}$$

*In particular the operator $\mathcal{L}(z) : H^s(\mathbb{T}^{d+1};\mathbb{C}) \times H^s(\mathbb{T}^{d+1};\mathbb{C}) \to H^s(\mathbb{T}^{d+1};\mathbb{C}) \times H^s(\mathbb{T}^{d+1};\mathbb{C})$ has the form*

$$\mathcal{L}(z) = \omega \cdot \partial_\varphi \mathbb{1} + i(E + A_2)\partial_{xx} + iA_1\partial_x + i(mE + A_0)\,, \tag{2.3.36}$$

*where*

$$E = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_i = A_i(\varphi, x, z) := \begin{pmatrix} a_i & b_i \\ -\bar{b}_i & -\bar{a}_i \end{pmatrix} \tag{2.3.37}$$

*with for $i = 0, 1, 2$, and $\forall z \in H^s(\mathbb{T}^{d+1};\mathbb{C})$,*

$$\begin{aligned} 2a_i(\varphi, x) &:= \varepsilon(\partial_{z_i}\mathtt{f})(\varphi, x, z(\varphi, x), z_x(\varphi, x), z_{xx}(\varphi, x)), \\ 2b_i(\varphi, x) &:= \varepsilon(\partial_{\bar{z}_i}\mathtt{f})(\varphi, x, z(\varphi, x), z_x(\varphi, x), z_{xx}(\varphi, x)), \end{aligned} \tag{2.3.38}$$

*where we denoted $\partial_{z_i} := \partial_{z_i^{(1)}} - i\partial_{z_i^{(2)}}$ and $\partial_{\bar{z}_i} := \partial_{z_i^{(1)}} + i\partial_{z_i^{(2)}}$ for $i = 0, 1, 2$.*

The operator $\mathcal{L}$ has further property. It is clearly Hamiltonian with respect to the symplectic form in (1.1.14) and the corresponding quadratic Hamiltonian has the form

$$\begin{aligned} H(u, \bar{u}) = &\int_{\mathbb{T}} (1 + a_2)|u_x|^2 + \frac{1}{2}\left[b_2\bar{u}_x^2 + \bar{b}_2u_x^2\right] - \frac{i}{2}\mathrm{Im}(a_1)(u_x\bar{u} - u\bar{u}_x)\mathrm{d}x \\ &+ \int_{\mathbb{T}} -m|u|^2 - \mathrm{Re}(a_0)|u|^2 - \frac{1}{2}(b_0\bar{u}^2 + \bar{b}_0u^2)\mathrm{d}x. \end{aligned} \tag{2.3.39}$$

Note that the symplectic form $\Omega$ in (1.1.14) is equivalent to the $2-$form $\widetilde{\Omega}$ in (2.3.26), i.e. given $u = u^{(1)} + iu^{(2)}, v = v^{(1)} + iv^{(2)} \in H^s(\mathbb{T}^{d+1};\mathbb{C})$, one has

$$\Omega(u, w) = \mathrm{Re}\int_{\mathbb{T}} iu\bar{v}dx = \int_{\mathbb{T}} (u^{(1)}v^{(2)} - v^{(1)}u^{(2)})dx = \widetilde{\Omega}(U, V), \tag{2.3.40}$$

where we set $U = (u^{(1)}, u^{(2)}), V = (v^{(1)}, v^{(2)}) \in H^s(\mathbb{T}^{d+1}; \mathbb{R}) \times H^s(\mathbb{T}^{d+1}; \mathbb{R})$. The (2.3.39) is the general form of a linear Hamiltonian operator as $\mathcal{L}$, and, the coefficients $a_i$ in (2.3.37) have the form

$$\begin{aligned} &a_2(\varphi, x) \in \mathbb{R}, && a_1(\varphi, x) = \frac{\mathrm{d}}{\mathrm{d}x} a_2(\varphi, x) + i\mathrm{Im}(a_1)(\varphi, x), \\ &b_1(\varphi, x) = \frac{\mathrm{d}}{\mathrm{d}x} b_2(\varphi, x), && a_0(\varphi, x) = \mathrm{Re}(a_0)(\varphi, x) + \frac{i}{2}\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{Im}(a_1)(\varphi, x) \end{aligned} \tag{2.3.41}$$

# 3. Abstract Nash-Moser Theorems

In this Chapter we present two Abstract Nash-Moser scheme on which are based the proofs of Theorems stated in Chapter 1. Theorem (3.2.39) in Section 3.2 is very general and is one of the novelty introduced in this Thesis. It is a results on "tame" vector fields, and it is adapted to treat autonomous cases. Clearly it should cover forced cases. Anyway we choose to prove Theorem 3.1.18 for functionals depending on external parameters for completeness. Moreover for forced cases the proof of a generalized implicit function theorem is much simpler and less sophisticated with respect to the one of Theorem (3.2.39).

## 3.1 An Abstract Existence Theorem for forced equations

In this Section we prove an Abstract Nash-Moser theorem in Banach spaces. This abstract formulation essentially shows a method to find solutions of implicit function problems. The aim is to apply the scheme to prove Proposition 4.0.46 to the functional $F$ defined in (2.1.14).

### 3.1.1 Nash-Moser scheme

Let us consider a scale of Banach spaces $(\mathcal{H}_s, \| \ \|_s)_{s\geq 0}$, such that

$$\forall s \leq s', \quad \mathcal{H}_{s'} \subseteq \mathcal{H}_s \ \text{ and } \ \|u\|_s \leq \|u\|_{s'}\,, \ \forall u \in \mathcal{H}_{s'}\,,$$

and define $\mathcal{H} := \cap_{s\geq 0}\mathcal{H}_s$.

We assume that there is a non-decreasing family $(E^{(N)})_{N\geq 0}$ of subspaces of $\mathcal{H}$ such that $\cup_{N\geq 0}E^{(N)}$ is dense in $\mathcal{H}_s$ for any $s \geq 0$, and that there are projectors

$$\Pi^{(N)} : \mathcal{H}_0 \to E^{(N)}$$

satisfying: for any $s \geq 0$ and any $\nu \geq 0$ there is a positive constant $C := C(s, \nu)$ such that

(P1) $\|\Pi^{(N)}u\|_{s+\nu} \leq CN^{\nu}\|u\|_s$ for all $u \in \mathcal{H}_s$, (P2) $\|(\mathbb{1} - \Pi^{(N)})u\|_s \leq CN^{-\nu}\|u\|_{s+\nu}$ for all $u \in \mathcal{H}_{s+\nu}$.

In the following we will work with parameter families of functions in $\mathcal{H}_s$, more precisely we consider $u = u(\lambda) \in \mathrm{Lip}(\Lambda, \mathcal{H}_s)$ where $\Lambda \subset \mathbb{R}$. We define:

- *sup norm:* $||f||_s^{sup} := ||f||_{s,\Lambda}^{sup} := \sup_{\lambda\in\Lambda} ||f(\lambda)||_s$,
- *Lipschitz semi-norm:* $||f||_s^{lip} := ||f||_{s,\Lambda}^{lip} := \sup_{\substack{\lambda_1,\lambda_2\in\Lambda \\ \lambda_1\neq\lambda_2}} \dfrac{||f(\lambda_1) - f(\lambda_2)||_s}{|\lambda_1 - \lambda_2|}$,

and for $\gamma > 0$ the weighted Lipschitz norm

$$||f||_{s,\gamma} := ||f||_{s,\Lambda,\gamma} := ||f||_s^{sup} + \gamma||f||_s^{lip}. \tag{3.1.1}$$

Let us consider a $C^2$ map $F : [0, \varepsilon_0] \times \Lambda \times \mathcal{H}_{\mathfrak{s}_0+\nu} \to \mathcal{H}_{\mathfrak{s}_0}$ for some $\nu > 0$ and assume the following

(F0) $F$ is of the form

$$F(\varepsilon, \lambda, u) = L_\lambda u + \varepsilon f(\lambda, u)$$

where, for all $\lambda \in \Lambda$, $L_\lambda$ is a linear operator which preserves all the subspaces $E^{(N)}$.

(F1) *reversibility* property:

$$\exists\, A_s, B_s \subseteq \mathcal{H}_s \text{ closed subspaces of } \mathcal{H}_s,\ s \geq 0, \text{ such that } F : A_{s+\nu} \to B_s\,.$$

We assume also the following tame properties: given $S' > \mathfrak{s}_0$, $\forall s \in [\mathfrak{s}_0, S')$, for all Lipschitz map $u(\lambda)$ such that $\|u\|_{\mathfrak{s}_0,\gamma} \leq 1$, $(\varepsilon, \lambda) \in [0, \varepsilon_0) \times \Lambda$,

(F2) $\|f(\lambda, u)\|_{s,\gamma}, \|L_\lambda u\|_{s,\gamma} \leq C(s)(1 + \|u\|_{s+\nu,\gamma})$,

(F3) $\|d_u f(\lambda, u)[h]\|_{s,\gamma} \leq C(s)\big(\|u\|_{s+\nu,\gamma}\|h\|_{\mathfrak{s}_0+\nu,\gamma} + \|h\|_{s+\nu,\gamma}\big)$,

(F4) $\|d_u^2 f(\lambda, u)[h, v]\|_{s,\gamma} \leq C(s)\big(\|u\|_{s+\nu,\gamma}\|h\|_{\mathfrak{s}_0+\nu,\gamma}\|v\|_{\mathfrak{s}_0+\nu,\gamma} +$

$$+\|h\|_{s+\nu,\gamma}\|v\|_{\mathfrak{s}_0+\nu,\gamma} + \|h\|_{\mathfrak{s}_0+\nu,\gamma}\|v\|_{s+\nu,\gamma}\big),$$

for any two Lipschitz maps $h(\lambda)$, $v(\lambda)$.

**Remark 3.1.16.** *Note that* $(F1)$ *implies* $d_u F(\varepsilon, \lambda, v) : A_{s+\nu} \to B_s$ *for all* $v \in A_s$.

We denote

$$\mathcal{L}(u) \equiv \mathcal{L}(\lambda, u) := L_\lambda + \varepsilon d_u f(\lambda, u)\,, \tag{3.1.2}$$

we have the following definition.

**Definition 3.1.17 (Good Parameters).** *Given $\mu > 0$, $N > 1$ let*

$$\kappa_1 = 6\mu + 12\nu, \quad \kappa_2 = 11\mu + 25\nu\,, \tag{3.1.3}$$

*for any Lipschitz family $u(\lambda) \in E^{(N)}$ with $||u||_{\mathfrak{s}_0+\mu,\gamma} \leq 1$, we define the set of good parameters $\lambda \in \Lambda$ as:*

$$\mathcal{G}_N(u) := \big\{\lambda \in \Lambda \ : \ ||\mathcal{L}^{-1}(u)h||_{\mathfrak{s}_0,\gamma} \leq C(\mathfrak{s}_0)\gamma^{-1}||h||_{\mathfrak{s}_0+\mu,\gamma}, \tag{3.1.4a}$$

$$||\mathcal{L}^{-1}(u)h||_{s,\gamma} \leq C(s)\gamma^{-1}\left(||h||_{s+\mu,\gamma} + ||u||_{s+\mu,\gamma}||h||_{\mathfrak{s}_0,\gamma}\right), \tag{3.1.4b}$$

$$\forall \mathfrak{s}_0 \leq s \leq \mathfrak{s}_0 + \kappa_2 - \mu, \quad \text{for all Lipschitz maps } h(\lambda)\big\}\,.$$

Clearly, Definition 3.1.17 depends on $\mu$ and $N$.
Given $N_0 > 1$ we set

$$N_n = (N_0)^{(\frac{3}{2})^n}\,, \quad \mathcal{H}_n := E^{(N_n)}\,, \quad A_n := A_s \cap \mathcal{H}_n$$

same for the subspace $B$. Also we define

$$\Pi^{(N_n)} := \Pi_n, \qquad (\mathbb{1} - \Pi^{(N_n)}) := \Pi_n^{\perp}.$$

In the following, we shall write $a \leq_s b$ to denote $a \leq C(s)b$, for some constant $C(s)$ depending on $s$. In general, we shall write $a \lessdot b$ if there exists a constant $C$, depending only on the data of the problem, such that $a \leq Cb$.

**Theorem 3.1.18.** *(Nash-Moser algorithm) Assume $F$ satisfies $(F0) - (F4)$ and fix $\gamma_0 > 0$, $\tau > d+1$. Then, there exist constants $\epsilon_0 > 0, C_\star > 0, N_0 \in \mathbb{N}$, such that for all $\gamma \leq \gamma_0$ and $\varepsilon\gamma^{-1} < \epsilon_0$ the following properties hold for any $n \geq 0$:*
*$(N1)_n$ there exists a function*

$$u_n : \mathcal{G}_n \subseteq \Lambda \to A_n, \qquad ||u_n||_{\mathfrak{s}_0+\mu,\gamma} \leq 1, \tag{3.1.5}$$

*where the sets $\mathcal{G}_n$ are defined inductively by $\mathcal{G}_0 := \Lambda$ and $\mathcal{G}_{n+1} := \mathcal{G}_n \cap \mathcal{G}_{N_n}(u_n)$, such that*

$$||F(u_n)||_{\mathfrak{s}_0,\gamma} \leq C_\star \varepsilon N_n^{-\kappa_1}. \tag{3.1.6}$$

*Moreover one has that $h_n := u_n - u_{n-1}$ (with $h_0 = 0$) satisfies*

$$||h_n||_{\mathfrak{s}_0+\mu,\gamma} \leq C_\star \varepsilon\gamma^{-1} N_n^{-\kappa_3}, \quad \kappa_3 := 9\nu + 2\mu. \tag{3.1.7}$$

*The Lipschitz norms are defined on the sets $\mathcal{G}_n$.*

$(N2)_n$ *the following estimates in high norms hold:*

$$||u_n||_{\mathfrak{s}_0+\kappa_2,\gamma} + \gamma^{-1}||F(u_n)||_{\mathfrak{s}_0+\kappa_2,\gamma} \le C_\star \varepsilon \gamma^{-1} N_n^{\kappa_1}. \tag{3.1.8}$$

*Finally, setting* $\mathcal{G}_\infty := \cap_{n\ge 0}\mathcal{G}_n$, *the sequence* $(u_n)_{n\ge 0}$ *converges in norm* $||\cdot||_{\mathfrak{s}_0+\mu,\mathcal{G}_\infty,\gamma}$ *to a function* $u_\infty$ *such that*

$$F(\lambda, u_\infty(\lambda)) \equiv 0, \qquad \sup_{\lambda\in\mathcal{G}_\infty} ||u_\infty(\lambda)||_{\mathfrak{s}_0+\mu} \le C\varepsilon\gamma^{-1}. \tag{3.1.9}$$

*Proof.* We proceed by induction.
We set $u_0 = h_0 = 0$, we get $(N1)_0$ and $(N2)_0$ by fixing

$$C_\star \ge \max\left\{||f(0)||_{\mathfrak{s}_0} N_0^{\kappa_1}, ||f(0)||_{\mathfrak{s}_0+\kappa_2} N_0^{-\kappa_1}\right\}.$$

We assume inductively $(Ni)_n$ for $i = 1, 2, 3$ for some $n \ge 0$ and prove $(Ni)_{n+1}$ for $i = 1, 2$.
By $(N1)_n$, $u_n \in A_n$ satisfies the conditions in Definition 3.1.17. Then, by definition, $\lambda \in \mathcal{G}_{n+1}$ implies that $\mathcal{L}_n := \mathcal{L}(u_n)$ is invertible with estimates (3.1.4), (used with $u = u_n$ and $N = N_n$).
Set

$$u_{n+1} := u_n + h_{n+1} \in A_{n+1}, \quad h_{n+1} := -\Pi_{n+1}\mathcal{L}_n^{-1}\Pi_{n+1}F(u_n), \tag{3.1.10}$$

which is well-defined. Indeed, $F(u_n) \in B_s$ implies, since $\mathcal{L}_n$ maps $A_{s+\nu} \to B_s$ that $h_{n+1} \in A_{n+1}$. By definition

$$F(u_{n+1}) = F(u_n) + \mathcal{L}_n h_{n+1} + \varepsilon\mathcal{Q}(u_n, h_{n+1}), \tag{3.1.11}$$

where, by condition $(F0)$ we have

$$\mathcal{Q}(u_n, h_{n+1}) := f(u_n + h_{n+1}) - f(u_n) - d_u f(u_n)h_{n+1}\,, \tag{3.1.12}$$

which is at least quadratic in $h_{n+1}$. Then, using the definition of $h_{n+1}$ in (3.1.10) we obtain

$$\begin{aligned} F(u_{n+1}) &= F(u_n) - \mathcal{L}_n\Pi_{n+1}\mathcal{L}_n^{-1}\Pi_{n+1}F(u_n) + \varepsilon\mathcal{Q}(u_n, h_{n+1}) \\ &= \Pi_{n+1}^\perp F(u_n) + \mathcal{L}_n\Pi_{n+1}^\perp\mathcal{L}_n^{-1}\Pi_{n+1}F(u_n) + \varepsilon\mathcal{Q}(u_n, h_{n+1}) \\ &= \Pi_{n+1}^\perp F(u_n) + \Pi_{n+1}^\perp\mathcal{L}_n\mathcal{L}_n^{-1}\Pi_{n+1}F(u_n) \\ &+ [\mathcal{L}_n, \Pi_{n+1}^\perp]\mathcal{L}_n^{-1}\Pi_{n+1}F(u_n) + \varepsilon\mathcal{Q}(u_n, h_{n+1}), \end{aligned} \tag{3.1.13}$$

hence, by using the fact that by $(F0)$ $[\mathcal{L}_n, \Pi_{n+1}^\perp] = \varepsilon[d_u f(\lambda, u_n), \Pi_{n+1}^\perp]$, one has

$$F(u_{n+1}) = \Pi_{n+1}^\perp F(u_n) + \varepsilon[d_u f(u_n), \Pi_{n+1}^\perp]\mathcal{L}_n^{-1}\Pi_{n+1}F(u_n) + \varepsilon\mathcal{Q}(u_n, h_{n+1}). \tag{3.1.14}$$

Now we need a technical Lemma to deduce the estimates (3.1.6) and (3.1.8) at the step $n + 1$. This Lemma guarantees that the scheme is quadratic, and the high norms of the approximate solutions and of the vector fields do not go to fast to infinity.

**Lemma 3.1.19.** *Set for simplicity*

$$K_n := ||u_n||_{\mathfrak{s}_0+\kappa_2,\gamma} + \gamma^{-1}||F(u_n)||_{\mathfrak{s}_0+\kappa_2,\gamma}, \quad k_n := \gamma^{-1}||F(u_n)||_{\mathfrak{s}_0,\gamma}. \tag{3.1.15}$$

*Then, there exists a constant $C_0 := C_0(\mu, d, \kappa_2)$ such that*

$$\begin{aligned} K_{n+1} &\leq C_0 N_{n+1}^{2\mu+4\nu}(1+k_n)^2 K_n, \\ k_{n+1} &\leq C_0 N_{n+1}^{-\kappa_2+\mu+2\nu} K_n(1+k_n) + C_0 N_{n+1}^{2\nu+2\mu} k_n^2 \end{aligned} \tag{3.1.16}$$

*Proof.* First of all, we note that, by conditions $(F2)-(F4)$, $\mathcal{Q}(u_n,\cdot)$ satisfies

$$||\mathcal{Q}(u_n,h)||_{s,\gamma} \leq ||h||_{\mathfrak{s}_0+\nu,\gamma}\left(||h||_{s+\nu,\gamma} + ||u_n||_{s+\nu,\gamma}||h||_{\mathfrak{s}_0+\nu,\gamma}\right), \tag{3.1.17a}$$

$$||\mathcal{Q}(u_n,h)||_{\mathfrak{s}_0+\nu,\gamma} \leq_s N_{n+1}^{2\nu}||h||^2_{\mathfrak{s}_0+\nu,\gamma}, \quad \forall h(\lambda) \in H_{n+1} \tag{3.1.17b}$$

where $h(\lambda) \in A_{n+1}$ is a Lipschitz family of functions depending on a parameter. The bound (3.1.17b) is nothing but the (3.1.17a) with $s = \mathfrak{s}_0 + \nu$, where we used the fact that $||u_n||_{\mathfrak{s}_0+\nu} \leq 1$ and the smoothing properties $(P1)$, that hold because $u_n \in A_n$ by definition and $h \in A_{n+1}$ by hypothesis.

Consider $h_{n+1}$ defined in (3.1.10). then we have

$$\begin{aligned} ||h_{n+1}||_{\mathfrak{s}_0+\kappa_2,\gamma} &\overset{(3.1.4b)}{\leq}_{\mathfrak{s}_0+\kappa_2} \gamma^{-1}N_{n+1}^{\mu}\Big(||F(u_n)||_{\mathfrak{s}_0+\kappa_2,\gamma} + \\ &\quad + ||u_n||_{\mathfrak{s}_0+\kappa_2,\gamma}||F(u_n)||_{\mathfrak{s}_0,\gamma}\Big), \end{aligned} \tag{3.1.18}$$

$$||h_{n+1}||_{\mathfrak{s}_0,\gamma} \overset{(3.1.4a)}{\leq}_{\mathfrak{s}_0} \gamma^{-1}N_{n+1}^{\mu}||F(u_n)||_{\mathfrak{s}_0,\gamma}. \tag{3.1.19}$$

Moreover, recalling that by (3.1.10) one has $u_{n+1} = u_n + h_{n+1}$, we get, by (3.1.18) and (3.1.19),

$$\begin{aligned} ||u_{n+1}||_{\mathfrak{s}_0+\kappa_2,\gamma} &\leq ||u_n||_{\mathfrak{s}_0+\kappa_2,\gamma}\left(1+\gamma^{-1}N_{n+1}^{\mu}||F(u_n)||_{\mathfrak{s}_0,\gamma}\right) \\ &+ \gamma^{-1}N_{n+1}^{\mu}||F(u_n)||_{\mathfrak{s}_0+\kappa_2,\gamma}. \end{aligned} \tag{3.1.20}$$

Now, we would like to estimate the norms of $F(u_{n+1})$. First of all, we can estimate the term $R_n := [d_u f(u_n), \Pi_{n+1}^{\perp}]\mathcal{L}_n^{-1}\Pi_{n+1}F(u_n)$ in (3.1.14), without using the commutator structure,

$$\begin{aligned} &||R_n||_{s,\gamma} \leq_s \gamma^{-1}N_{n+1}^{\mu+2\nu}\left(||F(u_n)||_{s,\gamma} + ||u_n||_{s,\gamma}||F(u_n)||_{\mathfrak{s}_0,\gamma}\right), \\ &||R_n||_{\mathfrak{s}_0,\gamma} \leq_{\mathfrak{s}_0+\kappa_2} \gamma^{-1}N_{n+1}^{-\kappa_2+\mu+2\nu}\Big(||F(u_n)||_{\mathfrak{s}_0+\kappa_2,\gamma} + ||u_n||_{\mathfrak{s}_0+\kappa_2,\gamma}||F(u_n)||_{\mathfrak{s}_0,\gamma}\Big), \end{aligned} \tag{3.1.21}$$

where we used the (3.1.4) to estimate $\mathcal{L}_n^{-1}$, the $(F3)$ for $d_u f$ and the smoothing estimates $(P1)-(P2)$. By (3.1.14), (3.1.21), (3.1.17b) and using $\varepsilon\gamma^{-1}\leq 1$ we obtain,

$$\begin{aligned} ||F(u_{n+1})||_{\mathfrak{s}_0,\gamma} \leq_{\mathfrak{s}_0} & \ ||\Pi^{\perp}_{N_{n+1}}F(u_n)||_{\mathfrak{s}_0,\gamma}+\varepsilon N_{n+1}^{2\nu}||h||^2_{\mathfrak{s}_0,\gamma} \\ & +\varepsilon\gamma^{-1}N_{n+1}^{-\kappa_2+\mu+2\nu}\left(||F(u_n)||_{\mathfrak{s}_0+\kappa_2,\gamma}+||u_n||_{\mathfrak{s}_0+\kappa_2,\gamma}||F(u_n)||_{\mathfrak{s}_0,\gamma}\right) \\ & \overset{(P2)}{\leq}_{\mathfrak{s}_0+\kappa_2} N_{n+1}^{-\kappa_2+\mu+2\nu}(||F(u_n)||_{\mathfrak{s}_0+\kappa_2,\gamma}+||u_n||_{\mathfrak{s}_0+\kappa_2,\gamma}||F(u_n)||_{\mathfrak{s}_0,\gamma}) \\ & + \quad \varepsilon\gamma^{-2}N_{n+1}^{2\nu+2\mu}||F(u_n)||^2_{\mathfrak{s}_0,\gamma}. \end{aligned} \tag{3.1.22}$$

Following the same reasoning as in (3.1.22), by using the estimates (3.1.21), (3.1.17a), (3.1.18), (3.1.19) and (P2), we get the estimate in high norm

$$\begin{aligned} ||F(u_{n+1})||_{\mathfrak{s}_0+\kappa_2,\gamma} & \leq (||F(u_n)||_{\mathfrak{s}_0+\kappa_2,\gamma}+||u_n||_{\mathfrak{s}_0+\kappa_2,\gamma}||F(u_n)||_{\mathfrak{s}_0,\gamma})\times \\ & \times \left(1+N_{n+1}^{\mu+2\nu}+N_{n+1}^{2\mu+4\mu}\gamma^{-1}||F(u_n)||_{\mathfrak{s}_0,\gamma}\right). \end{aligned} \tag{3.1.23}$$

From the (3.1.22) follows directly the second of the (3.1.16), while collecting together (3.1.20) and (3.1.23) one obtain the first of (3.1.16). ☐

By (3.1.6) we have that

$$k_n\leq\varepsilon\gamma^{-1}C_\star N_n^{-\kappa_1}\leq 1, \tag{3.1.24}$$

if $\varepsilon\gamma^{-1}$ is small enough. Then one has, for $N_0$ large enough,

$$K_{n+1}\overset{(3.1.24),(3.1.16)}{\leq} N_{n+1}^{4\nu+2\mu}2K_n\leq C_\star\varepsilon\gamma^{-1}N_n^{\frac{3}{2}(4\nu+2\mu)}N_n^{\kappa_1}\leq C_\star\varepsilon\gamma^{-1}N_{n+1}^{\kappa_1} \tag{3.1.25}$$

where we used the fact that, by formula (3.2.62), one has $3(2\nu+\mu)+\kappa_1=\frac{3}{2}\kappa_1$. This proves the $(N2)_{n+1}$. In the same way,

$$\begin{aligned} k_{n+1} & \overset{(N2)_n,(3.1.16)}{\leq} 2N_{n+1}^{-\kappa_2+\mu+2\nu}\varepsilon\gamma^{-1}N_n^{\kappa_1}C_0+\varepsilon^2\gamma^{-2}C_0N_n^{\frac{3}{2}(2\nu+2\mu)}N_n^{-2\kappa_1} \\ & \leq \quad \varepsilon\gamma^{-1}C_\star N_{n+1}^{-k_1}, \end{aligned} \tag{3.1.26}$$

where we used again the formula (3.2.62). This proves the $(N1)_{n+1}$. The bound (3.1.7) follows by $(N2)_n$ and by using Lemma 3.1.19 to estimate the norm of $h_n$. Then we get

$$||u_{n+1}||_{\mathfrak{s}_0+\mu,\gamma}\leq||u_0||_{\mathfrak{s}_0+\mu,\gamma}+\sum_{k=1}^{n+1}||h_k||_{\mathfrak{s}_0+\mu,\gamma}\leq\sum_{k=1}^{\infty}C_\star\varepsilon\gamma^{-1}N_k^{-\kappa_3}\leq 1, \tag{3.1.27}$$

if $\varepsilon\gamma^{-1}$ is small enough. This means that $(Ni)_{n+1}$, $i=1,2$, hold.

Now, if $\varepsilon\gamma^{-1}$ is small enough, we have by $(N1)_n$ that the sequence $(u_n)_{\geq 0}$ is a Cauchy sequence in norm $||\cdot||_{\mathfrak{s}_0+\mu,\gamma}$, on the set $\mathcal{G}_\infty = \cap_{n\geq 0}\mathcal{G}_n$. Hence, we have that $u_\infty := \lim_{n\to\infty} u_n$ solves the equation since

$$||F(u_\infty)||_{\mathfrak{s}_0+\mu,\gamma} \leq \lim_{n\to\infty} ||F(u_n)||_{\mathfrak{s}_0+\mu,\gamma} \leq \lim_{n\to\infty} N_n^{\mu} C_\star \varepsilon N_n^{-\kappa_1} = 0. \tag{3.1.28}$$

This concludes the proof of Theorem 3.1.18. □

## 3.2 An Existence Theorem for torus embeddings

In this Section we prove an abstract Theorem on a wide class of vector field. Roughly speaking we provide a scheme that, given a vector field, produces a set of parameters, inductively defined, for which it is possible to find a torus embedding. We introduce all the relevant notations and tools we need. In particular we define our phase space, a suitable subspace of vector fields for which we are able to deal with, and the type of change of variables we need in order to perform our algorithm.

### 3.2.1 The Phase Space

Our starting point is an infinite dimensional space with a product structure $V_{a,p} := \mathbb{C}^d \times \mathbb{C}^d \times \ell_{a,p}$. Here $\ell_{a,p}$ is a scale of separable Hilbert spaces endowed with norms $\|\cdot\|_{a,p}$, in particular this means that $\|f\|_{a,p} \le \|f\|_{a',p'}$ if $(a,p) \le (a',p')$ lexicographically. The space $\ell_{0,0}$ is endowed with a bilinear scalar product

$$f, g \in \ell_{0,0} \mapsto f \cdot g \in \mathbb{C}$$

such that

$$\|w\|_{0,0}^2 = w \cdot \overline{w}\,, \quad |g \cdot f| \le \|g\|_{0,0}\|f\|_{0,0} \le \|g\|_{a,q}\|f\|_{0,0}. \tag{3.2.29}$$

We denote the set of variables $\mathtt{V} := \{\theta_1, \ldots, \theta_d, y_1, \ldots, y_d, w\}$. Moreover we make the following assumption on the scale $\ell_{a,p}$.

**Hypothesis 5.** *We assume that there is a non-decreasing family $(F^{(K)})_{K\ge 0}$ of subspaces of $\ell_{a,p}$ such that $\cup_{K\ge 0} F^{(K)}$ is dense in $\ell_{a,p}$ for any $p \ge 0$, and that there are projectors*

$$\Pi_{F_K} : \ell_{0,0} \to F_K, \quad \Pi_{F_K}^{\perp} := \mathbb{1} - \Pi_{F_K}, \tag{3.2.30}$$

*such that one has for $\alpha, \beta \ge 0$*

$$\|\Pi_{F_K} w\|_{a+\alpha,p+\beta} \le e^{\alpha K} K^{\beta} \|w\|_{a,p} \ \ \forall w \in \ell_{a,p}, \tag{3.2.31a}$$

$$\|\Pi_{F_K}^{\perp} w\|_{a,p} \le e^{-\alpha K} K^{-\beta} \|w\|_{a+\alpha,p+\beta}, \quad \forall w \in \ell_{a+\alpha,p+\beta}. \tag{3.2.31b}$$

We shall need two parameters, $\mathfrak{p}_0 < \mathfrak{p}_1$. Precisely $\mathfrak{p}_0 > d/2$ is needed in order to have the Sobolev embedding and thus the algebra properties, while $\mathfrak{p}_1$ will be chosen very large and is needed in order to define the phase space.

**Definition 3.2.20** (Phase space)**.** *We consider the toroidal domain*

$$D_{a,p}(s,r) := \mathbb{T}_s^d \times D_{a,p}(r) = \mathbb{T}_s^d \times B_{r^2} \times \mathbf{B}_{r,a,p}\,, \subset V_{a,p} \tag{3.2.32}$$

*where*

$$\mathbb{T}^d_s := \left\{\theta \in \mathbb{C}^d \,:\, \mathrm{Re}(\theta) \in \mathbb{T}^d, \ \max_{h=1,\dots,d} |\mathrm{Im}\, \theta_h| < s\right\},$$

$$B_{r^2} := \left\{y \in \mathbb{C}^d \,:\, |y|_1 < r^2\right\}, \qquad \mathbf{B}_{r,a,p} := \left\{w \in \ell_{a,p} \,:\, \|w\|_{a,\mathfrak{p}_1} < r\right\},$$

*and we denote by* $\mathbb{T}^d := (\mathbb{R}/2\pi\mathbb{Z})^d$ *the* $d$*-dimensional torus.*

Fix some numbers $s_0, a_0 \geq 0$ and $r_0 > 0$. Given $s \leq s_0$, $a, a' \leq a_0$, $r \leq r_0$, $p, p' \geq \mathfrak{p}_0$ consider maps

$$\begin{aligned} f \,:\, \mathbb{T}^d_s \times D_{a',p'}(r) &\to \mathbb{C}^d \times \mathbb{C}^d \times \ell_{a,p} \\ (\theta, y, w) &\to (f^{(\theta)}(\theta, y, w), f^{(y)}(\theta, y, w), f^{(y)}(\theta, y, w)), \end{aligned} \tag{3.2.33}$$

with

$$f^{(\mathtt{v})}(\theta, y, w) = \sum_{l \in \mathbb{Z}^d} f^{(\mathtt{v})}_l(y, w) e^{\mathrm{i} l \cdot \theta}, \quad \mathtt{v} = \theta, y, w,$$

where $f^{(\mathtt{v})}_l(y, w) \in \mathbb{C}^d$ if $\mathtt{v} = \theta, y$ while $f^{(w)}_l(y, w) \in \ell_{a,p}$.

**Remark 3.2.21.** *We think of these maps as families of torus embeddings from* $\mathbb{T}^d_s$ *into* $V_{a,p}$ *depending parametrically on* $y, w \in D_{a',p'}(r)$*, and this is the reason behind the choice of the norm; see below.*

We define a norm (pointwise on $y, w$) by setting

$$\|f\|^2_{s,a,p} := \|f^{(\theta)}\|^2_{s,p} + \|f^{(y)}\|^2_{s,p} + \|f^{(w)}\|^2_{s,a,p} \tag{3.2.34}$$

where

$$\|f^{(\theta)}\|_{s,p} := \begin{cases} \dfrac{1}{s_0} \sup\limits_{i=1,\dots,d} \|f^{(\theta_i)}(\cdot, y, w)\|_{H^p(\mathbb{T}^d_s)} & s \leq s_0 \neq 0 \\ \sup\limits_{i=1,\dots,d} \|f^{(\theta_i)}(\cdot, y, w)\|_{H^p(\mathbb{T}^d)}, & s = s_0 = 0 \end{cases} \tag{3.2.35}$$

$$\|f^{(y)}\|_{s,p} := \frac{1}{r_0^2} \sum_{i=1}^{d} \|f^{(y_i)}(\cdot, y, w)\|_{H^p(\mathbb{T}^d_s)} \tag{3.2.36}$$

$$\|f^{(w)}\|_{s,a,p} := \frac{1}{r_0} \left[ \left( \sum_{l \in \mathbb{Z}^d} \|f^{(w)}_l(y, w)\|^2_{a,\mathfrak{p}_0} \langle l \rangle^{2p} e^{2s|l|} \right)^{\frac{1}{2}} + \left( \sum_{l \in \mathbb{Z}^d} \|f^{(w)}_l(y, w)\|^2_{a,p} \langle l \rangle^{2\mathfrak{p}_0} e^{2s|l|} \right)^{\frac{1}{2}} \right] \tag{3.2.37}$$

where $H^p(\mathbb{T}^d_s) = H^p(\mathbb{T}^d_s; \mathbb{C})$ is the standard Sobolev space with norm

$$\|u(\cdot)\|^2_{H^p(\mathbb{T}^d_s)} := \sum_{l \in \mathbb{Z}^d} |u_l|^2 e^{2s|l|} \langle l \rangle^{2p}, \qquad \langle l \rangle := \max\{1, |l|\}. \tag{3.2.38}$$

Note that trivially $\|\partial^{p'}_\theta u\|_{H^p(\mathbb{T}^d_s)} = \|u\|_{H^{p+p'}(\mathbb{T}^d_s)}$.

**Remark 3.2.22.** *Note that if $\ell_{a,p} = H^p(\mathbb{T}^r_a)$ then fixing $\mathfrak{p}_0 \geq (d+r)/2$ we have that $\|\cdot\|_{s,a,p}$ in* (3.2.37) *is equivalent to $\|\cdot\|_{H^p(\mathbb{T}^d_s \times \mathbb{T}^r_a)}$*

It is clear that any $f$ as in (3.2.33) can be identified with "unbounded" vector fields by writing[1]

$$f \leftrightarrow \sum_{\mathtt{v}=\theta,y,w} f^{(\mathtt{v})}(\theta, y, w)\partial_{\mathtt{v}}. \tag{3.2.39}$$

Similarly, provided that $|f^{(\theta)}(\theta, y, w)|$ is small for all $(\theta, y, w) \in \mathbb{T}^d_s \times D_{a,p}(r)$ we may lift $f$ to a map

$$\Phi := (\theta + f^{(\theta)}, y + f^{(y)}, w + f^{(w)}) : \mathbb{T}^d_s \times D_{a',p'} \to \mathbb{T}^d_{s_1} \times \mathbb{C}^d \times \ell_{a,p}\,, \quad \text{for some } s_1 \geq s\,, \tag{3.2.40}$$

and if we set $\|\theta\|_{s,a,p} := 1$ we can write

$$\|\Phi^{(\mathtt{v})}\|_{s,a,p} = \|\mathtt{v}\|_{s,a,p} + \|f^{(\mathtt{v})}\|_{s,a,p}\,,\ \mathtt{v} = \theta, y, w\,.$$

Note that

$$\|y\|_{s,a,p} = r_0^{-2}|y|_1\,, \quad \|w\|_{s,a,p} = r_0^{-1}\|w\|_{a,p}.$$

**Remark 3.2.23.** *Note that if $\|f\|_{s,a,\mathfrak{p}_1} \sim \rho$ is small enough one has*

$$\Phi : \mathbb{T}^d_s \times D_{a,p}(r) \to \mathbb{T}^d_{s+\rho s_0} \times D_{a,p}(r + \rho r_0).$$

**Definition 3.2.24.** *Fix $0 \leq \rho, \delta \leq 1/2$, and consider two differentiable maps $\Phi = \mathbb{1} + f$, $\Psi = \mathbb{1} + g$ as in* (3.2.40) *such that for all $p \geq \mathfrak{p}_0$, $2\rho s_0 \leq s \leq s_0$, $2\rho r_0 \leq r \leq r_0$ and $0 \leq a \leq a_0(1-2\delta)$ one has*

$$\Phi, \Psi : \mathbb{T}^d_s \times D_{a+\delta a_0,p}(r - \rho r_0) \to \mathbb{T}^d_{s+\rho s_0} \times D_{a,p}(r).$$

*If*

$$\begin{aligned} \mathbb{1} = \Psi \circ \Phi : \mathbb{T}^d_{s-2\rho s_0} \times D_{a+2\delta a_0,p}(r - 2\rho r_0) &\longrightarrow \mathbb{T}^d_s \times D_{a,p}(r) \\ (\theta, y, w) &\longmapsto (\theta, y, w) \end{aligned} \tag{3.2.41}$$

*we say that $\Psi$ is a left inverse of $\Phi$ and write $\Phi^{-1} := \Psi$.*

*Moreover fix $\nu \geq 0$, $0 \leq \delta' \leq 1/2$. Then for any $F : \mathbb{T}^d_s \times D_{a+\delta' a_0,p+\nu}(r) \to V_{a,p}$, with $0 \leq a \leq a_0(1 - 2\delta - \delta')$, we define the "pushforward" of $F$ as*

$$\Phi_* F := d\Phi(\Phi^{-1})[F(\Phi^{-1})] : \mathbb{T}^d_{s-2\rho s_0} \times D_{a+(2\delta+\delta')a_0,p+\nu}(r - 2\rho r_0) \to V_{a,p}\,. \tag{3.2.42}$$

[1] Vector fields are defined by giving their action on functions. In order to describe them as *vectors* we use the identification between differential and gradient given by the bilinear scalar product. On the $\theta, y$ components we use the euclidean product, while on $\ell_{a,p}$ we use the one defined in formula (3.2.29).

We need to introduce parameters $\xi \in \mathcal{O}_0$ a compact set in $\mathbb{R}^{\mathfrak{d}}$. Given any compact $\mathcal{O} \subseteq \mathcal{O}_0$ we consider Lipschitz families of vector fields

$$F : \mathbb{T}^d_s \times D_{a',p'}(r) \times \mathcal{O} \to V_{a,p}\,, \tag{3.2.43}$$

and say that they are *bounded* vector fields when $p = p'$ and $a = a'$. Given a positive number $\lambda$ we introduce the weighted Lipschitz norm

$$\|F\|_{\vec{v},p} = \|u\|_{\lambda,\mathcal{O},s,a,p} := \sup_{\xi\in\mathcal{O}} \|F(\xi)\|_{s,a,p} + \lambda \sup_{\xi\neq\eta\in\mathcal{O}} \frac{\|F(\xi) - F(\eta)\|_{s,a,p}}{|\xi-\eta|}\,. \tag{3.2.44}$$

and we shall drop the labels $\vec{v} = (\lambda, \mathcal{O}, s, a)$ when this does not cause confusion.

**Remark 3.2.25.** *Note that in some applications one might need to assume a higher regularity in $\xi$. In this case it is convenient to define the weighted $q_1$-norm*

$$\|F\|_{\vec{v},p} = \|F\|_{\lambda,\mathcal{O},s,a,p} := \sum_{\substack{h\in\mathbb{N}^{\mathfrak{d}}\\ |h|\le q_1}} \lambda^{|h|} \sup_{\xi\in\mathcal{O}} \|\partial^h_\xi F(\xi)\|_{s,a,p}.$$

*Througout the paper we shall always use the Lipschitz norm* (3.2.44)*, although all the properties hold verbatim also for the $q_1$-norm. The only delicate question is how to extend functions defined in a compact subset $\mathcal{O}' \subset \mathcal{O}$ to the whole domain. By the Kirszbraun theorem this is trivial in the Lipschitz case, while it requires some care for the $q_1$-norm.*

We are interested in vector fields defined on a scale of Hilbert spaces; precisely we shall fix $\delta, \nu, q \ge 0$ and consider vector fields

$$F : \mathbb{T}^d_s \times D_{a+\delta a_0,p+\nu}(r) \times \mathcal{O} \to V_{a,p}\,, \tag{3.2.45}$$

for some $s < s_0$, $a + \delta a_0 \le a_0$, $r \le r_0$ and all $p + \nu \le q$. We shall denote by $\mathcal{V}_{\vec{v},p}$ with $\vec{v} = (\lambda, \mathcal{O}, s, a, r)$ the space of vector fields as in (3.2.45) with $\delta = 0$.

We use the Lipschitz norm (3.2.44) also for maps $\Phi = (\theta + f^{(\theta)}, y + f^{(y)}, w + f^{(w)})$ with

$$\Phi : \mathbb{T}^d_s \times D_{a',p'}(r) \times \mathcal{O} \to \mathbb{T}^d_{s_1} \times D_{a,p}(r_1)\,.$$

We now define *tame vector fields*, i.e. vector fields behaving "tamely" when composed with maps $\Phi$.

**Definition 3.2.26.** *Fix parameters $\lambda$, $\delta, \nu \ge 0$, a large $q \ge \mathfrak{p}_1 + \nu + 3$ and a set $\mathcal{O}$ . Consider*

$$F : \mathbb{T}^d_s \times D_{a+\delta a_0,p+\nu}(r) \times \mathcal{O} \to V_{a,p}\,, \quad \forall\ p \le q$$

*for $s \le s_0, a \le a_0, r \le r_0$; for $s_1 \le s$ set*

$$\vec{v} = (\lambda, \mathcal{O}, s, a) \quad \vec{v}_1 = (\lambda, \mathcal{O}, s_1, a), \quad \vec{v}_2 = (\lambda, \mathcal{O}, s_1, a + \delta a_0).$$

*We say that $F$ is $C^3$*-tame *if there exists a scale of constants $C_{\vec{v},p}(F) = C_{\lambda,\mathcal{O},s,a,p}(F)$, such that the following holds.*

*For all $\mathfrak{p}_0 \le p \le p' \le q$ consider any map $\Phi = (\theta + f^{(\theta)}, y + f^{(y)}, w + f^{(w)})$ with*

$$\Phi : \mathbb{T}^d_{s_1} \times D_{a',p'}(r_1) \times \mathcal{O} \to \mathbb{T}^d_s \times D_{a+\delta a_0, p+\nu}(r) , \quad \text{for some} \quad r_1 \le r , s_1 \le s$$

*and any three vector fields*

$$h_1, h_2, h_3 : \mathbb{T}^d_{s_1} \times D_{a',p'}(r_1) \times \mathcal{O} \to V_{a+\delta a_0, p+\nu}.$$

*Then for all $y, w \in D_{a',p'}(r_1)$ one has*

$$\text{(F1)} \qquad \|F(\Phi)\|_{\vec{v}_1,p} \le C_{\vec{v},p}(F) + C_{\vec{v},\mathfrak{p}_0}(F)\|\Phi\|_{\vec{v}_2,p+\nu},$$

$$\begin{aligned}\text{(F2)} \qquad \|d_{\mathtt{V}}F(\Phi)[h_1]\|_{\vec{v}_1,p} \le\ & \big(C_{\vec{v},p+1}(F) + C_{\vec{v},\mathfrak{p}_0+1}(F)\|\Phi\|_{\vec{v}_2,p+\nu}\big)\|h_1\|_{\vec{v}_2,\mathfrak{p}_0+\nu} \\ & + C_{\vec{v},\mathfrak{p}_0+1}(F)\|h_1\|_{\vec{v}_2,p+\nu},\end{aligned}$$

$$\begin{aligned}\text{(F3)} \qquad \|d^2_{\mathtt{V}}F(\Phi)[h_1,h_2]\|_{\vec{v}_1,p} \le\ & \big(C_{\vec{v},p+2}(F) + C_{\vec{v},\mathfrak{p}_0+2}(F)\|\Phi\|_{\vec{v}_2,p+\nu}\big)\|h_1\|_{\vec{v}_2,\mathfrak{p}_0+\nu}\|h_2\|_{\vec{v}_2,\mathfrak{p}_0+\nu} \\ & + C_{\vec{v},\mathfrak{p}_0+2}(F)\big(\|h_1\|_{\vec{v}_2,p+\nu}\|h_2\|_{\vec{v}_2,\mathfrak{p}_0+\nu} + \|h_1\|_{\vec{v}_2,\mathfrak{p}_0+\nu}\|h_2\|_{\vec{v}_2,p+\nu}\big),\end{aligned}$$

$$\begin{aligned}\text{(F4)} \quad \|d^3_{\mathtt{V}}F(\Phi)[h_1,h_2,h_3]\|_{\vec{v}_1,p} \le\ & \big(C_{\vec{v},p+3}(F) + C_{\vec{v},\mathfrak{p}_0+3}(F)\|\Phi\|_{\vec{v}_2,p+\nu}\big)\|h_1\|_{\vec{v}_2,\mathfrak{p}_0+\nu}\|h_2\|_{\vec{v}_2,\mathfrak{p}_0+\nu}\|h_3\|_{\vec{v}_2,\mathfrak{p}_0+\nu} \\ & + C_{\vec{v},\mathfrak{p}_0+3}(F)\Big(\|h_1\|_{\vec{v}_2,p+\nu}\|h_2\|_{\vec{v}_2,\mathfrak{p}_0+\nu}\|h_3\|_{\vec{v}_2,\mathfrak{p}_0+\nu} \\ & + \|h_1\|_{\vec{v}_2,\mathfrak{p}_0+\nu}\|h_2\|_{\vec{v}_2,p+\nu}\|h_3\|_{\vec{v}_2,\mathfrak{p}_0+\nu} + \|h_1\|_{\vec{v}_2,\mathfrak{p}_0+\nu}\|h_2\|_{\vec{v}_2,\mathfrak{p}_0+\nu}\|h_3\|_{\vec{v}_2,p+\nu}\Big)\end{aligned}$$

*Here $d_{\mathtt{V}}F$ is the differential of $F$ w.r.t. the variables $\mathtt{V}$.*
*We say that a* bounded *vector field $F$ is tame with scale of constants $C_{\vec{v},p}(F)$ if (F1), (F2), (F3), (F4) above hold with $\nu = 0$. We call $C_{\vec{v},p}(F)$ the $p$*-tameness constants of $F$.

*More generally we say that $F$ is $C^k$-tame if it satisfies tame estimates as above up to its k'th derivatives.*

**Remark 3.2.27.** *Formula* (3.2.34) *depends on the point $(y, w)$, hence it is* not *a norm for vector fields and this is very natural in the context of* Sobolev regularity. *Indeed in the scale of domains $D_{a,p}(s, r)$ one controls only the $\mathfrak{p}_1$ norm of $w$ (see Definition 3.2.20), and hence there is no reason for which one may have*

$$\sup_{(y,w)\in D_{a,p}(r)} \|f\|_{s,a,p} < \infty .$$

*Naturally if one fixes* $\mathtt{p} = \mathfrak{p}_1$ *one may define as norm of* $F$ *the quantity* $\sup_{(y,w)\in D_{a,\mathfrak{p}_1}(r)} \|F\|_{s,a,\mathfrak{p}_1}$.

*The motivation for choosing the norm* (3.2.34) *instead of the standard operator norm is the following. Along the algorithm we need to control commutators of vector fields. In the analytic case, i.e. if* $s_0 \neq 0$, *one may keep* p *fixed and control the derivatives via* Cauchy estimates *by reducing the analyticity, so the phase space can be defined in terms of the fixed* p. *However, since we do not want to add the hypothesis* $s_0 \neq 0$, *we have to leave* p *as a parameter and use* tameness *properties of the vector field (see Definition 3.2.26) as in the Sobolev Nash-Moser schemes.*

**Remark 3.2.28.** *Definition 3.2.26 is quite natural if one has to deal with functions and vector fields which are merely differentiable. In order to clarify what we have in mind we consider the following example. Let* $L$ *be a linear operator*

$$L : H^p(\mathbb{T}^d) \to H^p(\mathbb{T}^d)\,.$$

*In principle there is no reason for* $L$ *to satisfy a bound like*

$$\|Lu\|_p \lessdot \|L\|_{\mathcal{L},p}\|u\|_{\mathfrak{p}_0} + \|L\|_{\mathcal{L},\mathfrak{p}_0}\|u\|_p \tag{3.2.46}$$

*where* $\|\cdot\|_{\mathcal{L},p}$ *is the* $H^p$*-operator norm. However if* $L = T_a$ *is a multiplication operator, i.e.* $T_a u = au$ *for some* $a \in H^p(\mathbb{T}^d)$ *then it is well known that*

$$\|T_a u\|_p \le \kappa_p(\|a\|_p\|u\|_{\mathfrak{p}_0} + \|a\|_{\mathfrak{p}_0}\|u\|_p)$$

*which is* (3.2.46) *since* $\|a\|_p = \|T_a\|_{\mathcal{L},p}$. *In this case we may set for all* $p \le q$ $C_p(T_a) = \kappa_q\|a\|_p$. *This is of course a trivial (though very common in the applications) example in which the tameness constants and the operator norm coincide; we preferred to introduce Definition 3.2.26 since it is the most general class in which we are able to prove our result.*

**Remark 3.2.29.** *The constants* $C(F)$ *essentially play the rôle of the* norm *of* $F$. *Indeed one could start by fixing* $C_{s,a,\mathfrak{p}_0}$ *as the norm of* $F$ *and then* minimize *over the scales of constants which satisfy all our constraints, this would most probably produce a well defined norm, but we did not pursue the subject.*

**Lemma 3.2.30.** *Consider any two* $C^k$-tame *vector fields* $F, G \in \mathcal{V}_{\vec{v},p}$, *then:*

*(i) for any* $0 \le j < q$ *one has that* $\partial^j_{\theta_l} F$ *is a* $C^k$*-tame vector field for* $p \le q - j$ *with tameness scale of constants* $C_{p+j}(F)$ *for any* $l = 1, \dots, d$.

*(ii) For any* $0 \le j \le k$ *one has that* $\partial^j_{y_l} F, \partial^j_w F[w^j]$ *are* $C^{k-j}$*-tame vector field for* $p \le q$ *with tameness scale of constants* $C_{p+j}(F)$. *Moreover if* $F$ *is a polynomial of degree* $\le k$ *in* $y, w$ *then it is* $C^h$*-tame for any* $h \ge 0$.

*(iii)* *The commutator* $[G, F]$ *is a* $C^{k-1}$*-tame vector field for* $p \leq q - 1$ *with scale of constants*

$$C_{\vec{v},p}([F, g]) \leq C_{\vec{v},p+1}(F) C_{\vec{v},\mathfrak{p}_0+\nu+1}(g) + C_{\vec{v},\mathfrak{p}_0+1}(F) C_{\vec{v},p+\nu+1}(g). \tag{3.2.47}$$

*Proof.* Items $(i)$ and $(ii)$ follows by the definition of the norm $\|\cdot\|_{s,a,p}$ and by Definition 3.2.26. Item $(iii)$ follows By Lemma B.182 in Appendix B. □

**Lemma 3.2.31** (**Conjugation).** *Consider a tame left invertible map* $\Phi = \mathbb{1} + f$ *with tame inverse* $\Phi^{-1} = \mathbb{1} + g$ *as in Definition 3.2.24 such that* (3.2.41) *holds. Assume that the fields* $f, g$ *are such that* $C_{\vec{v},p}(f) = C_{\vec{v},p}(g) \leq \rho$ *for* $\rho > 0$ *small. Consider a constant* $\mathfrak{p}_0 \geq 0$ *and any tame vector field* $F : D_{a,p+\nu}(s, r) \to V_{a,p}$*, then the pushforward*

$$G := \Phi_* F : D_{a,p+\nu}(s - 2\rho s_0, r - 2\rho r_0) \to V_{a-2\delta a_0,p} \tag{3.2.48}$$

*is tame with scale of constants*

$$C_{\vec{v}_1,p}(G) \leq (1 + C_{\vec{v}_2,\mathfrak{p}_0+\nu+1}(f)) C_{\vec{v},p}(F) + C_{\vec{v},\mathfrak{p}_0}(F)(1 + C_{\vec{v}_2,\mathfrak{p}_0+\nu+1}(f)) C_{\vec{v}_2,p+\nu+1}(f), \tag{3.2.49}$$

*with* $\vec{v} := (\lambda, \mathcal{O}, s, a)$*,* $\vec{v}_1 := (\lambda, \mathcal{O}, s - 2\rho s_0, a - 2\delta a_0)$ *and* $\vec{v}_2 := (\lambda, \mathcal{O}, s - \rho s_0, a - \delta a_0)$*.*

*Proof.* The proof is in Appendix B. □

**Definition 3.2.32.** *It will be convenient to extend the projection* $\Pi_{F_k}$ *to* $\mathbb{C}^{2n} \times \ell_{a,p}$ *by setting*

$$\Pi_{F_k}(\theta, y, w) = (\theta, y, \Pi_{F_k} w).$$

*Given* $K > 0$ *and a vector field* $f \in \mathcal{V}_{\vec{v},p}$ *we define the projection* $\Pi_K f \in \mathcal{V}_{\vec{v},p}$ *as*

$$\Pi_K f := \sum_{|l| \leq K} \Pi_{F_K} f_l(\theta, y, \Pi_{F_K} w) e^{il\cdot\theta} \tag{3.2.50}$$

*Then we define* $E^K$ *as the subspace of* $\mathcal{V}_{\vec{v},p}$ *where* $\Pi_K$ *acts as the identity.*

**Lemma 3.2.33.** *Given a* $C^k$*-tame vector field* $F \in \mathcal{V}_{\vec{v},p}$ *one has for any* $p_1, d_1, d_2 \geq 0$ *and for* $K$ *large*

*(P1)* $\Pi_{(K)} F$ *is a* $C^k$*-tame vector field with* $C_{s+d_1,a+d_2,p+p_1}(\Pi_{(K)} f) \leq e^{d_1 K + d_2 K} K^{p_1} C_{s,a,p}(F)$*.*

*(P2)* $(\mathbb{1} - \Pi_{(K)}) F$ *is a* $C^{k-1}$*-tame vector field with* $C_{s,a,p}((\mathbb{1} - \Pi_{(K)}) F) \leq K^{-p_1} e^{-d_1 K - d_2 K} C_{s+d_1,a+d_2,p+p_1}(F)$*, provided* $s + d_1 \leq s_0, a + d_2 \leq a_0, p + p_1 \leq q$*.*

*Proof.* Consider a map $\Phi$ as in definition 3.2.26 and the composition of $\Pi_K F$ with $\Phi$. For this purpose set for instance

$$G^{(w)}(\theta, y, w) = (\Pi_K F^{(w)}) \circ \Phi = \sum_{|l| \le K} \Pi_{F_K} \Big[ F^{(w)}(\Phi^{(\theta)}, \Phi^{(y)}, \Pi_{F_K}\Phi^{(w)}) \Big]_l e^{\mathrm{i} l \cdot \theta} \,,$$

we have

$$\begin{aligned} \|G^{(w)}\|^2_{s+d_1, a+d_2, p+p_1} &= \sum_{|l|\le K} \Big( \|\Pi_{F_k} G_l^{(w)}\|^2_{a+d_2, \mathfrak{p}_0} \langle l \rangle^{2p+2p_1} e^{2|l|(s+d_1)} \\ &\quad + \|\Pi_{F_k} G_l^{(w)}\|^2_{a+d_1, p+p_1} \langle l \rangle^{2\mathfrak{p}_0} e^{2|l|(s+d_1)} \Big) \\ &\le K^{p_1} e^{(d_1+d_2)K} \|F^{(w)}(\Phi^{(\theta)}, \Phi^{(y)}, \Pi_{F_K}\Phi^{(w)})\|^2_{s,a,p} \\ &\le K^{p_1} e^{(d_1+d_2)K} (C_{\vec{v},p}(F) + C_{\vec{v},\mathfrak{p}_0} \|\Pi_{F_K}\Phi)\|_{\vec{v}_2, p+\nu}) \end{aligned} \tag{3.2.51}$$

For the other components and to estimate the orthogonal projector $\Pi_K^{\perp}$ one can follows the same reasoning. One one has the results it is sufficient to apply Definition (3.2.26) to obtain the results on the tameness constants. □

Given a vector field $F \in \mathcal{V}_{\vec{v},p}$, we introduce the notation

$$F^{(v,0)}(\theta) := F^{(v)}(\theta, 0, 0), \quad F^{(v,v')}(\theta)[\cdot] := d_{v'} F^{(v)}(\theta, 0, 0)[\cdot], \quad v = \theta, y, w \,, \; v' = y, w \tag{3.2.52}$$

and the subspaces

$$\begin{aligned} &\mathcal{N} := \Big\{ V \in \mathcal{V}_{\vec{v},p} : V = V^{(\theta,0)}(\theta) \cdot \partial_\theta + V^{(w,w)}(\theta) w \cdot \partial_w) \Big\} \,, \\ &\mathcal{A}^{(v,0)} := \{ v \in \mathcal{V}_{\vec{v},p}(s,r) \; : \; V = V^{(v,0)}(\theta) \cdot \partial_v \} \,, \; v = y, w \\ &\mathcal{A}^{(v,v')} := \{ v \in \mathcal{V}_{\vec{v},p} \; : \; V = V^{(v,v')}(\theta) v' \cdot \partial_v \} \,, \; v = \theta, y, w \,, \quad v' = y, w \\ &\mathcal{A} := \mathcal{A}^{(y,0)} \oplus \mathcal{A}^{(y,y)} \oplus \mathcal{A}^{(y,w)} \oplus \mathcal{A}^{(w,0)} \\ &\mathcal{R} := \big\{ V \in \mathcal{V}_{\vec{v},p} : V = V - \Pi_{\mathcal{N}} V - \Pi_{\mathcal{A}} V \big\} \,, \end{aligned} \tag{3.2.53}$$

where $\Pi_{\mathcal{N}}$ and $\Pi_{\mathcal{A}}$ are the projections on the subspaces $\mathcal{N}$ and $\mathcal{A}$. Note that a vector field in $\mathcal{A}$ can be extended to analytic functions of $y, w \in \mathbb{C}^n \times \ell_{a,p}$. Clearly the subspaces in (3.2.53) are disjoint so that for any vector field $F \in \mathcal{V}_{a,p}(s,r) := \mathcal{N} \oplus \mathcal{A} \oplus \mathcal{R}$ we may split

$$F = N + X + R = \Pi_{\mathcal{N}} F + \Pi_{\mathcal{A}} F + \Pi_{\mathcal{R}} F \,. \tag{3.2.54}$$

Another important class of vector fields is the following.

**Definition 3.2.34 (Linear vector fields).** *We consider vector fields $f \in \mathcal{A}$ with $f : \mathcal{O} \times D_{a,p+\nu}(s,r) \to V_{a,p}$ such that*

$$f = \sum_{v=y,w} f^{(v,0)}\partial_v + (f^{(y,y)}y + f^{(y,w)} \cdot w)\partial_y\,, \quad f^{(y,y)} \in \mathrm{Mat}(d \times d)\,, \quad f^{(y_i,w)} \in \ell_{a,p} \tag{3.2.55}$$

*It is convenient to define:*

$$|f|_{\vec{v},p} := \sum_{k=y,w} \|f^{(k,0)}\|_{\vec{v},p} + \sup_{i,j=1,\dots,d} \|f^{(y_i y_j)}\|_{\vec{v},p} + \sup_i \|f^{(y_i,w)}\|_{\vec{v},p} \tag{3.2.56}$$

*We also denote by $\mathcal{B}$ the subspace of* **bounded** *vector fields in $\mathcal{A}$ which satisfy* (3.2.55).

**Remark 3.2.35.** *Note that $E^{(K)}$ are dense in the subset of $\mathcal{A}$ which satisfy* (3.2.55) *w.r.t. the norm* (3.2.56).

As explained before, we must deal with a class of non-linear tame vector field $F$ with some special properties. Here we define the "subspace" of such fields.

**Definition 3.2.36 (Subspaces).** *Consider $\mathcal{E}$ a subspace of ${\mathcal{V}_{\vec{v},p}}^2$ such that for any $F \in \mathcal{E} \cap \mathcal{A}$ satisfies 3.2.34. We say that the map $\Phi$ is $\mathcal{E}$–*preserving *if $\Phi_*$ maps $\mathcal{E}$ in itself. We denote by $\hat{\mathcal{B}} \subset \mathcal{B}$ the subset of linear bounded vector fields defined as follows:*

$$\hat{\mathcal{B}} := \{f \in \mathcal{B} \ : \ \Phi^1_f \quad \text{is } \mathcal{E} \text{ preserving}\}, \tag{3.2.57}$$

For this dynamical system we wish to construct an invariant torus. This is done by producing a change of variables $\Phi$ generated by a vector field in the Lie algebra $\hat{\mathcal{B}}$ such that

$$\Pi_{\mathcal{A}}\Phi_*(F) = 0.$$

This construction is performed iteratively, by a quadratic scheme where the input is $F$ and a set of parameters $\mathcal{O}$, the output is a change of variables $\Phi$ defined for all $\xi \in \mathcal{O}$ set of *good parameters* $\xi \in \mathcal{O}_\infty \subseteq \mathcal{O}$ such that $\Pi_{\mathcal{A}}\Phi_*(F) = 0$ holds for $\xi \in \mathcal{O}_\infty$ the main point is that $\mathcal{O}_\infty$ is described iteratively in a relatively explicit way. Then the point becomes verifying that such set is not empty. We now give the important definition of set of parameters for which we are able to run the algorithm. Fix parameters

$$\tau > 0, \ \ \mu \geq \tau, \ \ \alpha < \frac{1}{4}, \ \ \gamma, \rho > 0, K > 0 \text{ large} \quad \mathfrak{r} \ll 1, \tag{3.2.58}$$

[2]For instance $\mathcal{E}$ is the subspace of Hamiltonian vector fields.

**Definition 3.2.37 (Good Parameters $J_{K,\mathfrak{r},\gamma}(F)$ ).** *Given a tame vector field*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p}$$

*such that for $\xi \in \mathcal{O} \subseteq \mathcal{O}_0$ one has $F \in \mathcal{E}$. We say that a compact set $\mathcal{O}' \subset \mathcal{O}$ is a set of* good parameters*, $\mathcal{O}' \in J_{K,\mathfrak{r},\gamma}(F)$ if the following holds.*
*For all $\xi \in \mathcal{O}'$ there exist a vector field $g$ such that*

*(a)* $g \in \hat{\mathcal{B}} \cap \mathcal{V}_{\vec{v}_1,p} \cap E^{(K)}$ *with* $\vec{v}_1 = (\gamma, \mathcal{O}', s, a)$

*(b) one has*

$$|g|_{\vec{v}_1,p} \lessdot \gamma^{-1} K^{\mu}(|\Pi_{\mathcal{A}} G|_{\vec{v}_1,p} + K^{\alpha(p-\mathfrak{p}_1)} |\Pi_{\mathcal{A}} G|_{\vec{v}_1,\mathfrak{p}_1} \gamma^{-1} C_{\vec{v},p}(G)). \tag{3.2.59}$$

*(c) setting* $X = \Pi_K \Pi_{\mathcal{A}} F$ *and* $u := (\Pi_K \Pi_{\mathcal{A}} \mathrm{ad}(\Pi_{\mathcal{A}^\perp} F)[g] - X)$*, one has*

$$\begin{aligned} |u|_{\vec{v},\mathfrak{p}_1} &\le \gamma^{-1} \mathfrak{r} K^{\mu} C_{\vec{v},\mathfrak{p}_1}(G) |X|_{\vec{v},\mathfrak{p}_1}, \\ |u|_{\vec{v},\mathfrak{p}_2} &\le \gamma^{-1} K^{\mu} \Big( |X|_{\vec{v},\mathfrak{p}_2} C_{\vec{v},\mathfrak{p}_1}(G) + K^{\alpha(\mathfrak{p}_2 - \mathfrak{p}_1)} |X|_{\vec{v},\mathfrak{p}_1} C_{\vec{v},\mathfrak{p}_2}(G) \Big) ; \end{aligned} \tag{3.2.60}$$

Now we set some notations we need to state our Theorem.
Given $\varepsilon \ll 1$, $K_0 > 1$, $r_0 > 0$ and $a_0, s_0 \ge 0$, we set for all $\gamma_0, \Theta_0, \mathcal{A}_0 > 0$,

$$\begin{aligned} &K_n = (K_0)^{(3/2)^n},\ \gamma_n = \gamma_0(1 + \frac{1}{2^n}), \quad \mathfrak{A}_n = \mathfrak{A}_0 2^{-n-2} + \mathfrak{A}_{n-1}, \quad a_n = a_0(1 - \frac{1}{2}\sum_{j=1}^n 2^{-j}), \\ &r_n = r_0(1 - \frac{1}{2}\sum_{j=1}^n 2^{-j}), \quad s_n = s_0(1 - \frac{1}{2}\sum_{j=1}^n 2^{-j}), \Pi_n := \Pi^{(K_n)}, \quad \Pi_n^\perp := \mathbb{1} - \Pi_n, \\ &E_n = E^{(K_n)}, \quad \rho_n := \frac{1}{2^{n+8}}, n \ge 1, \ \ \rho_0 = 0, \quad \mathfrak{r}_n = \delta K_n^{-\kappa_1}, \end{aligned} \tag{3.2.61}$$

We will also use the shorter notation

$$\|\cdot\|_{\vec{v}_{(n,i)},p} = \|\cdot\|_{\gamma_n,\mathcal{O}_n,s_n - i\rho_n s_0, a_n - i\rho_n a_0, p} \quad \|\cdot\|^{(n,i)}_{p,\mathcal{G}} := \|\cdot\|_{s_n - i\rho_n s_0, a_n - i\rho_n a_0, p, \lambda_n, \mathcal{G}}, \quad i = 0, \dots, 6$$

where $\mathcal{O}_n \subseteq \mathcal{O}_0$. We use the second notation when we want to specify a different domain $\mathcal{G}$ for the parameters $\xi$. The The same notation is used for the tameness constants $C_{\vec{v}_{(n,i)},p}(\cdot)$ . For all $\alpha \in [0, 1/4)$, $p \ge \mathfrak{p}_1$, and given $\mu = \mu(\alpha, \mathfrak{p}_1, d)$ we choose constants $\kappa_i$ such that

$$\begin{aligned} &\kappa_2 \ge \kappa_3 + \mu + 1, \quad \kappa_2 \ge 3\mu + 2\kappa_1 + 4, \quad \kappa_1 \ge 6\mu + 6\nu + 24, \quad \alpha \le \frac{1}{3}, \\ &\kappa_1 \ge 27\mu + 6\nu + 24 \\ &\alpha\kappa_2 \le \frac{\kappa_1}{2}, \quad 2\mu + \nu + 4 + \alpha\kappa_2 + \kappa_3 + 1 \le (5/3)\kappa_1 \\ &2\mu + \nu + \alpha\kappa_2 - \alpha\mu \le \kappa_1 \end{aligned} \tag{3.2.62}$$

Note that if

$$\alpha = \frac{1}{20}, \qquad \kappa_2 = 2\kappa_1 + 6\mu + 8, \qquad \kappa_3 \le \kappa_1 + 3\mu, \tag{3.2.63}$$

then conditions in (3.2.62) simplify. We can ask for

$$\kappa_1 \ge 27\mu + 6\nu + 24 \tag{3.2.64}$$

Set moreover $\mathfrak{p}_0 = (d+1)/2$ and $\mathfrak{p}_1 = \mathfrak{p}_0 + \mu$, $\mathfrak{p}_2 = \mathfrak{p}_0 + \kappa_2$

At each step we change the coordinates in order to simplify the *homological equation.* This is done in two ways:

- We use quasi-identical transformations $\Phi_n$ in order to decouple the $\theta, y$ from the $w$ components and in order to translate the approximate solutions at zero. The vector fields which generate such changes of coordinates are in $\hat{\mathcal{B}}$ and are found by solving a *homological equation.* We will show that this can be done provided that we restrict $\xi$ to the *good parameters.*

- We introduce as free parameters a convergent sequence of $\mathcal{E}$ preserving change of variables $\mathcal{T}_{n+1} = \mathcal{T}_{n+1}(\xi) : D_{a_n,\mathfrak{p}_1}(s_n, r_n) \to D_{a_n - \rho_n a_0,\mathfrak{p}_1}(s_n + \rho_{n+1}s_0, r_n + \rho_{n+1}r_0)$ such that for all $n \ge 0$ and $\delta > 0$ small one has

  (i) $\mathcal{T}_{n+1}$ is left-invertible and for any tame $G \in \mathcal{E} \cap \mathcal{V}_{\vec{v}_{(n,0)},p}$ one has $\hat{G} := (\mathcal{T}_{n+1})_* G$ is tame with scale of constants
  $$\begin{aligned} C^{(n,2)}_{\mathfrak{p}_1,\mathcal{O}_0}(\hat{G}) &\lessdot C^{(n,0)}_{\mathfrak{p}_1,\mathcal{O}_0}(G)(1 + \delta K_n^{-1}) \\ C^{(n,2)}_{\mathfrak{p}_2,\mathcal{O}_0}(\hat{G}) &\lessdot C^{(n,0)}_{\mathfrak{p}_2,\mathcal{O}_0}(G) + C^{(n,0)}_{\mathfrak{p}_0,\mathcal{O}_0}(G)\delta K_n^{\kappa_3} \end{aligned} \tag{3.2.65}$$

  (ii)
  $$\begin{aligned} \|(\mathcal{T}_{n+1} - \mathbb{1})w\|^{(n,1)}_{\mathfrak{p}_1,\mathcal{O}_0} &\lessdot \delta K_n^{-1} \|w\|^{(n,0)}_{\mathfrak{p}_1,\mathcal{O}_0}, \\ \|(\mathcal{T}_{n+1} - \mathbb{1})w\|^{(n,1)}_{\mathfrak{p}_2,\mathcal{O}_0} &\lessdot \delta(\|w\|^{(n,0)}_{\mathfrak{p}_2,\mathcal{O}_0} + K_n^{\kappa_3}\|w\|^{(n,0)}_{\mathfrak{p}_1,\mathcal{O}_0}) \end{aligned} \tag{3.2.66}$$

  (iii) for $n \ge 0$ one has
  $$\Pi_{\mathcal{A}}(\mathcal{T}_{n+1})_* F = 0, \qquad \forall F \in \mathcal{N} \cap \mathcal{R} \tag{3.2.67}$$

**Remark 3.2.38.** *Note that the almost-identical changes of variables $\Phi_n$ are essentially fixed by the algorithm. On the other hand the iterative algorithm converges for any choice of $T_n$ as above. In particular this extra* degree of freedom *can be used in the applications, when one can produce a privileged system of coordinates in which it is simpler to solve the homological equation. Indeed in classic KAM schemes the $T_n$ are iteratively fixed so that they diagonalize (up to order $\varepsilon_n$) the linearized vector field in the $w$ component and diagonalize $\omega(\theta)$.*

We also introduce the following notation. Given a sequence of invertible linear maps $\mathcal{H}_n$ defined and Lipschitz in $\mathcal{O}_0$ with $\mathcal{H}_0 = \mathbb{1}$ we set $F_n := (\mathcal{H}_n)_* F$ for $n \geq 0$ and

$$\hat{F}_{n-1} := (\mathcal{T}_n)_* F_{n-1} = \hat{N}_{n-1} + \hat{X}_{n-1} + \hat{R}_{n-1}, \qquad n \geq 1,$$

(where the splitting (3.2.54) is used with obvious meaning of symbols). Our main abstract result is the following

**Theorem 3.2.39 (Abstract KAM).** *Let $\mathfrak{p}_1, \mathfrak{p}_2$ large enough and consider $F = F(\xi) \in \mathcal{E} \cap \mathcal{V}_{\vec{v}_{(0,0)},p}$ for some $s_0, a_0 \geq 0$, $p \in [\mathfrak{p}_1, \mathfrak{p}_2)$ with $\xi \in \mathcal{O}_0 \subseteq \mathbb{R}^r$ a Lipschitz family of tame vector fields of the form*

$$F := N_0 + G := N_0 + N_1 + X + R, \tag{3.2.68}$$

*where $N_0 \in \mathcal{N}$ preserves all the subspace $E^{(K)}$ and $N_1 := \Pi_{\mathcal{N}} G$, $X := \Pi_{\mathcal{A}} G$, $R := \Pi_{\mathcal{R}} G$. Fix $\gamma_0 > 0$, set $\vec{v}_0 := (\gamma_0, \mathcal{O}_0, s_0, a_0)$ and*

$$\Theta := \gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(\Pi_{\mathcal{N}^\perp} G), \quad \mathfrak{A}_0 = \gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(G) \quad \delta := \gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_1}(\Pi_{\mathcal{A}} G), \tag{3.2.69}$$

*There exist $C_0(d, \mathfrak{p}_0)$ and $C_1(d, \mathfrak{p}_0)$ and $\epsilon_0 = \epsilon_0(d, \mathfrak{p}_0)$ (small) such that if*

$$\Theta \leq \epsilon_0, \qquad K_0^{C_0} \delta \leq \epsilon_0, \qquad \mathfrak{A}_0 K_0^{-C_1} \leq \epsilon_0, \tag{3.2.70}$$

*then there exists a sequence of changes of variables $\mathcal{H}_n$ defined and Lipschitz on $\mathcal{O}_0$ such that $\mathcal{H}_0 = \mathbb{1}$ and the $\mathcal{H}_n = \Phi_n \circ \mathcal{T}_n \circ \mathcal{H}_{n-1}$ that converges for all $\xi \in \mathcal{O}_0$ to some change of variables*

$$\overline{\mathcal{H}} = \overline{\mathcal{H}}(\xi) : D_{a_0,p}(s_0/2, r_0/2) \longrightarrow D_{\frac{a_0}{2},p}(s_0, r_0), \tag{3.2.71}$$

*such that the following holds.*

*Setting $F_n := (\mathcal{H}_n)_* F := N_0 + G_n$, then for any sequence of sets $\mathcal{O}_n$ defined inductively for $n \geq 1$*

$$\mathcal{O}_n \in \mathcal{O}_{n-1} \cap J_{K_{n-1},\mathfrak{r}_{n-1},\gamma_{n-1}}((\mathcal{T}_n)_*((\mathcal{H}_{n-1})_* F)), \tag{3.2.72}$$

*with $\mathfrak{r}_n := \gamma_0 \delta K_{n-1}^{-\kappa_1}$, one has that $\Phi_n = \mathbb{1} + f_n$ is generated by $g_n \in \hat{\mathcal{B}}$ given in Definition (3.2.37) and such that with*

$$C_{\vec{v}_{(n-1,2)},\mathfrak{p}_1}(g_n) \leq \delta K_{n-1}^{-\kappa}, \quad C_{\vec{v}_{(n-1,2)},\mathfrak{p}_2}(g_n) \leq \delta K_{n-1}^{\kappa}, \qquad \kappa = \kappa_1 - \mu, \tag{3.2.73}$$

*and $\mu = 5(\eta + \nu + 3)$ where $\eta$ the loss of regularity in Definition 3.2.37. Moreover one has*

$$\begin{gathered}
\gamma_n^{-1} C_{\vec{v}_n,\mathfrak{p}_1}(\Pi_{\mathcal{N}^\perp} G_n) \leq \delta \mathfrak{A}_n, \quad \gamma_n^{-1} C_{\vec{v}_n,\mathfrak{p}_2}(\Pi_{\mathcal{N}^\perp} G_n) \leq \delta \mathfrak{A}_0 K_n^{\kappa_1} \\
\gamma_n^{-1} C_{\vec{v}_n,\mathfrak{p}_1}(G_n) \leq \mathfrak{A}_n, \quad \gamma_n^{-1} C_{\vec{v}_n,\mathfrak{p}_2}(G_n) \leq \mathfrak{A}_0 K_n^{\kappa_1}, \\
\gamma_n^{-1} C_{\vec{v}_n,\mathfrak{p}_1}(\Pi_{\mathcal{A}} G_n) \leq C_\star \delta K_n^{-\kappa_1},
\end{gathered} \tag{3.2.74}$$

*and defining* $F_\infty := (\overline{\mathcal{H}})_* F$ *one has*

$$\Pi_{\mathcal{A}} F_\infty = 0 \quad \forall \xi \in \overline{\mathcal{G}} := \bigcap_{n \geq 0} \mathcal{O}_n \tag{3.2.75}$$

*and*

$$\gamma_0^{-1} C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{N}} F_\infty - N_0) \lessdot 2\mathfrak{A}_0, \quad \gamma_0^{-1} C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{R}} F_\infty) \lessdot 2\epsilon_0$$

*with* $\vec{v} := (\lambda, \overline{\mathcal{G}}, s_0/2, a_0/2)$.

**Remark 3.2.40.** *Note that we do not require that the* $\mathcal{H}_n$ *are of the form* $\mathbb{1} + \Psi_n$ *with* $\Psi_n$ *in* $\hat{\mathcal{B}}$*. In principle one can impose this condition but this gives a further constraint on the good parameter set, so that one may not be able to prove that it is non empty. Moreover in the applications, we need to prove the following two properties on the good parameters sets* $\mathcal{O}_n$*:*

- $\mathrm{meas}(\mathcal{O}_n^c) \sim \gamma_0^a$ *for some* $a > 0$*;*
- $\mathrm{meas}(\mathcal{O}_n \backslash \mathcal{O}_{n+1}) \sim \frac{\gamma_0^b}{n^2}$ *for* $b \leq a$*.*

We have formulated our theorem in a very general form, as a draw-back we define our sets of good parameters in a very implicit way. If we add to the definition of $\mathcal{E}$ the condition that $\mathcal{E}$ is compatible with the projections on $\mathcal{A}, \mathcal{N}, \mathcal{R}$ then a more explicit expression is possible.

**Definition 3.2.41 (Property $\mathcal{P}$).** *Let* $\mathcal{E}$ *be a subspace as in Definition 3.2.36. We say that* $\mathcal{E}$ *has the property* $\mathcal{P}$ *if*

$$\forall g \in \hat{\mathcal{B}}, F \in \mathcal{E}: \quad [g, F] \in \mathcal{E}, \quad \forall g, h \in \hat{\mathcal{B}}: \quad \Pi_{\mathcal{A}}[g, h] \in \hat{\mathcal{B}} \tag{3.2.76}$$

*and moreover that* $\mathcal{E}$ *is compatible with the projections on* $E^{(K)}, \mathcal{A}, \mathcal{R}, \mathcal{N}$*, i.e. that for all* $F \in \mathcal{E}$, $\Pi_{\mathcal{U}} F \in \mathcal{E}$ *with* $\mathcal{U}$ *one of the above subspaces.*

As in Definition 3.2.37 let us fix a tame vector field $F \in \mathcal{V}_{\vec{v},p}$, namely $F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p}$, such that $F \in \mathcal{E}$ for all $\xi \in \mathcal{O}$. We denote

$$\Omega(\theta) = \Omega_F(\theta) := \partial_w F^{(w)}(\theta, 0, 0), \qquad \omega = \omega(\theta) = \omega_F(\theta) := F^{(\theta)}(\theta, 0, 0), \qquad N := \omega(\theta) \cdot \partial_\theta + \Omega(\theta) w \cdot \partial_w. \tag{3.2.77}$$

**Definition 3.2.42** (Melnikov conditions )**.** *We say that a compact set* $\mathcal{O}' \subset \mathcal{O}$ *satisfies the Melnikov conditions for* $F$, $\mathcal{O}' \in \mathcal{M}_{K,\mathfrak{r},\gamma}(F)$ *if the following holds.*
*There exist linear operators* $W_\pm : \mathcal{O}' \times E^{(K)} \cap \mathcal{A}^{(w,0)} \to E^{(K)} \cap \mathcal{A}^{(w,0)}$ *and* $W_0 : \mathcal{O}' \times E^K \cap (\mathcal{A}^{(y,0)} \cup \mathcal{A}^{(y,y)}) \to \mathcal{A}^{(y,0)} \cup \mathcal{A}^{(y,y)}$ *such that, for any vector field* $X \in \mathcal{A} \cap E^{(K)} \cap \mathcal{E} \cap \mathcal{V}_{\vec{v},p}$ *setting*

$$\mathcal{W}X := W_0\big[X^{(y)}(\theta) + X^{(y,y)}(\theta) y\big] \cdot \partial_y + X^{(y,w)}(\theta)[W_+ w]) \cdot \partial_y + W_- X^{(w)}(\theta) \cdot \partial_w, \tag{3.2.78}$$

*one has:*

*(a) the vector field* $\mathcal{W}X$ *is in* $\hat{\mathcal{B}}$

*(b) one has*

$$|\mathcal{W}X|_{\vec{v}_1,p} \lessdot \gamma^{-1}K^{\eta}(|X|_{\vec{v}_1,p} + K^{\alpha(p-\mathfrak{p}_1)}|X|_{\vec{v}_1,\mathfrak{p}_1}C_{\vec{v},p}(G)). \tag{3.2.79}$$

*(c) setting* $u := (\Pi_K \mathrm{ad}(\Pi_{\mathcal{N}}F)[\mathcal{W}X] - X)$, and $\vec{v}_1 = (\gamma, \mathcal{O}', s, a)$ *one has*

$$\begin{aligned} |u|_{\vec{v},\mathfrak{p}_1} &\le \gamma^{-1}\mathfrak{r}K^{\eta}C_{\vec{v},\mathfrak{p}_1}(G)|X|_{\vec{v},\mathfrak{p}_1}, \\ |u|_{\vec{v},\mathfrak{p}_2} &\le \gamma^{-1}K^{\eta}\left(|X|_{\vec{v},\mathfrak{p}_2}C_{\vec{v},\mathfrak{p}_1}(G) + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)}|X|_{\vec{v},\mathfrak{p}_1}C_{\vec{v},\mathfrak{p}_2}(G)\right); \end{aligned} \tag{3.2.80}$$

**Lemma 3.2.43.** *Take a tame vector field* $F \in \mathcal{E}$ *and assume that* $\mathcal{E}$ *has property* $\mathcal{P}$. *Assume also that* $F = N_0 + G$ *with* $\gamma^{-1}C_{\vec{v},\mathfrak{p}_1}(G) \le K$, *then the sets* $\mathcal{O}' \in \mathcal{M}_{K,\mathfrak{r},\gamma}(F)$ *are in* $J_{K,\mathfrak{r},\gamma}(F)$ *provided that we fix* $\mu = 5(\eta + \nu + 5)$.

*Proof.* In the Appendix C.3. □

### 3.2.2 The iterative scheme

We first analyze one step of the iteration we want to perform in order to prove Theorem 3.2.39. Essentially given the field $F_n$, we show that for parameters in the set $\mathcal{G}_n$, we are able to construct a change of variables such that the vector field in such new variables has the size of the projection on $\mathcal{A}$ much smaller than the previous one. For simplicity from now on we drop for all the objects indexed by $n$ while we will write $+$ instead of $n+1$.

**Lemma 3.2.44 (The KAM step).** *Consider constants* $\gamma_0 > 0$, $\mathfrak{A}_0 > 0$, $\delta \ge 0$, *and* $K_0$ *of Theorem 3.2.39. Take constant* $K \ge K_0$ *and set* $K_+ = K^{\frac{3}{2}}$ *and a set* $\mathcal{O} \subseteq \mathcal{O}_0$. *Take some constants* $\gamma, a, r, s$ *such that*

$$\gamma_0 \le \gamma \le 2\gamma_0,\ \frac{a_0}{16} \le a \le a_0,\ \frac{s_0}{16} \le s \le s_0,\ \frac{r_0}{16} \le r \le r_0, \tag{3.2.81}$$

*and* $1 > \rho_+ > 0$ *such that*

$$r - 16\rho_+ r_0 > 0,\ \text{ if } s_0 \neq 0 \text{ then } s - 16\rho_+ s_0 > 0,\ \text{ if } a_0 \neq 0 \text{ then } a - 16\rho_+ a_0 > 0. \tag{3.2.82}$$

*Consider also a vector field*

$$F : D_{a,p+\nu}(s,r) \to V_{a,p}, \tag{3.2.83}$$

*such that* $F \in \mathcal{E}$ *for* $\xi \in \mathcal{O}$. *Write* $F = N_0 + G = N_0 + \Pi_{\mathcal{N}}G + \Pi_{\mathcal{A}}G + \Pi_{\mathcal{R}}G := N_0 + M + X + R$. *Set* $\vec{v} := (\gamma, \mathcal{O}, s, a)$ *and*

$$\Gamma_p := \gamma^{-1}C_{\vec{v},p}(G), \quad \Theta_p := \gamma^{-1}C_{\vec{v},p}(\Pi_{\mathcal{N}^{\perp}}G), \quad \delta_p := \gamma^{-1}C_{\vec{v},p}(\Pi_{\mathcal{A}}G), \tag{3.2.84}$$

*Assume that*

$$\Gamma_{\mathfrak{p}_1} \le 2\mathfrak{A}_0, \quad \Theta_{\mathfrak{p}_1} < 2\delta\mathfrak{A}_0, \quad \rho_+^{-1} K_+^{\mu+\nu+6}\delta_{\mathfrak{p}_1} \le \epsilon, \tag{3.2.85}$$

*where $\mu$ is the loss of regularity in Definition 3.2.37 and $\epsilon$ is a constant depending only on $\mathfrak{p}_0$ and $d$. Consider the map $\mathcal{T}_+$ that satisfies formulæ*(3.2.65), (3.2.66) *and* (3.2.67) *with $\mathcal{T}_{n+1} \rightsquigarrow \mathcal{T}_+$,$\rho_{n+1} \rightsquigarrow \rho_+$, $a_n, r_n, s_n \rightsquigarrow a, r, s$ $K_n \rightsquigarrow K$ and set*

$$\hat{F} := (\mathcal{T}_+)_* F = N_0 + \hat{M} + \hat{X} + \hat{R} : D_{a-2\rho_+ a_0, p+\nu}(s - 2\rho_+ s_0, r - 2\rho_+ r_0)$$

*If $\epsilon$ is small enough then the following hold:*

*for any $\mathcal{O}_+ \subseteq \mathcal{O} \cap J_{K_+,\mathfrak{r},\gamma}(\hat{F})$ with $\mathfrak{r} := \delta K^{-\kappa_1}$ one has that, for $\xi \in \mathcal{O}_+$, there exists an invertible (see Def. 3.2.24) $\mathcal{E}$-preserving change of coordinates $\widetilde{\Phi}_+ := \mathbb{1} + \widetilde{f}_+$ such that*
*(i)*

$$\widetilde{\Phi}_+ := \mathbb{1} + \widetilde{f}_+ : D_{a-4\rho_+ a_0, p}(s - 4\rho_+ s_0, r - 4\rho_+ r_0) \longrightarrow D_{a-2\rho_+ a_0, p}(s - 2\rho_+ s_0, r - 2\rho_+ r_0), \tag{3.2.86}$$

*with $\widetilde{f}_+ := \Phi_+ - \mathbb{1} \in E^{K_+}$;*
*(ii) there exists a Lipschitz extension $f_+$ of $\widetilde{f}_+$ such that $C^{(3)}_{p,\mathcal{O}_0}(f_+) \le C^{(3)}_{p,\mathcal{O}_+}(\widetilde{f}_+)$. Moreover, setting $\Phi_+ := \mathbb{1} + f_+$, one has, for any $s_+, a_+, r_+$ with*

$$s - 16\rho_+ s_0 \le s_+ \le s - 10\rho_+ s_0, \quad r - 16\rho_+ r_0 \le r_+ \le r - 10\rho_+ r_0, \quad a - 16\rho_+ a_0 \le a_+ \le a - 10\rho_+ a_0, \tag{3.2.87}$$

*that*

$$F_+ := (\Phi_+)_* \hat{F} = N_0 + G_+ : D_{a_+, p+\nu}(s_+, r_+), \tag{3.2.88}$$

*(iii) writing $F_+ := N_0 + M_+ + X_+ + R_+$with*

$$M_+ := \Pi_{\mathcal{N}}(\Phi_+)_* \hat{F} - N_0, \quad X_+ := \Pi_{\mathcal{A}}(\Phi_+)_* \hat{F}, \quad R_+ := \Pi_{\mathcal{R}}(\Phi_+)_* \hat{R}, \tag{3.2.89}$$

*and setting $\gamma/2 \le \gamma_+ \le \gamma$, $\vec{v}_+ := (\gamma_+, \mathcal{O}_+, a_+, s_+)$, one has that*

$$\Gamma_{+,p} := \gamma_+^{-1} C_{\vec{v}_+,p}(G_+), \quad \Theta_{+,p} := \gamma_+^{-1} C_{\vec{v}_+,p}(\Pi_{\mathcal{N}^\perp} G_+), \quad \delta_{+,p} := \gamma_+^{-1} C_{\vec{v}_+,p}(\Pi_{\mathcal{A}} G_+)$$

*satisfy the following estimates:*

$$\begin{aligned} \Gamma_{+,\mathfrak{p}_2} &\le (1 + K_+^{\mu+2}\delta_{\mathfrak{p}_1}) \left[ \Gamma_{\mathfrak{p}_2} + K_+^{\mu+\nu+2}\Theta_{\mathfrak{p}_2} + \delta_{\mathfrak{p}_1} K_+^{2\mu+\nu+4+\alpha(\kappa_2-\mu)}(\Gamma_{\mathfrak{p}_2} + K^{\kappa_3+1}) \right], \\ \Theta_{+,\mathfrak{p}_2} &\le (1 + K_+^{\mu+2}\delta_{\mathfrak{p}_1}) \left[ \Theta_{\mathfrak{p}_2} K_+^{\mu+\nu+4} + \delta_{\mathfrak{p}_1} K_+^{\mu+\nu+4+\alpha(\kappa_2-\mu)}(\Gamma_{\mathfrak{p}_2} + \delta K^{\kappa_3+1}) \right] \end{aligned} \tag{3.2.90}$$

*and*

$$\Gamma_{+,\mathfrak{p}_1} \le (1 + K_+^{\mu+\nu+6}\delta_{\mathfrak{p}_1})\Gamma_{\mathfrak{p}_1}, \qquad \Theta_{+,\mathfrak{p}_1} \le (1 + K_+^{\mu+\nu+6}\delta_{\mathfrak{p}_1})\Theta_{\mathfrak{p}_1}, \tag{3.2.91}$$

*In addiction to this one has*

$$\delta_{+,\mathfrak{p}_1} \le K_+^{-(1-\alpha)(\kappa_2-\mu)+\mu+2}(\Gamma_{\mathfrak{p}_2} + \delta K^{\kappa_3+1})\delta_{\mathfrak{p}_1} + K_+^{-\kappa_2+\mu+3}\Theta_{\mathfrak{p}_2} \\ + K_+^{\mu+2-\frac{2}{3}\kappa_1}\delta_{\mathfrak{p}_1} + \delta_{\mathfrak{p}_1}^2 K_+^{2\mu+2\nu+8}\left[1 + (\Gamma_{\mathfrak{p}_2} + \delta K^{\kappa_3+1})K_+^{-(\kappa_2-\mu)}\right] \tag{3.2.92}$$

*Proof.* First of all we note that, by (3.2.65), we have

$$\gamma^{-1} C^{(2)}_{\mathfrak{p}_1,\mathcal{O}_+}(\hat{X}) \le \gamma^{-1} C^{(0)}_{\mathfrak{p}_1,\mathcal{O}_+}(X)(1 + \delta K_+^{-1}) \le 2\delta_{\mathfrak{p}_1}, \tag{3.2.93}$$

for $K_+$ large. Our aim is to define the vector field $g_+$ as the "approximate" solution of the equation

$$\Pi_{\mathcal{A}}\Pi_{K_+}[g_+, N_0 + \hat{M} + \hat{R}] = \Pi_{\mathcal{A}}\Pi_{K_+}\hat{X} \tag{3.2.94}$$

By Definition if $\xi \in \mathcal{O}_+$ then $\xi \in J_{K_+,\mathfrak{r},\gamma}(\hat{F})$ and hence we can find $g_+$ with properties (a),(b),(c) of Definition 3.2.37.

By Definition 3.2.37 of "good" parameters, we see that for $\xi \in \mathcal{O}_+$ it is possible to find an approximate solution of (3.2.94) if $\hat{R} \equiv 0$. However we use the results contained in Appendix C.3 to define a solution of the whole equation:

we set

$$g_+ := \left(\sum_{k=0}^{4} (\mathcal{W}_+ \Pi_{\mathcal{A}}\Pi_{K_+}\mathrm{ad}\hat{R})^k\right) \mathcal{W}_+ \Pi_{K_+}\hat{X} \tag{3.2.95}$$

where $\mathcal{W}_+$ satisfies conditions $(a), (b), (c)$ of Definition 3.2.37.

By using estimates (3.2.59), one gets

$$C^{(2)}_{p,\mathcal{O}_+}(g_+) \lessdot \gamma^{-1} K_+^{\mu}\left(C^{(2)}_{p,\mathcal{O}_+}(\Pi_{K_+}\hat{X}) + K_+^{\alpha(p-\mathfrak{p}_1)} C^{(2)}_{\mathfrak{p}_1,\mathcal{O}_+}(\Pi_{K_+}\hat{X})\gamma^{-1} C^{(2)}_{p,\mathcal{O}_+}(\hat{G})\right), \tag{3.2.96}$$

and hence, using (3.2.65) and (3.2.67) and $\Gamma_{\mathfrak{p}_1} \le K_0$,

$$C^{(2)}_{\mathfrak{p}_1,\mathcal{O}_+}(g_+) \le \gamma^{-1} K_+^{\mu} C^{(0)}_{\mathfrak{p}_1,\mathcal{O}_+}(X)(1 + \delta K^{-1})(1 + \Gamma_{\mathfrak{p}_1}) \le K_+^{\mu+2}\delta_{\mathfrak{p}_1}, \\ C^{(2)}_{\mathfrak{p}_2,\mathcal{O}_+}(g_+) \lessdot K_+^{\mu}\left[(\Theta_{\mathfrak{p}_2} + \delta_{\mathfrak{p}_1}(\Gamma_{\mathfrak{p}_2} + \Gamma_{\mathfrak{p}_1}\delta K^{\kappa_3})K_+^{\alpha(\mathfrak{p}_2-\mu)})\right] \tag{3.2.97}$$

Moreover, by condition $(a)$ of definition 3.2.37, one has that $g_+ \in \hat{\mathcal{B}}$. Now, if $\epsilon$ in (3.2.85) is small enough , one can apply Lemma B.178 and conclude that the transformation of coordinates $\widetilde{\Phi}_+ = \mathbb{1} + \widetilde{f}_+$ generated by the flow of the vector field $g_+$ is well defined and moreover $C^{(3)}_{p,\mathcal{O}_+}(\widetilde{f}_+) \le 2C^{(2)}_{p,\mathcal{O}_+}(g_+)$. Finally, by taking $\epsilon$ possibly smaller, one choose $\epsilon$ small enough in order to fit condition (B.15) of Lemma B.179. Hence (3.2.86) hold. This proves item (i).

Let us check (ii). In the following we write $\vec{v}_1 := (\gamma_+, \mathcal{O}_+, a - 8\rho a_0, s - 8\rho s_0)$ and $\vec{v}_+ = (\gamma_+, \mathcal{O}_+, a_+, s_+)$. By Kirstzbraun Theorem (see for instance [57]) there exists a Lipschitz extensions $f_+$ of $\widetilde{f}_+$ defined for all $\xi \in \mathcal{O}_0$ with the same Lipschitz norm. Moreover $f_+$ is a tame vector field with the same scale of constants as $\widetilde{f}_+$. We define $\Phi_+ = \mathbb{1} + f_+$ for $\xi \in \mathcal{O}_0$. Clearly it coincides with $\widetilde{\Phi}_+$ in the set of good parameters $\mathcal{O}_+$. Then the vector field $F_+$ is well defined. The conditions in (3.2.87) are simply smallness conditions on $\epsilon$ in (3.2.85). Conditions in (3.2.87) just say us that $s_+$ is smaller that $s$ but it is very close to it.

Let us check (iii) recalling the definitions in (3.2.89). First we note that

$$\begin{aligned} G_+ &:= (\Phi_+)_* N_0 + (\Phi_+)_* \hat{G} - N_0 = \int_0^1 (\Phi_+)^t_* [g_+, N_0] + (\Phi_+)_* \hat{G} \\ &= \int_0^1 (\Phi_+)^t_* \Pi_{\mathcal{A}} \Pi_{K_+} [g, N_0 + \hat{M} + \hat{R}] - \int_0^1 (\Phi_+)^t_* \Pi_{\mathcal{A}} \Pi_{K_+} [g, \hat{M} + \hat{R}] + (\Phi_+)_* \hat{G} \\ &= \int_0^1 (\Phi_+)^t_* (\Pi_{K_+} \hat{X} + r) - \Pi_{\mathcal{N}} \int_0^1 (\Phi_+)^t_* \Pi_{\mathcal{A}} \Pi_{K_+} [g, \hat{M} + \hat{R}] + (\Phi_+)_* \hat{G} \end{aligned} \tag{3.2.98}$$

where

$$r := \Pi_{\mathcal{A}} \Pi_{K_+} [g_+, N_0 + \hat{M} + \hat{R}] - \Pi_{K_+} \hat{X}. \tag{3.2.99}$$

Note that $r$ in (3.2.99) satisfy bound (3.2.60) with $\mathfrak{r} \rightsquigarrow \delta K^{-\kappa_1}$. By Lemma 3.2.31 we estimate the third summand in (3.2.98) and obtain

$$\begin{aligned} C_{\vec{v}_1, \mathfrak{p}_2}((\Phi_+)_* \hat{G}) &\le \gamma (1 + K_+^{\mu+2} \delta_{\mathfrak{p}_1}) [\Gamma_{\mathfrak{p}_2} + \Gamma_{\mathfrak{p}_1} K_+^{\mu+2} (\Theta_{\mathfrak{p}_2} + \delta_{\mathfrak{p}_1} (\Gamma_{\mathfrak{p}_2} + \Gamma_{\mathfrak{p}_1} \delta K^{\kappa_3}) K_+^{\alpha(\mathfrak{p}_2 - \mathfrak{p}_1)})] \\ &\le \gamma (1 + K_+^{\mu+2} \delta_{\mathfrak{p}_1}) \left[ \Gamma_{\mathfrak{p}_2} + K_+^{\mu+1} \Theta_{\mathfrak{p}_2} + \delta_{\mathfrak{p}_1} (\Gamma_{\mathfrak{p}_2} K_+^{\mu + \alpha(\kappa_2 - \mu) + 1} + K_+^{\mu + \frac{2}{3}\kappa_3 + 1 + \alpha(\kappa_2 - \mu)}) \right] \end{aligned} \tag{3.2.100}$$

and

$$C_{\vec{v}_1, \mathfrak{p}_1}(\Phi_*(\hat{G})) \overset{(3.2.97),(3.2.65)}{<} \gamma (1 + K_+^{\mu+2} \delta_{\mathfrak{p}_1}) \Gamma_{\mathfrak{p}_1} \tag{3.2.101}$$

for $\epsilon$ in (3.2.85) small enough. Moreover

$$\begin{aligned} C_{\vec{v}_1, \mathfrak{p}_2} \left( \int_0^1 \Phi^t_* (\Pi_{K_+} \hat{X} + r) \right) &\le (1 + K_+^{\mu} \delta_{\mathfrak{p}_1}) (C_{\vec{v}_1, \mathfrak{p}_2}(\hat{X}) + C_{\mathfrak{p}_1}(\hat{X}) C_{\vec{v}_1, \mathfrak{p}_2}(f_+)) \\ &\quad + (1 + K_+^{\mu} \delta_{\mathfrak{p}_1}) (C_{\vec{v}_1, \mathfrak{p}_2}(r) + C_{\mathfrak{p}_1}(r) C_{\vec{v}_1, \mathfrak{p}_2}(f_+)) \\ &\overset{(3.2.80)}{\le} (1 + K_+^{\mu} \delta_{\mathfrak{p}_1}) \left[ C_{\mathfrak{p}_2}(\hat{X}) + C_{\mathfrak{p}_1}(\hat{X}) (C_{\vec{v}_1, \mathfrak{p}_2}(\hat{G}) + 3 C_{\vec{v}_1, \mathfrak{p}_1}(\hat{G}) C_{\vec{v}_1, \mathfrak{p}_2}(f_+)) \right] \end{aligned} \tag{3.2.102}$$

where we used the fact that $\kappa_1 > (3/2)\mu + 1$ and $\delta \le 1$ and $\Gamma_{\mathfrak{p}_1} \le K_0$. Now by using (3.2.65) and (3.2.97) we get

$$\begin{aligned} C_{\vec{v}_1, \mathfrak{p}_2} \left( \int_0^1 \Phi^t_* (\Pi_{K_+} \hat{X} + r) \right) &\le \gamma (1 + K_+^{\mu+2} \delta_{\mathfrak{p}_1}) \left[ \Theta_{\mathfrak{p}_2} + \delta_{\mathfrak{p}_1} \left( \delta K^{\kappa_3 + 2} + \Gamma_{\mathfrak{p}_2} + \Theta_{\mathfrak{p}_2} K_+^{\mu+2} \right) \right] \\ &\quad + \gamma (1 + K_+^{\mu+2} \delta_{\mathfrak{p}_1}) \left[ \delta_{\mathfrak{p}_1}^2 K_+^{\mu + \alpha(\kappa_2 - \mu)} (\Gamma_{\mathfrak{p}_2} + \delta K^{\kappa_3 + 1}) + \delta_{\mathfrak{p}_1} K_+^{\mu + \alpha(\kappa_2 - \mu)} \Gamma_{\mathfrak{p}_2} \right] \end{aligned} \tag{3.2.103}$$

In low norm we have

$$C_{\vec{v}_1,\mathfrak{p}_1}\left(\int_0^1 \Phi_*^t(\Pi_{K_+}\hat{X}+r)\right) \le (1+K_+^{\mu+2}\delta_{\mathfrak{p}_1})C_{\vec{v}_1,\mathfrak{p}_1}(\hat{X})(1+C_{\vec{v}_1,\mathfrak{p}_1}(f_+))(1+\gamma^{-1}\delta K^{-\kappa_1}K_+^{-\mu}C_{\vec{v}_1,\mathfrak{p}_1}(\hat{G})) \overset{(3.2.65),(3.2.96)}{\le} \gamma(1+K_+^{\mu}\delta_{\mathfrak{p}_1})2\delta_{\mathfrak{p}_1}, \tag{3.2.104}$$

where we also used the fact that $\kappa_1 > (3/2)\mu+1$ and the smallness of $\epsilon$. The last term can be estimated as follows. First we note that the vector field $\Pi_{\mathcal{A}}\Pi_{K_+}[g_+,\hat{M}+\hat{R}]$ is tame, with tameness constant given by

$$\begin{aligned} C_{\vec{v}_1,p}(\Pi_{\mathcal{A}}\Pi_{K_+}[g,\hat{M}+\hat{R}]) &\overset{(P1)}{\le} K_+C_{\vec{v}_1,p-1}(\Pi_{\mathcal{A}}[g,\hat{M}+\hat{R}]) \\ &\le K_+(C_{\vec{v}_1,p+\nu}(g_+)C_{\vec{v},\mathfrak{p}_0}(\hat{G})+C_{\vec{v}_1,\mathfrak{p}_0+\nu}(g_+)C_{\vec{v},p}(\hat{G})), \end{aligned} \tag{3.2.105}$$

hence,

$$\begin{aligned} C_{\vec{v}_+,\mathfrak{p}_2}(\Pi_{\mathcal{A}}\Pi_{K_+}[g,\hat{M}+\hat{R}]) &\le \gamma K_+^{\nu}\left[C^{(2)}_{\mathfrak{p}_2,\mathcal{O}_+}(g_+)\Gamma_{\mathfrak{p}_1}+C^{(2)}_{\mathfrak{p}_1,\mathcal{O}_+}(g_+)(\Gamma_{\mathfrak{p}_2}+\Gamma_{\mathfrak{p}_1}\delta K^{\kappa_3})\right] \\ &\overset{(3.2.97)}{\le} \gamma K_+^{\mu+\nu+4}\left[\Theta_{\mathfrak{p}_2}+\delta_{\mathfrak{p}_1}(K_+^{\alpha(\kappa_2-\mu)+1}\Gamma_{\mathfrak{p}_2}+\delta K_+^{\alpha(\kappa_2-\mu)+\frac{2}{3}\kappa_3+2})\right] \end{aligned} \tag{3.2.106}$$

and

$$C_{\vec{v}_+,\mathfrak{p}_1}(\Pi_{\mathcal{A}}\Pi_{K_+}[g,\hat{M}+\hat{R}]) \le \gamma K_+^{\mu+\nu+2}\delta_{\mathfrak{p}_1} \tag{3.2.107}$$

Hence we have obtained

$$\begin{aligned} &C_{\vec{v}_1,\mathfrak{p}_2}(\int_0^1(\Phi_+)_*^t\Pi_{\mathcal{A}}\Pi_{K_+}[g,\hat{M}+\hat{R}]) \le (1+K_+^{\mu+2}\delta_{\mathfrak{p}_1})K_+^{\nu+\mu+2}(\Theta_{\mathfrak{p}_2}+\delta_{\mathfrak{p}_1}(K_+^{\alpha(\kappa_2+\mu)+1}\Gamma_{\mathfrak{p}_2}+2\delta K_+^{\alpha(\kappa_2-\mu)+\frac{2}{3}\kappa_3+2})) \\ &C_{\vec{v}_1,\mathfrak{p}_1}(\int_0^1(\Phi_+)_*^t\Pi_{\mathcal{A}}\Pi_{K_+}[g,\hat{M}+\hat{R}]) \le (1+K_+^{\mu+2}\delta_{\mathfrak{p}_1})\gamma K_+^{\mu+\nu+2}\delta_{\mathfrak{p}_1} \end{aligned} \tag{3.2.108}$$

The first of (3.2.90) follows by collecting together bounds (3.2.100), (3.2.103) and (3.2.108). Moreover bounds (3.2.101), (3.2.104) and (3.2.108) imply the first bound in (3.2.91). Let us study $\Theta_{+,p}$. By linearity one has that

$$\Pi_{\mathcal{N}^\perp}(\Phi_+)_*G_+ = \Pi_{\mathcal{N}^\perp}\Big((\Phi_+)_*N_0+(\Phi_+)_*(\Pi_{\mathcal{N}}\hat{G})+(\Phi_+)_*(\Pi_{\mathcal{N}^\perp}\hat{G})\Big) \tag{3.2.109}$$

First of all we have

$$\begin{aligned} &C_{\vec{v}_1,\mathfrak{p}_2}((\Phi_+)_*(\Pi_{\mathcal{N}^\perp}\hat{G})) \overset{(3.2.49),(3.2.65)}{\le} \gamma(1+K_+^{\mu+2}\delta_{\mathfrak{p}_1})\Big(\Theta_{\mathfrak{p}_2}(1+K_+^{\mu+2})+\delta_{\mathfrak{p}_1}K_+^{\mu+\alpha(\kappa_2-\mu)+1}(\Gamma_{\mathfrak{p}_2}+\delta K_+^{\kappa_3+1})\Big), \\ &C_{\vec{v}_1,\mathfrak{p}_1}((\Phi_+)_*(\Pi_{\mathcal{N}^\perp}G)) \le \gamma(1+K_+^{\mu+2}\delta_{\mathfrak{p}_1})\Theta_{\mathfrak{p}_1}(1+K_+^{\mu+2}\delta_{\mathfrak{p}_1}), \end{aligned} \tag{3.2.110}$$

and we can use Lemma B.181 to estimate the projection on $\mathcal{N}^\perp(\Phi_+)_*(\Pi_{\mathcal{N}^\perp} G)$. Now we use Lemma (B.185) to obtain

$$\begin{aligned} C_{\vec{v}_1,\mathfrak{p}_2}(\Pi_{\mathcal{N}^\perp}(\Phi_+)_*(\Pi_{\mathcal{N}}\hat{G})) &\le C_{\vec{v},\mathfrak{p}_2}(\hat{G}) C_{\vec{v}_2,\mathfrak{p}_1+1}(f_+) + 2C_{\vec{v},\mathfrak{p}_1}(\hat{G}) C_{\vec{v}_2,\mathfrak{p}_2+1}(f_+) \\ &\overset{(3.2.65)}{\le} \gamma K_+^{\mu+2}\left[\Theta_{\mathfrak{p}_2} + \delta_{\mathfrak{p}_1} K_+^{\alpha(\kappa_2-\mu)}(\Gamma_{\mathfrak{p}_2} + \delta K^{\kappa_3+1})\right], \end{aligned} \tag{3.2.111}$$

and

$$C_{\vec{v}_1,\mathfrak{p}_1}(\Pi_{\mathcal{N}^\perp}(\Phi_+)_*(\Pi_{\mathcal{N}}\hat{G})) \le K_+^{\mu+4}\delta_{\mathfrak{p}_1}, \tag{3.2.112}$$

Finally to estimate the term $\Pi_{\mathcal{N}^\perp}(\Phi_+)_* N_0$ one can use bounds (3.2.103), (3.2.104) and (3.2.108) which, together with (3.2.110), (3.2.111) and (3.2.112) imply the estimates on $\Theta_{+,\mathfrak{p}_2}, \Theta_{+,\mathfrak{p}_1}$ in (3.2.90) and (3.2.91) for $\epsilon$ small enough.

In order to prove the estimate in low norm we first write

$$\Pi_{\mathcal{A}} F_+ = \Pi_{\mathcal{A}}(\hat{F} + [\hat{F}, g_+] + Q), \qquad Q := (\Phi_+)_*\hat{F} - (\hat{F} + [\hat{F}, g_+]) \tag{3.2.113}$$

and hence

$$\begin{aligned} \Pi_{\mathcal{A}} F_+ &= \hat{X} + \Pi_{\mathcal{A}}[\hat{N} + \hat{R}, g_+] + \Pi_{\mathcal{A}}[\hat{X}, g_+] + \Pi_{\mathcal{A}} Q \\ &= \Pi_{K_+}\Big(\hat{X} + \Pi_{\mathcal{A}}[\hat{N} + \hat{R}, g_+]\Big) + \Pi_{K_+}^\perp\Big(\hat{X} + \Pi_{\mathcal{A}}[\hat{N} + \hat{R}, g_+]\Big) + \Pi_{\mathcal{A}}[\hat{X}, g_+] + \Pi_{\mathcal{A}} Q. \end{aligned} \tag{3.2.114}$$

Consider the tern $\Pi_{\mathcal{A}} Q$. Using Lemma B.180 and B.184 one can reason as in (3.2.98) and write

$$\Pi_{\mathcal{A}} Q = \Pi_{\mathcal{A}}\left[\int_0^1\int_0^t (\Phi_+)_*^s \Big([g_+, [g_+, \hat{G}]] + [g_+, \Pi_{K_+}\hat{X} + r - \Pi_{\mathcal{A}}\Pi_{K_+}[g_+, \hat{M} + \hat{R}]]\Big)\right] \tag{3.2.115}$$

where $r$ is defined in (3.2.99). On the first summand in (3.2.115) one applies Lemma B.183 and obtains

$$\begin{aligned} C_{\vec{v}_1,\mathfrak{p}_1}\left(\Pi_{\mathcal{A}}\int_0^1\int_0^t (\Phi_+)_*^s\Big([g_+, [g_+, \hat{G}]]\Big)\right) &\le C_{\vec{v}_1,\mathfrak{p}_1+2}(\hat{G}) C^2_{\vec{v}_1,\mathfrak{p}_1}(g_+) \\ &\le C^2_{\vec{v}_1,\mathfrak{p}_1}(g_+)\Big(C_{\vec{v}_1,\mathfrak{p}_1+2}(\Pi_{K_+}\hat{G}) + C_{\vec{v}_1,\mathfrak{p}_1+2}(\Pi_{K_+}^\perp\hat{G})\Big) \\ &\overset{(3.2.65),(3.2.97)}{\le} \gamma\delta^2_{\mathfrak{p}_1} K_+^{2\mu+7}\left[1 + (\Gamma_{\mathfrak{p}_2} + \delta K^{\kappa_3+1}) K_+^{-(\kappa_2-\mu)}\right], \end{aligned} \tag{3.2.116}$$

Now using Lemmata 3.2.31, C.186 and the bounds (3.2.104), (3.2.108) we obtain

$$C_{\vec{v}_1,\mathfrak{p}_1}(\int_0^1\int_0^t (\Phi_+)_*^s[g_+, \Pi_{K_+}\hat{X} + r - \Pi_{\mathcal{A}}\Pi_{K_+}[g_+, \hat{M} + \hat{R}]]) \le \gamma(1 + K_+^{\mu+2}\delta_{\mathfrak{p}_1}) K_+^{2\mu+2\nu+6}\delta^2_{\mathfrak{p}_1}. \tag{3.2.117}$$

For $\epsilon$ small enough one has that

$$C_{\vec{v}_+,\mathfrak{p}_1}(\Pi_{\mathcal{A}} Q) \le \gamma\delta^2_{\mathfrak{p}_1} K_+^{2\mu+2\nu+8}\left[1 + (\Gamma_{\mathfrak{p}_2} + \delta K^{\kappa_3+1}) K_+^{-(\kappa_2-\mu)}\right]. \tag{3.2.118}$$

Moreover one has

$$C_{\vec{v}_1,\mathfrak{p}_1}(\Pi_K^{\perp}\Big(\hat{X}+\Pi_{\mathcal{A}}[\hat{N}+\hat{R},g_+]\Big)) \stackrel{(P2)}{\le} K_+^{-(\mathfrak{p}_2-\mathfrak{p}_1-1)}C_{\vec{v}_1,\mathfrak{p}_2-1}(\hat{X}+\Pi_{\mathcal{A}}[\hat{N}+\hat{R},g_+]) \stackrel{(3.2.65),(B.33)}{\le} \gamma K_+^{-(\kappa_2-\mu-1)}\Big(\Theta_{\mathfrak{p}_2}K_+^{\mu+2}+\delta_{\mathfrak{p}_1}K_+^{\mu+2+\alpha(\kappa_2-\mu)}(\Gamma_{\mathfrak{p}_2}+\delta K^{\kappa_3+1})\Big) \tag{3.2.119}$$

Finally, using the condition $(c)$ of Definition 3.2.37 with $\mathfrak{r}=\varepsilon K^{-\kappa_1}$ one obtain

$$C_{\vec{v}_+,\mathfrak{p}_1}\left(\Pi_{K_+}\Big(\hat{X}+\Pi_{\mathcal{A}}[\hat{N}+\hat{R},g_+]\Big)\right)\le \gamma K_+^{\mu+2-\frac{2}{3}\kappa_1}\delta_{\mathfrak{p}_1}. \tag{3.2.120}$$

In conclusion one has

$$C_{\vec{v}_1,\mathfrak{p}_1}(\Pi_{\mathcal{A}}F_+)\le \gamma K_+^{-(1-\alpha)(\kappa_2-\mu)+\mu+2}(\Gamma_{\mathfrak{p}_2}+\delta K^{\kappa_3+1})\delta_{\mathfrak{p}_1}+K_+^{-\kappa_2+\mu+3}\Theta_{\mathfrak{p}_2} + \gamma K_+^{\mu+2-\frac{2}{3}\kappa_1}\delta_{\mathfrak{p}_1}+\gamma\delta_{\mathfrak{p}_1}^2K_+^{2\mu+2\nu+8}\Big[1+(\Gamma_{\mathfrak{p}_2}+\delta K^{\kappa_3+1})K_+^{-(\kappa_2-\mu)}\Big] \tag{3.2.121}$$

that is bound (3.2.92). □

In order to prove Theorem 3.2.39 we use the following iterative lemma:

**Proposition 3.2.45 (Iterative Lemma).** *Consider a tame vector field $F$ as in Theorem 3.2.39. Then there exist constants $\delta>0, C_\star>0$, $K_0\in\mathbb{N}$ , $C_0(d,\mathfrak{p}_0)$ and $C_1(d,\mathfrak{p}_0)$, $\epsilon_0=\epsilon_0(\mathfrak{p}_0,d)$ such that, denoting*

$$\Gamma_{0,p}:=\gamma_0^{-1}C_{\vec{v}_0,p}(G)\,,\quad \Theta_{0,p}:=\gamma_0^{-1}C_{\vec{v}_0,p}(\Pi_{\mathcal{N}^\perp}G),\quad \delta_{0,p}:=\gamma_0^{-1}C_{\vec{v}_0,p}(\Pi_{\mathcal{A}}G), \tag{3.2.122}$$

*with $\vec{v}_0:=(\lambda,\mathcal{O}_0,s_0,a_0)$, and assuming*

$$\Gamma_{0,\mathfrak{p}_2}\le\mathfrak{A}_0,\quad \delta_{0,\mathfrak{p}_1}\le\delta\quad \Theta_{0,\mathfrak{p}_2}=\delta\mathfrak{A}_0,\quad C_\star K_0^{C_0}\delta_{0,\mathfrak{p}_1}\le\epsilon_0,\ \ \mathfrak{A}_0K_0^{-C_1}\le\epsilon_0 \tag{3.2.123}$$

*if $\epsilon_0$ is small enough then the following holds:*

$(\mathbf{N1})_n$ *For $0\le j\le n-1$ we have the vector fields $F_j:\mathcal{O}_0\times D_{a_j,p+\nu}(s_j,r_j)\to V_{a_j,p}$ and the sets $\mathcal{O}_j\subseteq\mathcal{O}_0$ which satisfy the hypotheses of Lemma 3.2.44 with the parameters $\rho_{j+1},K_j$ and $\gamma_j$. To define $F_n$ we apply Lemma 3.2.44 to $F_{n-1}$. In particular this implies fixing $\mathcal{O}_{j+1}\in\mathcal{O}_j\cap J_{K_{j+1},\mathfrak{r}_j,\gamma_j}(\hat{F}_j)$ with $\mathfrak{r}_j$*

$(\mathbf{N2})_n$ *for $n\ge 0$ setting for $\vec{v}_n:=(\gamma_n,\mathcal{O}_n,a_n,s_n)$*

$$\Gamma_{n,p}:=\gamma_n^{-1}C_{\vec{v}_n,p}(G_n)\,,\quad \Theta_{n,p}:=\gamma_n^{-1}C_{\vec{v}_n,p}(\Pi_{\mathcal{N}^\perp}G_n),\quad \delta_{n,p}:=\gamma_n^{-1}C_{\vec{v}_n,p}(\Pi_{\mathcal{A}}G_n), \tag{3.2.124}$$

*one has the following estimate in high norm*

$$\Gamma_{n,\mathfrak{p}_2}\le\mathfrak{A}_0K_n^{\kappa_1},\qquad \Theta_{n,\mathfrak{p}_2}\le\delta\mathfrak{A}_0K_n^{\kappa_1}, \tag{3.2.125}$$

*and in low norm*

$$\Gamma_{n,\mathfrak{p}_1} \le \mathfrak{A}_n, \qquad \Theta_{n,\mathfrak{p}_1} \le \delta\mathfrak{A}_n, \qquad \delta_{n,\mathfrak{p}_1} \le C_\star \delta K_n^{-\kappa_1}\,. \tag{3.2.126}$$

$(\mathbf{N3})_n$ *for* $n \ge 0$ *one has*

$$\mathcal{K}_n = \Phi_n \circ \mathcal{T}_n : D_{a'_n,p}(s'_n, r'_n) \to D_{a_{n-1}-2\rho_n a_0,p}(s_{n-1} - 2\rho_n s_0, r_{n-1} - 2\rho_n r_0) \tag{3.2.127}$$

*with* $s'_n = s_{n-1} - 5\rho_n s_0$, $a'_n = a_{n-1} - 5\rho_n a_0$, $r'_n = r_{n-1} - 5\rho_n r_0$ *is* $\mathcal{E}-$ *preserving and we define*

$$\mathcal{H}_n := \mathcal{K}_n \circ \mathcal{K}_{n-1} \circ \ldots \mathcal{K}_1 \circ \mathcal{K}_0. \tag{3.2.128}$$

*For* $\vec{v}\,' = (\gamma_n, \mathcal{O}_n, s'_n, a'_n)$ *and* $u \in D_{a'_n,p}(s'_n, r'_n)$ *one has that*

$$\|\mathcal{H}_n - \mathbb{1}\|_{\vec{v}\,',\mathfrak{p}_1} \lessdot \gamma_0^{-1}\varepsilon_0 \sum_{k=0}^{n} \frac{1}{2^k}, \quad \|\mathcal{H}_n - \mathcal{H}_{n-1}\|_{\vec{v}\,',\mathfrak{p}_1} \le \gamma_0^{-1}\varepsilon_0 K_n^{-1}\,. \tag{3.2.129}$$

*Same bounds hold for the inverse.*

*Proof.* We prove the result by induction on $n$.
$(\mathbf{N1})_0$. The field $F_0$ satisfies the Hypotheses of Lemma 3.2.44. Indeed $\mathcal{O} = \mathcal{O}_0$, $F \in \mathcal{E}$ for any $\xi \in \mathcal{O}_0$ and equations (3.2.81) and (3.2.82) are trivially true. Moreover the first two conditions of (3.2.85) are trivial, while

$$\rho_1^{-1} K_1^{\mu+\nu+6} \delta_{0,\mathfrak{p}_1} = 2^{-9} K_0^{\frac{3}{2}(\mu+\nu+6)} \delta_{0,\mathfrak{p}_1} \le \epsilon_0, \tag{3.2.130}$$

by (3.2.123) for suitable $C(\mathfrak{p}_0, d)$. Hence we can apply Lemma 3.2.44 to $F_0$ and define $F_1$ and $\mathcal{O}_1$ as the output $F_+$ and $\mathcal{O}_+$ of the lemma.
$(\mathbf{N2})_0$. Equations (3.2.125) and (3.2.126) follows simply by (3.2.123). The bounds on $\delta_{0,\mathfrak{p}_1}$ in (3.2.126) follows by choosing $C_\star \ge K_0^{\kappa_1}$.
$(\mathbf{N3})_0$. This is trivial because $\mathcal{T}_0 = \mathbb{1}$ and all the bounds hold with $\Phi_0 = \mathbb{1}$.
Let us assume inductively $(\mathbf{Ni})_m$ for $i = 1, 2, 3, 4$ for all $m \le n$ and prove $(\mathbf{Ni})_{n+1}$ for $i = 1, 2, 3$.
$(\mathbf{N1})_{n+1}$. By inductive hypothesis we have defined the field $F_n$ as as the output of Lemma 3.2.44 applied to $F_{n-1}$. We need to check that also $F_n$ satisfies the hypotheses of Lemma (3.2.44). Clearly $K_n > K_0$ and $\mathcal{O}_n \subseteq \mathcal{O}_0$. Clearly by (3.2.61) $\gamma_0 \le \gamma_n \le 2\gamma_0$, moreover

$$r_0 \ge r_n = r_0 \left(1 - \frac{1}{2}\sum_{j=1}^{n} \frac{1}{2^j}\right) \ge r_0 \frac{1}{2},$$

and same for $a_n$ and $s_n$, hence the (3.2.81) holds true. Moreover one has that $\rho_n + 1 := 2^{-(n+8)}$ satisfies the (3.2.82). If $\xi \in \mathcal{O}_n$ then $\xi \in \cap_{i=1}^n \mathcal{O}_i$ and hence thanks to $(\mathbf{N1})_n$ we have that $F_j \in \mathcal{E}$ for $\xi \in \mathcal{O}_n$ and

$0 \le j \le n-1$. We conclude that $F_n \in \mathcal{E}$ on $\mathcal{O}_n$. It remains to prove (3.2.85) for $F_n$. We obtain the first two bounds in (3.2.85) using $(\mathbf{N2})_n$, and moreover

$$\rho_{n+1}^{-1} K_{n+1}^{\mu+\nu+6} \delta_{n,\mathfrak{p}_1} \overset{(N2)_n}{\le} 2^{-(n+9)} K_{n+1}^{-\frac{2}{3}\kappa_1+\mu+\nu+6} C_\star \delta, \tag{3.2.131}$$

hence by (3.2.61) one can choose $K_0$ large in order to fulfill (3.2.85).

$(\mathbf{N2})_{n+1}$. To prove this point we use item *(iii)* of Lemma 3.2.44. Let us check the (3.2.125). First one has

$$\begin{aligned}\Gamma_{n+1,\mathfrak{p}_2} &\overset{(3.2.90)}{\le} (1+K_{n+1}^{\mu+2}\delta C_\star K_n^{-\kappa_1}) \Big[\mathfrak{A}_0 K_n^{\kappa_1} + K_{n+1}^{\mu+\nu+2+\frac{2}{3}\kappa_1}\delta\mathfrak{A}_0 + \\ &\quad + \delta C_\star K_{n+1}^{2\mu+\nu+4+\alpha(\kappa_2-\mu)}\mathfrak{A}_0 + \delta C_\star K_{n+1}^{-\frac{2}{3}\kappa_1+\frac{2}{3}\kappa_3+\frac{2}{3}+2\mu+\nu+4+\alpha(\kappa_2-\mu)}\Big], \\ &\overset{(3.2.62)}{\le} \mathfrak{A}_0 K_{n+1}^{\kappa_1},\end{aligned} \tag{3.2.132}$$

and moreover

$$\begin{aligned}\Theta_{n+1,\mathfrak{p}_2} &\le (1+K_{n+1}^{\mu+2}C_\star\delta K_n^{-\kappa_1}) \Big[\delta\mathfrak{A}_0 K_{n+1}^{\frac{2}{3}\kappa_1+\mu+\nu+4} + \delta C_\star \mathfrak{A}_0 K_{n+1}^{\mu+\nu+4+\alpha(\kappa_2-\mu)} + \\ &\quad + \delta C_\star K_{n+1}^{-\frac{2}{3}\kappa_1+\mu+\nu+4+\alpha(\kappa_2-\mu)+\frac{2}{3}\kappa_3+\frac{2}{3}}\Big] \\ &\overset{(3.2.62)}{\le} \delta\mathfrak{A}_0 K_{n+1}^{\kappa_1}.\end{aligned} \tag{3.2.133}$$

Let us now check the estimates in low norm. We have

$$\Gamma_{n+1,\mathfrak{p}_1} \le (1+K_{n+1}^{\mu+\nu+6}C_\star\delta K_n^{-\kappa_1})\Gamma_{n,\mathfrak{p}_1} \overset{(N2)_n}{\le} \mathfrak{A}_n + \mathfrak{A}_0 K_{n+1}^{-1} \le \mathfrak{A}_{n+1} \tag{3.2.134}$$

for $K_0$ large enough. Moreover we obtain

$$\Theta_{n+1,\mathfrak{p}_1} \overset{(3.2.90)}{\le} (1+K_{n+1}^{\mu+\nu+6}C_\star\delta K_n^{-\kappa_1})\Theta_{n,\mathfrak{p}_1} \overset{(N2)_n}{\le} \delta(\mathfrak{A}_n + \mathfrak{A}_0 K_{n+1}^{\mu+\nu+6} C_\star K_n^{-\kappa_1}) \le \delta\mathfrak{A}_{n+1}. \tag{3.2.135}$$

Finally

$$\begin{aligned}\delta_{n+1,\mathfrak{p}_1} &\overset{(3.2.62)}{\le} K_{n+1}^{-(1-\alpha)(\kappa_2-\mu)+\mu+2-\frac{2}{3}\kappa_1}\delta C_\star(\mathfrak{A}_0 K_n^{\kappa_1} + \delta K_n^{\kappa_3+1}) + K_{n+1}^{-\kappa_2+\mu+3+\frac{2}{3}\kappa_1}\delta\mathfrak{A}_0 \\ &\quad + K_{n+1}^{\mu+2-\frac{4}{3}\kappa_1}\delta C_\star + \delta^2 C_\star^2 K_+^{-\frac{4}{3}\kappa_1+2\mu+2\nu+8}\Big[1 + (\mathfrak{A}_0 K_n^{\kappa_1} + \delta K^{\kappa_3+1})K_{n+1}^{-(\kappa_2-\mu)}\Big] \\ &\overset{(3.2.62)}{\le} C_\star\delta K_{n+1}^{-\kappa_1}\end{aligned} \tag{3.2.136}$$

$(\mathbf{N3})_{n+1}$. One has that for $\vec{v}\,' = (\gamma_{n+1}, \mathcal{O}_{n+1}, s_n - 5\rho_{n+1}s_0, a_n - 5\rho_{n+1}a_0)$,

$$\begin{aligned}
\|(\mathcal{H}_{n+1} - \mathbb{1})(u)\|_{\vec{v}\,',\mathfrak{p}_1} &\le \|(\mathcal{H}_n - \mathbb{1})(u)\|_{\vec{v}\,',\mathfrak{p}_1} + \|f_{n+1}(\mathcal{T}_{n+1}\circ\mathcal{H}_n(u))\|_{\vec{v}\,',\mathfrak{p}_1} + \|(\mathcal{T}_{n+1} - \mathbb{1})\mathcal{H}_n(u)\|_{\vec{v}\,',\mathfrak{p}_1} \\
&\overset{(N3)_n,(3.2.66)}{\le} \gamma_0^{-1}\varepsilon_0 \sum_{k=0}^{n} \frac{1}{2^k} + \gamma_n^{-1}\varepsilon_n K_n^{-\kappa_3}\|\mathcal{H}_n(u)\|_{\vec{v}\,',\mathfrak{p}_1} + 2C_{\vec{v}\,',\mathfrak{p}_1}(g_{n+1})\|\mathcal{T}_{n+1}\circ\mathcal{H}_n(u)\|_{\vec{v}\,',\mathfrak{p}_1} \\
&\overset{(??)}{\le} \gamma_0^{-1}\varepsilon_0 \sum_{k=0}^{n} \frac{1}{2^k} + (1 + 8\gamma_0^{-1}\varepsilon_0)(2K_n^{-\kappa_3} + K_n^{\mu-2/3\kappa_1})\gamma_n^{-1}\varepsilon_n \\
&\le \gamma_0^{-1}\varepsilon_0 \sum_{k=0}^{n} \frac{1}{2^k} + \frac{1}{2^{n+1}}
\end{aligned} \tag{3.2.137}$$

for $K_0$ large and $\gamma_0^{-1}\varepsilon_0$ small enough. Finally

$$\begin{aligned}
\|(\mathcal{H}_{n+1} - \mathcal{H}_n)u\|_{\vec{v}\,',\mathfrak{p}_1} &\le \|(\mathcal{T}_{n+1} - \mathbb{1})\mathcal{H}_n(u)\|_{\vec{v}\,',\mathfrak{p}_1} + \|f_{n+1}(\mathcal{T}_{n+1}\circ\mathcal{H}_n)(u)\|_{\vec{v}\,',\mathfrak{p}_1} \\
&\overset{(\mathbf{N4})_n,(3.2.66)}{\lessdot} \mathcal{A}_0 C_\star \gamma_n^{-1}\varepsilon_n (K_n^{-\kappa_3} + K_n^{\mu-2/3\kappa_1}),
\end{aligned} \tag{3.2.138}$$

implies the second of (3.2.129) for for $K_0$ large and $\gamma_0^{-1}\varepsilon_0$ small enough. This concludes the proof of Proposition 3.2.45. □

**Proof of Theorem 3.2.39** We first show that there exists a limit map $\bar{\mathcal{H}} = \lim_{n\to\infty}\mathcal{H}_n$. This simply follows by (3.2.129). Indeed

$$\begin{aligned}
\|(\bar{\mathcal{H}} - \mathbb{1})(u)\|_{\gamma_\infty,\mathcal{O}_\infty,\frac{s_0}{2},\frac{a_0}{2},\mathfrak{p}_1} &\le \|(\mathcal{H}_1 - \mathbb{1})(u)\|_{\gamma_\infty,\mathcal{O}_\infty,\frac{s_0}{2},\frac{a_0}{2},\mathfrak{p}_1} + \sum_{n\ge 2}\|(\mathcal{H}_n - \mathcal{H}_{n-1})(u)\|_{\gamma_\infty,\mathcal{O}_\infty,\frac{s_0}{2},\frac{a_0}{2},\mathfrak{p}_1} \\
&\le 2\gamma_0^{-1}\varepsilon_0.
\end{aligned} \tag{3.2.139}$$

Hence for $\varepsilon_0$ small enough also the (3.2.71) hold. Bound (3.2.73) follow by (3.2.97). It remains to check (3.2.75). By definition the final vector field is

$$F_\infty := \lim_{n\to\infty} F_n, \quad F_n := (\Phi_n \circ \mathcal{T}_n)_* F_{n-1}. \tag{3.2.140}$$

On the other hand we have

$$\mathcal{F}_\infty := (\overline{\mathcal{H}})_* F_0$$

We need that $\mathcal{F}_\infty = F_\infty$. Setting $\mathcal{F}_n := (\mathcal{H}_n)_* F_0$ we show that $\mathcal{F}_n = F_n$ for any $n \ge 1$. We proceed by induction. For $n = 1$ clearly $\mathcal{H}_1 = \Phi_1 \circ \mathcal{T}_1$, hence $F_1 = \mathcal{F}_1$. Now assume that $\mathcal{F}_{n-1} = F_{n-1}$. By the $(\mathbf{N4})_n$ one has that

$$\mathcal{H}_n = \mathcal{K}_n \circ \mathcal{H}_{n-1} = (\Phi_n \circ \mathcal{T}_n) \circ \mathcal{H}_{n-1}, \qquad \mathcal{H}_n^{-1} = \mathcal{H}_{n-1}^{-1} \circ \mathcal{K}_n^{-1}.$$

Hence we have

$$\begin{aligned} F_n - \mathcal{F}_n &= (\mathcal{K}_n)_* F_{n-1} - (\mathcal{H}_n)_* F_0 = d\mathcal{K}_n F_{n-1} \circ \mathcal{K}_n^{-1} - d\mathcal{H}_n F_0 \circ \mathcal{H}_n^{-1} \\ &= d\mathcal{K}_n F_{n-1} \circ \mathcal{K}_n^{-1} - d\mathcal{K}_n d\mathcal{H}_{n-1} F_0 \circ \mathcal{H}_{n-1}^{-1} \circ \mathcal{K}_n^{-1} \\ &= d\mathcal{K}_n (F_{n-1} - d\mathcal{H}_{n-1} F_0 \circ \mathcal{H}_{n-1}^{-1}) \circ \mathcal{K}_n^{-1} = 0. \end{aligned} \tag{3.2.141}$$

# 4. Forced NLS: reversible case

In this Chapter we give the proof of Theorem 1.1.1. In order to make clear ous strategy we divided the proof into several technical Propositions. The general strategy is essentially the same in both cases and it is the one followed in [5] e [31]. It is based on a Nash-Moser iteration. We consider the operator $F$ in 2.1.14 for the reversible case and our aim is to show that there exists a sequence of functions that converges, in some Sobolev space, to a *solution* of (1.1.3). As explained in the Introduction the first we prove an (essentially standard) Nash-Moser iteration scheme which produces a Cauchy sequence of functions converging to a *solution* on a possibly empty Cantor like set.

**Proposition 4.0.46 (Nash-Moser).** *Fix $\gamma \leq \gamma_0, \mu > \tau > d$. There exist $q \in \mathbb{N}$, depending only on $\tau, d, \mu$, such that for any nonlinearity $\mathtt{f} \in C^q$ satisfying Hypothesis 1 the following holds. Consider $F(\mathbf{u})$ defined in Definition 2.1.8, (or $\mathcal{F}$ defined in* (2.1.11)*), then there exists a small constant $\epsilon_0 > 0$ such that for any $\varepsilon$ with $0 < \varepsilon\gamma^{-1} < \epsilon_0$, there exist constants $C_\star, N_0 \in \mathbb{N}$, a sequence of functions $\mathbf{u}_n$ and a sequence of sets $\mathcal{G}_n(\gamma, \tau, \mu) \equiv \mathcal{G}_n \subseteq \Lambda$ such that $\mathbf{u}_n : \mathcal{G}_n \to \mathbf{X}^0$ (or $\mathbf{u}_n : \mathcal{G}_n \to \mathrm{H}^0$),*

$$||\mathbf{u}_n||_{\mathfrak{s}_0+\mu,\gamma} \leq 1, \quad ||\mathbf{u}_n - \mathbf{u}_{n-1}||_{\mathfrak{s}_0+\mu,\gamma} \leq C_\star \varepsilon \gamma^{-1} (N_0)^{-\left(\frac{3}{2}\right)^n (18+2\mu)}. \tag{4.0.1}$$

*Here $||\cdot||_{s,\gamma}$ is the weighted Lipschitz norm in* (2.1.6)*. Moreover the sequence converges in $||\cdot||_{\mathfrak{s}_0+\mu,\gamma}$ to a function $\mathbf{u}_\infty$ such that*

$$F(\mathbf{u}_\infty) = 0, \quad \forall\, \lambda \in \mathcal{G}_\infty := \cap_{n\geq 0} \mathcal{G}_n. \tag{4.0.2}$$

The proof of Proposition 4.0.46 of course uses the abstract result we obtained in Section 3.1. We state the result also for the Hamiltonian equation because essentially there are no important differences. In Section 5.4 we will show that also the functional $\mathcal{F}$ in (2.1.11) satisfies the hypotheses of Lemma 3.1.18, hence Proposition 4.0.46 holds.

## 4.1 Reversible Operators

We now specify to $\mathcal{H}_s = \mathbf{H}^s := H^s(\mathbb{T}^{d+1}, \mathbb{C}) \times H^s(\mathbb{T}^{d+1}, \mathbb{C}) \cap \mathcal{U}$, with the notations of Definition 2.2.9; recall that

$$||\mathbf{h}||_{s,\gamma} := ||\mathbf{h}||_{H^s \times H^s,\gamma} = \max\{||h^+||_{s,\gamma}, ||h^-||_{s,\gamma}\}. \tag{4.1.3}$$

Since we are working one the space of functions which are odd in space, it is more convenient to use the sine basis in space instead of the exponential one. Namely for $u$ odd in space we have the two equivalent representations:

$$u(\varphi, x) = \sum_{\ell \in \mathbb{Z}^d, j \in \mathbb{Z}} u_j(\ell) e^{i(\ell \cdot \varphi + jx)} = \sum_{\ell \in \mathbb{T}^d, j \in \mathbb{N}} \widetilde{u}_j(\ell) e^{i\ell \cdot \varphi} \sin jx,$$

setting $\widetilde{u}_j(\ell) = 2iu_j(\ell)$, since $u_j = -u_{-j}$. Then we have also two equivalent $H^s$ norms differing by a factor 2. In the following we will use the second one which we denote by $\|\cdot\|_s$, because it is more suitable to deal with odd functions and odd operators. The same remark holds also for even functions, in that case we will use the cosine basis of $L^2_x$.

We will also use this notation. From a dynamical point of view our solution $\mathbf{u}(\varphi, x) \in \mathbf{H}^s(\mathbb{T}^d \times \mathbb{T})$ can be seen as a map

$$\mathbb{T}^d \ni \varphi \to h(\varphi) := \mathbf{u}(\varphi, x) \in \mathbf{H}^s_x := H^s_x(\mathbb{T}) \times H^s_x(\mathbb{T}) \cap \mathcal{U}. \tag{4.1.4}$$

In other words we look for a curve in the phase space $\mathbf{H}^s_x$ that solves (2.1.14). We will denote the norm of $h(\varphi) := (u(\varphi, x), \bar{u}(\varphi, x))$

$$||h(\varphi)||^2_{\mathbf{H}^s_x} := \sum_{j \in \mathbb{Z}} |u_j(\varphi)|^2 \langle j \rangle^{2s}. \tag{4.1.5}$$

It can be interpreted as the norm of the function at time a certain time $t$, with $\omega t \leftrightarrow \varphi$. The same notation is used also if the function $u$ belongs to some subspaces of even or odd functions in $H^s_x$.

Let $a_{i,j} \in H^s(\mathbb{T}^d \times \mathbb{T})$, on the multiplication operator $A = (a_{i,j})_{i,j=\pm 1} : \mathcal{H}_s \to \mathcal{H}_s$, we define the norm

$$||A||_s := \max_{i,j=\pm 1}\{||a_{i,j}||_s\}, \quad ||A||_{s,\gamma} := \max_{i,j=\pm 1}\{||a_{i,j}||_{s,\gamma}\} \tag{4.1.6}$$

Recalling the definitions (2.2.23), we set,

**Definition 4.1.47.** *An operator $R : H^s \to H^s$ is "*reversible*" with respect to the reversibility (2.2.20) if*

$$R : X^s \to Z^s, \quad s \geq 0 \tag{4.1.7}$$

*We say that $R$ is "*reversibility-preserving*" if*

$$R : G^s \to G^s, \quad \text{for} \quad G^s = X^s, Y^s, Z^s, \quad s \geq 0. \tag{4.1.8}$$

*In the same way, we say that $A : \mathbf{X}^s \to \mathbf{Z}^s$, for $s \geq 0$ is "reversible", while $A : \mathbf{G}^s \to \mathbf{G}^s$, for $\mathbf{G}^s = \mathbf{X}^s, \mathbf{Y}^s, \mathbf{Z}^s$, $s \geq 0$ is "reversibility-preserving".*

**Remark 4.1.48.** *Note that, since $\mathbf{X}^s = X^s \times X^s \cap \mathcal{U}$, Definition 4.1.47 guarantees that a reversible operator preserves also the subspace $\mathcal{U}$, namely $(u, \bar{u}) \overset{R}{\to} (z, \bar{z}) \in H^s \times H^s \cap \mathcal{U}$.*

**Lemma 4.1.49.** *Consider operators $A, B, C$ of the form*

$$A := \begin{pmatrix} a_1^1(\varphi, x) & a_1^{-1}(\varphi, x) \\ a_{-1}^1(\varphi, x) & a_{-1}^{-1}(\varphi, x) \end{pmatrix}, \ B := i \begin{pmatrix} a_1^1(\varphi, x) & a_1^{-1}(\varphi, x) \\ -a_{-1}^1(\varphi, x) & -a_{-1}^{-1}(\varphi, x) \end{pmatrix}, \ C := B\partial_x.$$

*One has that $A$ is reversibility-preserving if and only if $a_\sigma^{\sigma'} \in Y^s$ for $\sigma, \sigma' = \pm 1$. Moreover $B$ is reversible if and only if $A$ is reversibility-preserving. Finally $C$ is reversible if and only if $a_\sigma^{\sigma'} \in X^s$.*

*Proof.* The Lemma is proved by simply noting that for $u \in X^s$

$$a_\sigma^{\sigma'} \cdot u \in X^s, \ \ i\sigma a_\sigma^{\sigma'} \cdot u \in Z^s, \ \ \forall\, a_\sigma^{\sigma'} \in Y^s, \ ia_\sigma^{\sigma'} \cdot u_x \in Z^s, \ \forall\, a_\sigma^{\sigma'} \in X^s, \tag{4.1.9}$$

using that $u_x \in Y^s$ if $u \in X^s$. The fact that the subspace $\mathcal{U}$ is preserved, follows by the hypothesis that $a_\sigma^{\sigma'} = \overline{a_{\sigma'}^{\sigma}}$, that guarantees, for instance $R\mathbf{u} = (z_1, z_2)$ with $z_1 = \overline{z_2}$. □

### 4.1.1 Proof of Proposition 4.0.46

We now prove that our equation (2.1.14) satisfies the hypotheses of the abstract Nash-Moser theorem. We fix $\nu = 2$ and consider the operator $F : \mathbf{H}^s \to \mathbf{H}^{s-2}$,

$$F(\lambda, \mathbf{u}) := \begin{pmatrix} \lambda\bar{\omega} \cdot \partial_\varphi u + i\partial_{xx} u \\ \lambda\bar{\omega} \cdot \partial_\varphi \bar{u} - i\partial_{xx}\bar{u} \end{pmatrix} + \varepsilon \begin{pmatrix} if_1(\varphi, x, u, \bar{u}, u_x, \bar{u}_x, u_{xx}, \bar{u}_{xx}) \\ -if_2(\varphi, x, \bar{u}, u, \bar{u}_x, u_x, \bar{u}_{xx}, u_{xx}) \end{pmatrix} \tag{4.1.10}$$

For simplicity we write

$$F(\mathbf{u}) := F(\lambda, \mathbf{u}) = L_\omega \mathbf{u} + \varepsilon f(\mathbf{u}) \tag{4.1.11}$$

where (recall $\omega = \lambda\bar{\omega}$)

$$\begin{aligned} L_\lambda \equiv L_\omega \ &:= \begin{pmatrix} \omega \cdot \partial_\varphi + i\partial_{xx} & 0 \\ 0 & \omega \cdot \partial_\varphi - i\partial_{xx} \end{pmatrix}, \\ f(\mathbf{u}) \ &:= \begin{pmatrix} if_1(\varphi, x, u, \bar{u}, u_x, \bar{u}_x, u_{xx}, \bar{u}_{xx}) \\ -if_2(\varphi, x, \bar{u}, u, \bar{u}_x, u_x, \bar{u}_{xx}, u_{xx}) \end{pmatrix} \end{aligned} \tag{4.1.12}$$

Hypothesis $(F0)$ is trivial. Hypothesis $(F1)$ holds true with $A_s = \mathbf{X}^s$, $B_s = \mathbf{Z}^s$ by Hypothesis 1.

Hypotheses $(F2)-(F4)$ follow from the fact that $\mathtt{f}$ is a $C^q$ composition operator, see Lemmata A.165, A.166. Let us discuss in detail the property $(F3)$, which we will use in the next section.

Take $\mathbf{u}\in\mathbf{X}^s$, then by our extension rules we have

$$\varepsilon d_{\mathbf{u}}f(\mathbf{u}):=iA_0(\varphi,x,\mathbf{u})+iA_1(\varphi,x,\mathbf{u})\partial_x+iA_2(\varphi,x,\mathbf{u})\partial_{xx}, \tag{4.1.13}$$

where, by (2.1.15), the coefficients of the linear operators $A_j=A_j(\varphi,x,\mathbf{u})$ have the form

$$A_2:=\begin{pmatrix} a_2 & b_2 \\ -\bar{b}_2 & -a_2\end{pmatrix},\quad A_1:=\begin{pmatrix} a_1 & b_1 \\ -\bar{b}_1 & -\bar{a}_1\end{pmatrix},\quad A_0:=\begin{pmatrix} a_0 & b_0 \\ -\bar{b}_0 & -\bar{a}_0\end{pmatrix}. \tag{4.1.14}$$

with

$$\begin{aligned} a_i(\varphi,x)&:=\varepsilon(\partial_{z_i^+}f_1)(\varphi,x,u,\bar{u},u_x,\bar{u}_x,u_{xx},\bar{u}_{xx}),\\ b_i(\varphi,x)&:=\varepsilon(\partial_{z_i^-}f_1)(\varphi,x,u,\bar{u},u_x,\bar{u}_x,u_{xx},\bar{u}_{xx}).\end{aligned} \tag{4.1.15}$$

Thanks to Hypothesis 1, and Remark 3.1.16 one has that $d_{\mathbf{u}}f(\mathbf{u}):\mathbf{X}^0\to\mathbf{Z}^0$ and hence

$$a_i,b_i\in Y^s,\ \ i=0,2,\qquad a_1,b_1\in X^s. \tag{4.1.16}$$

By (4.1.14) and Lemma 4.1.49, the (4.1.16) implies

$$iA_2,iA_0:\mathbf{X}^0\to\mathbf{Z}^0,\quad iA_1\partial_x:\mathbf{X}^0\to\mathbf{Z}^0. \tag{4.1.17}$$

then the operator $\mathcal{L}=d_{\mathbf{u}}F$ maps $\mathbf{X}^0$ to $\mathbf{Z}^0$, i.e. it is *reversible* according to Definition 4.1.47.

The coefficients $a_i$ and $b_i$ and their derivative $d_{u^\sigma}a_i(\mathbf{u})[h]$ with respect to $u^\sigma$ in the direction $h$, for $h\in H^s$, satisfy the following tame estimates.

**Lemma 4.1.50.** *For all $\mathfrak{s}_0\le s\le q-2$, $||u||_{\mathfrak{s}_0+2}\le 1$ we have, for any $i=0,1,2$, $\sigma=\pm1$*

$$||b_i(\mathbf{u})||_s,||a_i(\mathbf{u})||_s\le\varepsilon C(s)(1+||\mathbf{u}||_{s+2}), \tag{4.1.18a}$$

$$||d_{u^\sigma}b_i(\mathbf{u})[h]||,||d_{u^\sigma}a_i(\mathbf{u})[h]||_s\le\varepsilon C(s)(||h||_{s+2}+||\mathbf{u}||_{s+2}||h||_{\mathfrak{s}_0+2}). \tag{4.1.18b}$$

*If moreover $\lambda\to\mathbf{u}(\lambda)\in\mathbf{H}^s$ is a Lipschitz family such that $||\mathbf{u}||_{s,\gamma}\le1$, then*

$$||b_i(\mathbf{u})||_{s,\gamma},||a_i(\mathbf{u})||_{s,\gamma}\le\varepsilon C(s)(1+||\mathbf{u}||_{s+2,\gamma}). \tag{4.1.19}$$

*Proof.* To prove the (4.1.18a) it is enough to apply Lemma A.165(i) to the function $\partial_{z_i^\sigma}f_1$, for any $i=0,1,2$ and $\sigma=\pm1$ which holds for $s+1\le q$. Now, let us write, for any $i=0,1,2$ and $\sigma,\sigma'=\pm$,

$$\begin{aligned} d_{u^\sigma}a_i(\mathbf{u})[h]&\overset{(4.1.14)}{=}\varepsilon\sum_{k=0}^{2}(\partial^2_{z_k^\sigma z_i^+}f_1)(\varphi,x,u,\bar{u},u_x,\bar{u}_x,u_{xx},\bar{u}_{xx})\partial_x^k h,\\ d_{u^\sigma}b_i(\mathbf{u})[h]&\overset{(4.1.14)}{=}\varepsilon\sum_{k=0}^{2}(\partial^2_{z_k^\sigma z_i^-}f_1)(\varphi,x,u,\bar{u},u_x,\bar{u}_x,u_{xx},\bar{u}_{xx})\partial_x^k h,\end{aligned} \tag{4.1.20}$$

Then, by Lemma A.165(i) applied on $\partial^2_{z_k^{\sigma'} z_i^{\sigma}} f_1$ we obtain

$$||(\partial^2_{z_k^{\sigma'} z_i^{\sigma}} f_1)(\varphi, x, u, \bar{u}, u_x, \bar{u}_x, u_{xx}, \bar{u}_{xx})||_s \leq C(s)||f||_{C^{s+2}}(1 + ||u^{\sigma}||_{s+2}), \tag{4.1.21}$$

for $s + 2 \leq q$. The bound (4.1.18b) follows by (A.5) using the (4.1.21). To prove the (4.1.19) one can reason similarly. □

This Lemma ensures property $(F3)$. Properties $(F2)$ and $(F4)$ are proved in exactly in the same way, for property $(F4)$ just consider derivatives of $f$ of order 3.

We have verified all the Hypotheses of Theorem 3.1.18, which ensures the existence of a solution defined on some *possibly empty* set of parameters $\mathcal{G}_\infty$. This concludes the proof of Proposition 4.0.46.

To prove Theorem 1.1.1 we clearly has to show that the set $\mathcal{G}_\infty$ has "large" measure. Hence we need a more explicit formulation of such set. In order to do this we produce a set o parameters defined in terms of the eigenvalues of the linearized operator $\mathcal{L}(\mathbf{u})$ on which we have some estimates on the inverse of $\mathcal{L}$. In order to do this we use a *reducibility* argument. In other words we show that in a rather explicit set of parameters it is possible to conjugate the operator $\mathcal{L}$ to a diagonal linear operator. In this way the problem of the invertibility becomes trivial. Sections 4.2 and 4.3 are devoted to the proof of the following Proposition.

**Proposition 4.1.51 (Diagonalization: reversible case).** *Fix $\gamma \leq \gamma_0, \tau > d$. There exist $\eta, q \in \mathbb{N}$, depending only on $\tau, d$, such that for any nonlinearity $\mathtt{f} \in C^q$ satisfying the Hypotheses 1, there exists $\epsilon_0 > 0$ such that for any $\varepsilon$ with $0 < \varepsilon\gamma^{-1} < \epsilon_0$, for any set $\Lambda_o \subseteq \Lambda \subset \mathbb{R}$ and for any Lipschitz family $\mathbf{u}(\lambda) \in \mathbf{X}^0$ defined on $\Lambda_o$ with $||\mathbf{u}||_{\mathfrak{s}_0+\eta,\gamma} \leq 1$ the following holds. There exist Lipschitz functions $\mu_h^\infty : \Lambda \to i\mathbb{R}$ of the form*

$$\mu_h^\infty := \mu_{\sigma,j}^\infty = -\sigma i m j^2 + r_{\sigma,j}^\infty, \qquad \sup_{h \in \boldsymbol{C} \times \mathbb{N}} |r_h^\infty|_\gamma \leq C\varepsilon, \tag{4.1.22}$$

*with $m \in \mathbb{R}$, $h = (\sigma, j) \in \boldsymbol{C} \times \mathbb{N}$ and $\boldsymbol{C} := \{+1, -1\}$, such that $\mu_{\sigma,j}^\infty = -\mu_{-\sigma,j}^\infty$. Setting*

$$\Lambda_\infty^{2\gamma}(\mathbf{u}) := \left\{ \begin{array}{ll} \lambda \in \Lambda_o : & |\lambda\bar{\omega} \cdot \ell + \mu_{\sigma,j}^\infty(\lambda) - \mu_{\sigma',j'}^\infty(\lambda)| \geq \frac{2\gamma|\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^\tau} \\ & \forall \ell \in \mathbb{Z}^d, \forall (\sigma, j), (\sigma', j') \in \boldsymbol{C} \times \mathbb{N} \end{array} \right\}, \tag{4.1.23}$$

*we have:*
*(i) For $\lambda \in \Lambda_\infty^{2\gamma}$ there exist linear bounded operators $W_1, W_2 : \mathbf{X}^{\mathfrak{s}_0} \to \mathbf{X}^{\mathfrak{s}_0}$ with bounded inverse, such that $\mathcal{L}(\mathbf{u})$ defined in (4.2.30) satisfies*

$$\mathcal{L}(\mathbf{u}) = W_1 \mathcal{L}_\infty W_2^{-1}, \ \mathcal{L}_\infty = \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}_\infty, \ \mathcal{D}_\infty = \operatorname{diag}\{\mu_h^\infty\}_{h \in \boldsymbol{C} \times \mathbb{N}}. \tag{4.1.24}$$

*Moreover, for any* $s \in (\mathfrak{s}_0, q-\eta)$, *if* $||\mathbf{u}||_{s+\eta,\gamma} < +\infty$, *then* $W_i^{\pm 1}$ *are bounded operators* $\mathbf{X}^s \to \mathbf{X}^s$.
*(ii) under the same assumption of* $(i)$, *for any* $\varphi \in \mathbb{T}^d$ *the* $W_i$ *define changes of variables on the phase space*

$$W_i(\varphi), W_i^{-1}(\varphi) : \mathbf{X}^s_x \to \mathbf{X}^s_x, \quad i = 1, 2, \tag{4.1.25}$$

*see Remark 2.2.10. Such operators satisfy the bounds*

$$||(W_i^{\pm 1}(\varphi) - \mathbb{1})\mathbf{h}||_{\mathbf{H}^s_x} \leq \varepsilon\gamma^{-1}C(s)(||\mathbf{h}||_{\mathbf{H}^s_x} + ||\mathbf{u}||_{s+\eta+\mathfrak{s}_0}||\mathbf{h}||_{\mathbf{H}^1_x}). \tag{4.1.26}$$

**Remark 4.1.52.** *The purpose of item (ii) is to prove that a function* $\mathbf{h}(t) \in \mathbf{X}^s_x$ *is a solution of the linearized NLS* (4.1.29) *if and only if the function* $\mathbf{v}(t) := W_2^{-1}(\omega t)[\mathbf{h}(t)] \in \mathbf{H}^s_x$ *solves the constant coefficients dynamical system*

$$\begin{pmatrix} \partial_t v \\ \partial_t \bar{v} \end{pmatrix} + \begin{pmatrix} \mathcal{D}_\infty & 0 \\ 0 & -\mathcal{D}_\infty \end{pmatrix} \begin{pmatrix} v \\ \bar{v} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \quad \dot{v}_j + \mu^\infty_{+,j} v_j = 0 \quad j \in \mathbb{N}, \tag{4.1.27}$$

*Since the eigenvalues are all imaginary we have that*

$$||v(t)||^2_{H^s_x} = \sum_{j\in\mathbb{N}} |v_j(t)|^2 \langle j \rangle^{2s} = \sum_{j\in\mathbb{N}} |v_j(0)|^2 \langle j \rangle^{2s} = ||v(0)||^2_{H^s_x}, \tag{4.1.28}$$

*that means that the Sobolev norm in the space of functions depending on* $x$, *is constant in time.*

Proposition 4.1.51 provides a quite explicit set of parameters (see (4.1.23)) for which it is possible to diagonalize the linearized operator $\mathcal{L}$ at some point $\mathbf{u}$.

To prove a Proposition like 4.0.46 is quite standard when the non linearity $\mathtt{f}$ does not contain derivatives. In this simpler case $\mathcal{L}(\mathbf{u})$ is a diagonal matrix plus a small bounded perturbation. Hence one can use a classical reducibility scheme *à la KAM*. In our case this is not true, indeed

$$\mathcal{L}(\mathbf{u}) = \omega \cdot \partial_\varphi \mathbb{1} + i(\mathbb{1} + A_2(\varphi, x))\partial_{xx} + iA_1(\varphi, x)\partial_x + iA_0(\varphi, x) \tag{4.1.29}$$

where $A_i : \mathbf{H}^s \to \mathbf{H}^s$ are defined in (4.1.14) and $\mathbb{1}$ is the $2 \times 2$ identity. Hence the reduction requires a careful analysis. In particular in Section 4.2 we perform a series of changes of variables which conjugate $\mathcal{L}$ to an operator $\mathcal{L}_4$ which is the sum of an unbounded *diagonal* operator plus a small *bounded* remainder. As we will see the transformation used in Section 4.2 are deeply different from the usual KAM transformations. Then in section 4.3 we perform a KAM reduction algorithm.

## 4.2 The diagonalization algorithm: regularization

For $\mathbf{u} \in \mathbf{X}^0$ we consider the linearized operator

$$\mathcal{L}(\mathbf{u}) := L_\omega + \varepsilon d_{\mathbf{u}} f(\mathbf{u}) = \omega \cdot \partial_\varphi \mathbb{1} + i\big(E + A_2(\varphi, x, \mathbf{u})\big)\partial_{xx} + iA_1\partial_x + iA_0\,, \tag{4.2.30}$$

with $E = \mathrm{diag}(1, -1)$, $d_{\mathbf{u}} f(\mathbf{u})$ defined in formula (4.1.13) and $\|\mathbf{u}\|_{\mathfrak{s}_0+2}$ small. In this Section we prove

**Lemma 4.2.53.** *Let* $\mathtt{f} \in C^q$ *satisfy the Hypotheses of Proposition 4.0.46 and assume* $q > \eta_1 + \mathfrak{s}_0$ *where*

$$\eta_1 := d + 2\mathfrak{s}_0 + 10. \tag{4.2.31}$$

*There exists* $\epsilon_0 > 0$ *such that, if* $\varepsilon\gamma_0^{-1} \leq \epsilon_0$ *(see (1.1.2) for the definition of* $\gamma_0$*) then, for any* $\gamma \leq \gamma_0$ *and for all* $\mathbf{u} \in \mathbf{X}^0$ *depending in a Lipschitz way on* $\lambda \in \Lambda$*, if*

$$||\mathbf{u}||_{\mathfrak{s}_0+\eta_1,\gamma} \leq 1, \tag{4.2.32}$$

*then, for* $\mathfrak{s}_0 \leq s \leq q - \eta_1$*, the following holds.*
*(i) There exist invertible maps* $\mathcal{V}_1, \mathcal{V}_2 : \mathbf{H}^0 \to \mathbf{H}^0$ *such that* $\mathcal{L}_4 := \mathcal{V}_1^{-1}\mathcal{L}\mathcal{V}_2$ *with*

$$\mathcal{L}_4 := \omega \cdot \partial_\varphi \mathbb{1} + i\begin{pmatrix} m & 0 \\ 0 & -m \end{pmatrix}\partial_{xx} + i\begin{pmatrix} 0 & q_1 \\ -\bar{q}_1 & 0 \end{pmatrix}\partial_x + i\begin{pmatrix} q_2 & q_3 \\ -\bar{q}_3 & -\bar{q}_2 \end{pmatrix}, \tag{4.2.33}$$

*where the* $q_i \equiv q_i(\varphi, x)$*.*
*The* $\mathcal{V}_i$ *are reversibility-preserving and moreover for all* $\mathbf{h} \in \mathbf{X}^0$ *and* $i = 1, 2$

$$||\mathcal{V}_i\mathbf{h}||_{s,\gamma} + ||\mathcal{V}_i^{-1}\mathbf{h}||_{s,\gamma} \leq C(s)(||\mathbf{h}||_{s+2,\gamma} + ||\mathbf{u}||_{s+\eta_1,\gamma}||\mathbf{h}||_{\mathfrak{s}_0+2,\gamma})\,. \tag{4.2.34}$$

*(ii) The coefficient* $m := m(\mathbf{u})$ *of* $\mathcal{L}_4$ *satisfies*

$$|m(\mathbf{u}) - 1|_\gamma \leq \varepsilon C, \tag{4.2.35a}$$

$$|d_{\mathbf{u}} m(\mathbf{u})[\mathbf{h}]| \leq \varepsilon C||\mathbf{h}||_{\eta_1}. \tag{4.2.35b}$$

*(iii) The operators* $q_i := q_i(\mathbf{u})$*, are such that*

$$||q_i||_{s,\gamma} \leq \varepsilon C(s)(1 + ||\mathbf{u}||_{s+\eta_1,\gamma}), \tag{4.2.36a}$$

$$||d_{\mathbf{u}}(q_i)(\mathbf{u})[\mathbf{h}]||_s \leq \varepsilon C(s)(||\mathbf{h}||_{s+\eta_1} + ||\mathbf{u}||_{s+\eta_1} + ||\mathbf{h}||_{\mathfrak{s}_0+\eta_1}), \tag{4.2.36b}$$

*Finally* $\mathcal{L}_4$ *is reversible.*

The rest of the Section is devoted to the proof of this Lemma. We divide it in four steps. at each step we construct a *reversibility-preserving* change of variable $\mathcal{T}_i$ that conjugates[1] $\mathcal{L}_i$ to $\mathcal{L}_{i+1}$ where $\mathcal{L}_0 := \mathcal{L}$ and $\mathcal{L}_i :=$

$$\omega\cdot\partial_\varphi \mathbb{1} + i\begin{pmatrix} 1+a_2^{(i)} & b_2^{(i)} \\ -\bar{b}_2^{(i)} & -1-a_2^{(i)} \end{pmatrix}\partial_{xx} + i\begin{pmatrix} a_1^{(i)} & b_1^{(i)} \\ -\bar{b}_1^{(i)} & -\bar{a}_i^{(i)} \end{pmatrix}\partial_x + \begin{pmatrix} a_0^{(i)} & b_0^{(i)} \\ -\bar{b}_0^{(i)} & -\bar{a}_0^{(i)} \end{pmatrix}, \tag{4.2.37}$$

possibly renaming the space and time variables. On the transformation we need to prove bounds like

$$||\mathcal{T}_i(\mathbf{u})\mathbf{h}||_{s,\gamma} \le C(s)(||\mathbf{h}||_{s,\gamma} + ||\mathbf{u}||_{s+\kappa_i,\gamma}||\mathbf{h}||_{\mathfrak{s}_0}), \tag{4.2.38a}$$

$$\begin{aligned} ||d_{\mathbf{u}}\mathcal{T}_i(\mathbf{u})[\mathbf{h}]\mathbf{g}||_s \le \varepsilon C(s)\,(&||\mathbf{g}||_{s+1}||\mathbf{h}||_{\mathfrak{s}_0+\kappa_i} + ||\mathbf{g}||_2||\mathbf{h}||_{s+\kappa_i}+ \\ &+||\mathbf{u}||_{s+\kappa_i}||\mathbf{g}||_2||\mathbf{h}||_{\mathfrak{s}_0})\,, \end{aligned} \tag{4.2.38b}$$

for suitable $\kappa_i$. We prove the same for $\mathcal{T}_i^{-1}$. Moreover the coefficients in (4.2.37) satisfy

$$||a_j^{(i)}(\mathbf{u})||_{s,\gamma} \le \varepsilon C(s)(1 + ||\mathbf{u}||_{s+\kappa_i,\gamma}), \tag{4.2.39a}$$

$$||d_{u^\sigma} a_j^{(i)}(\mathbf{u})[h]||_s \le \varepsilon C(s)(||\mathbf{h}||_{s+\kappa_i} + ||\mathbf{u}||_{s+\kappa_i} + ||\mathbf{h}||_{\mathfrak{s}_0+\kappa_i}), \tag{4.2.39b}$$

for $j = 0, 1, 2$ and $i = 1, \dots, 4$. We prove the same for $b_j^{(i)}$.

**Step 1. Diagonalization of the second order coefficient**

We first diagonalize the term $E + A_2$ in (4.2.30). By a direct calculation, one can see that the matrix $(E + A_2)$ has eigenvalues $\lambda_{1,2} = \pm\sqrt{(1 + a_2)^2 - |b_2|^2}$. Hence we set $a_2^{(1)}(\varphi, x) = \lambda_1 - 1$. We have that $a_2^{(1)} \in \mathbb{R}$ because $a_2 \in \mathbb{R}$ and $a_i, b_i$ are small. The diagonalizing matrix is

$$\mathcal{T}_1^{-1} := \frac{1}{2}\begin{pmatrix} 2 + a_2 + a_2^{(1)} & b_2 \\ -\bar{b}_2 & -(2 + a_2 + a_2^{(1)}) \end{pmatrix}. \tag{4.2.40}$$

The tame estimates (4.2.39) for $a_2^{(1)}$ and the (4.2.38) on $\mathcal{T}_1^{-1}$ follow with $\kappa_1 = 2$ by (4.1.18a), (4.2.32) and (A.5). The bound on $\mathcal{T}_1$ follows since

$$\det\mathcal{T}_1^{-1} = (|b_2|^2 - (2 + a_2 + a_2^{(1)})^2)/4,$$

and by using the same strategy as for $a_2^{(1)}$. One has

$$\begin{aligned} \mathcal{L}_1 \;\; := \;\; & \mathcal{T}_1^{-1}\mathcal{L}\mathcal{T}_1 = \omega\cdot\partial_\varphi\mathbb{1} + i\mathcal{T}_1^{-1}(E + A_2)\mathcal{T}_1\partial_{xx} + \\ + \;\; & i\left[2\mathcal{T}_1^{-1}(E + A_2)\partial_x\mathcal{T}_1 + \mathcal{T}_1^{-1}A_1\mathcal{T}_1\right]\partial_x \\ + \;\; & i\left[-i\mathcal{T}_1^{-1}(\omega\cdot\partial_\varphi\mathcal{T}_1) + \mathcal{T}_1^{-1}(E + A_2)\partial_{xx}\mathcal{T}_1 + \mathcal{T}_1^{-1}A_1\partial_x\mathcal{T}_1 + \mathcal{T}_1^{-1}A_0\mathcal{T}_1\right]; \end{aligned} \tag{4.2.41}$$

[1] Actually in the third step we only are able to conjugate $\mathcal{L}_2$ to $\rho\mathcal{L}_3$, where $\rho$ is a suitable function. This is the reason why $\mathcal{L}$ is semi-conjugated to $\mathcal{L}_4$.

the (6.4.56) has the form (4.2.37) and this identifies the coefficients $a_j^{(1)}, b_j^{(1)}$. Note that the matrix of the second order operator is now diagonal. Moreover, by (A.5), (4.2.39) on $a_2^{(1)}$, (4.1.18a) and (4.1.18b) one obtains the bounds (4.2.39) for the remaining coefficients $a_j^{(1)}, b_j^{(1)}$ with $\kappa_1 := 5$. Then we can fix $\kappa_1 = 5$ in all the bounds (4.2.38a)-(4.2.39b) even if for some of the coefficients there are better bounds.

Finally, since the matrix $\mathcal{T}_1^{-1}$ is $E + A_2$ plus a diagonal matrix with even components, it has the same parity properties of $A_2$, then maps $\mathbf{Y}^s$ to $\mathbf{Y}^s$ and $\mathbf{X}^s$ to $\mathbf{X}^s$, this means that it is reversibility-preserving and hence $\mathcal{L}_1$ is reversible. In particular one has that $a_2^{(1)}, a_0^{(1)}, b_0^{(1)} \in Y^0$ and $a_1^{(1)}, b_1^{(1)} \in X^0$ then by Lemma 4.1.49.

**Remark 4.2.54.** *We can note that in the* quasi-linear *case this first step can be avoided. Indeed in that case one has $\partial_{\bar{z}_2} f \equiv 0$, so that the matrix $A_2$ is already diagonal, with real coefficients.*

### Step 2. Change of the space variable

We consider a $\varphi-$dependent family of diffeomorphisms of the $1-$dimensional torus $\mathbb{T}$ of the form

$$y = x + \xi(\varphi, x), \tag{4.2.42}$$

where $\xi$ is as small real-valued function, $2\pi$ periodic in all its arguments. The change of variables (4.2.42) induces on the space of functions the invertible linear operator

$$(\mathcal{T}_2 h)(\varphi, x) := h(\varphi, x + \xi(\varphi, x)), \text{ with } \quad (\mathcal{T}_2^{-1} v)(\varphi, y) = v(\varphi, y + \widehat{\xi}(\varphi, y)), \tag{4.2.43}$$

where $y \to y + \widehat{\xi}(\varphi, y)$ is the inverse diffeomorphism of (4.2.42). With a slight abuse of notation we extend the operator to $\mathbf{H}^s$:

$$\mathcal{T}_2 : \mathbf{H}^s \to \mathbf{H}^s, \quad \mathcal{T}_2 \mathbf{h} = \begin{pmatrix} (\mathcal{T}_2 h)(\varphi, x) \\ (\mathcal{T}_2 \bar{h})(\varphi, x) \end{pmatrix}. \tag{4.2.44}$$

Now we have to calculate the conjugate $\mathcal{T}_2{}^{-1} \mathcal{L}_1 \mathcal{T}_2$ of the operator $\mathcal{L}_1$ in (6.4.56).
The conjugate $\mathcal{T}_2^{-1} a \mathcal{T}_2$ of any multiplication operator $a : h(\varphi, x) \to a(\varphi, x) h(\varphi, x)$ is the multiplication operator $(\mathcal{T}_2^{-1} a) : v(\varphi, y) \to (\mathcal{T}_2^{-1} a)(\varphi, y) v(\varphi, y)$. The conjugate of the differential operators will be

$$\begin{aligned} \mathcal{T}_2^{-1} \omega \cdot \partial_\varphi \mathcal{T}_2 &= \omega \cdot \partial_\varphi + [\mathcal{T}_2^{-1}(\omega \cdot \partial_\varphi \xi)] \partial_y, \quad \mathcal{A}^{-1} \partial_x \mathcal{T}_2 = [\mathcal{T}_2^{-1}(1 + \xi_x)] \partial_y, \\ \mathcal{T}_2^{-1} \partial_{xx} \mathcal{A} &= [\mathcal{T}_2^{-1}(1 + \xi_x)^2] \partial_{yy} + [\mathcal{T}_2^{-1}(\xi_{xx})] \partial_y, \end{aligned} \tag{4.2.45}$$

where all the coefficients are periodic functions of $(\varphi, x)$. Thus we have obtained $\mathcal{L}_2 = \mathcal{T}_2{}^{-1} \mathcal{L}_1 \mathcal{T}_2$ where $\mathcal{L}_2$ has the form (4.2.37) in the variable $y$ instead of $x$. Note that the second rows are the complex conjugates of the first, this is due to the fact that $\mathcal{T}_2$ trivially preserves the subspace $\mathcal{U}$. We have

$$\begin{aligned} 1+a_2^{(2)}(\varphi,y) &= \mathcal{T}_2^{-1}[(1+a_2^{(1)})(1+\xi_x)^2], \qquad b_1^{(2)}(\varphi,y)=\mathcal{T}_2^{-1}[b_1^{(1)}(1+\xi_x)], \\ a_1^{(2)}(\varphi,y) &= \mathcal{T}_2^{-1}((1+a_2^{(1)})\xi_{xx})-i\mathcal{T}_2^{-1}(\omega\cdot\partial_\varphi\xi)+\mathcal{T}_2^{-1}[a_1^{(1)}(1+\xi_x)], \\ a_0^{(2)}(\varphi,y) &= \mathcal{T}_2^{-1}[a_0^{(1)}], \qquad b_0^{(2)}(\varphi,y)=\mathcal{T}_2^{-1}[b_0^{(1)}]. \end{aligned} \tag{4.2.46}$$

We are looking for $\xi(\varphi,x)$ such that the coefficient of the second order differential operator does not depend on $y$, namely

$$\mathcal{T}_2^{-1}[(1+a_2^{(1)})(1+\xi_x)^2]=1+a_2^{(2)}(\varphi), \tag{4.2.47}$$

for some function $a_2^{(2)}(\varphi)$. Since $\mathcal{T}_2$ operates only on the space variables, the (6.4.90) is equivalent to

$$(1+a_2^{(1)}(\varphi,x))(1+\xi_x(\varphi,x))^2=1+a_2^{(2)}(\varphi). \tag{4.2.48}$$

Hence we have to set

$$\xi_x(\varphi,x)=\rho_0, \qquad \rho_0(\varphi,x):=(1+a_2^{(2)})^{\frac12}(\varphi)(1+a_2^{(1)}(\varphi,x))^{-\frac12}-1, \tag{4.2.49}$$

that has solution $\xi$ periodic in $x$ if and only if $\int_{\mathbb{T}}\rho_0 dy=0$. This condition implies

$$a_2^{(2)}(\varphi)=\left(\frac{1}{2\pi}\int_{\mathbb{T}}(1+a_2^{(1)}(\varphi,x))^{-\frac12}\right)^{-2}-1 \tag{4.2.50}$$

Then we have the solution (with zero average) of (6.4.92)

$$\xi(\varphi,x):=(\partial_x^{-1}\rho_0)(\varphi,x), \tag{4.2.51}$$

where $\partial_x^{-1}$ is defined by linearity as

$$\partial_x^{-1}e^{ikx}:=\frac{e^{ikx}}{ik}, \quad \forall\, k\in\mathbb{Z}\backslash\{0\}, \quad \partial_x^{-1}=0. \tag{4.2.52}$$

In other word $\partial_x^{-1}h$ is the primitive of $h$ with zero average in $x$. Thus, conjugating $\mathcal{L}_1$ through the operator $\mathcal{T}_2$ in (4.2.44) , we obtain the operator $\mathcal{L}_2$ in (4.2.37).

Now we start by proving that the coefficient $a_2^{(2)}$ satisfies tame estimates like (4.2.39) with $\kappa_2=2$. Let us write

$$\begin{aligned} a_2^{(2)}(\varphi) &= \psi\left(G[g(a_2^{(1)})-g(0)]\right)-\psi(0), \\ \psi(t) &:= (1+t)^{-2}, \quad Gh:=\frac{1}{2\pi}\int_{\mathbb{T}}hdx, \quad g(t):=(1+t)^{-\frac12}. \end{aligned} \tag{4.2.53}$$

Then one has, for $\varepsilon$ small,

$$||a_2^{(2)}||_s \overset{(A.10)}{\leq} C(s)||G[g(a_2^{(1)}) - g(0)]||_s \leq C(s)||g(a_2^{(1)}) - g(0)||_s \overset{(A.10)}{\leq} C(s)||a_2^{(1)}||_s. \tag{4.2.54}$$

In the first case we used (A.10) on the function $\psi$ with $u = 0, p = 0, h = G[g(a_2^{(1)}) - g(0)]$, while in the second case we have set $u = 0, p = 0, h = a_2^{(1)}$ and used the estimate on $g$. Then we used the (4.2.39) and the bound (4.1.18a), with $s_0 = \mathfrak{s}_0$ which holds for $s + 2 \leq q$. By (4.2.53), we get for $\sigma = \pm 1$

$$d_{u^\sigma} a_2^{(2)}(\mathbf{u})[h] = \psi' \left( G[g(a_2^{(1)}) - g(0)] \right) G \left[ g'(a_2^{(1)}) d_{u^\sigma} a_2^{(1)}[h] \right] \tag{4.2.55}$$

Using (A.5) with $s_0 = \mathfrak{s}_0$, Lemma A.165(i) to estimate the functions $\psi'$ and $g'$, as done in (4.2.54)), and by the (4.1.18b) we get (4.2.39b). The (4.2.39a) follows by (4.2.54), (4.2.39b) and Lemma A.166. The second step is to give tame estimates on the function $\xi = \partial_x^{-1} \rho_0$ defined in (6.4.92) and (6.4.94). It is easy to check that, estimates (4.2.39) are satisfied also by $\rho_0$. They follow by using the estimates on $a_2^{(2)}$ and the estimates (4.2.39), (4.1.18a), (4.1.18b), (4.1.19) for $a_2^{(1)}$. By defining $|u|_s^\infty := ||u||_{W^{s,\infty}}$ and using Lemma A.164(i) we get

$$|\xi|_s^\infty \leq C(s)||\xi||_{s+\mathfrak{s}_0} \leq C(s)||\rho_0||_{s+\mathfrak{s}_0} \leq \varepsilon C(s)(1 + ||\mathbf{u}||_{s+\mathfrak{s}_0+2}), \tag{4.2.56a}$$

$$|d_{u^\sigma} \xi(\mathbf{u})[h]|_s^\infty \leq \varepsilon C(s)(||h||_{s+\mathfrak{s}_0+2} + ||\mathbf{u}||_{s+\mathfrak{s}_0+2}||h||_{\mathfrak{s}_0+2}), \tag{4.2.56b}$$

and hence, by Lemma A.166 one has

$$|\xi|_{s,\gamma}^\infty \leq \varepsilon C(s)(1 + ||\mathbf{u}||_{s+\mathfrak{s}_0+2,\gamma}), \tag{4.2.57}$$

for any $s + \mathfrak{s}_0 + 2 \leq q$. The diffeomorphism $x \mapsto x + \xi(\varphi, x)$ is well-defined if $|\xi|_{1,\infty} \leq 1/2$, but it is easy to note that this condition is implied requiring $\varepsilon C(s)(1 + ||\mathbf{u}||_{\mathfrak{s}_0+3}) \leq 1/2$. Let us study the inverse diffeomorphism $(\varphi, y) \mapsto (\varphi, y + \widehat{\xi}(\varphi, y))$ of $(\varphi, x) \mapsto (\varphi, x + \gamma(\varphi, x))$. Using Lemma A.167(i) on the torus $\mathbb{T}^{d+1}$, one has

$$|\widehat{\xi}|_s^\infty \leq C|\xi|_s^\infty \leq \varepsilon C(s)(1 + ||\mathbf{u}||_{s+\mathfrak{s}_0+2}). \tag{4.2.58}$$

By definition we have that $\widehat{\xi}(\varphi, y) + \xi(\varphi, y + \widehat{\xi}(\varphi, y)) = 0$, which implies, for $\sigma = \pm 1$,

$$|d_{u^\sigma} \widehat{\xi}(\mathbf{u})[h]|_s^\infty \leq \varepsilon C(||h||_{\mathfrak{s}_0+2} + ||\mathbf{u}||_{s+\mathfrak{s}_0+3}||h||_{\mathfrak{s}_0+2}). \tag{4.2.59}$$

Now, thanks to bounds (4.2.58) and (4.2.59), using again Lemma A.166 with $p = \mathfrak{s}_0 + 3$, we obtain

$$|\widehat{\xi}|_{s,\gamma}^\infty \leq \varepsilon C(s)(1 + ||\mathbf{u}||_{s+\mathfrak{s}_0+3,\gamma}). \tag{4.2.60}$$

We have to estimate $\mathcal{T}_2(\mathbf{u})$ and $\mathcal{T}_2^{-1}(\mathbf{u})$. By using (A.16c), (4.2.57) and (4.2.60), we get the (4.2.38a) with $\kappa_2 = \mathfrak{s}_0 + 3$, Now, since

$$d_{\mathbf{u}}(\mathcal{T}_2(\mathbf{u})g)[\mathbf{h}] := d_{\mathbf{u}}g(\varphi, x + \xi(\varphi, x; \mathbf{u})) = (\mathcal{T}_2(\mathbf{u})g_x)d_{\mathbf{u}}\xi(\mathbf{u})[\mathbf{h}],$$

we get the (4.2.38b) using the (A.7), (4.2.56b) and (4.2.38a). The (4.2.38b) on $\mathcal{T}_2^{-1}$ follows by the same reasoning. Finally, using the bounds (A.7), (4.2.38), (4.2.60), (4.1.19), Lemma 4.1.50 and $||\mathbf{u}||_{\mathfrak{s}_0+\eta_1,\gamma} \leq 1$, one has the (4.2.39a) on the coefficients $a_j^{(2)}, b_j^{(2)}$ for $j = 0, 1$ in (4.2.46). Now, by definition (4.2.46), we can write

$$a_1^{(2)} = \mathcal{T}_2^{-1}(\mathbf{u})\rho_1, \quad \rho_1 := (1 + a_2^{(1)})\xi_{xx} - i\omega \cdot \partial_\varphi \xi + a_1^{(1)}(1 + \xi_x), \tag{4.2.61}$$

so that, thanks to bounds in Lemma 4.1.50, and (4.2.56a), (4.2.56b), (A.7) and recalling that $||\mathbf{u}||_{\mathfrak{s}_0+\eta_1} \leq 1$, we get the (4.2.39a) on $\rho_1$. Now, the (4.2.39b) on $a_1^{(2)}$ follows by using the chain rule, setting $\kappa_2 = \mathfrak{s}_0+5$ and for $s + \mathfrak{s}_0 + 5 \leq q$. The same bounds on the coefficients $a_0^{(2)}, b_0^{(2)}$ are obtained in the same way.

**Remark 4.2.55.** *Note that $\xi$ is a real function and $\xi(\varphi, x) \in X^0$ since $a \in Y^0$. This implies that the operators $\mathcal{T}_2$ and $\mathcal{T}_2^{-1}$ map $X^0 \to X^0$ and $Y^0 \to Y^0$, namely preserves the parity properties of the functions. Moreover we have that $a_2^{(2)}, a_0^{(2)}, b_0^{(2)} \in Y^0$, while $a_1^{(2)}, b_1^{(2)} \in X^0$. Then then by Lemma 4.1.49, one has that the operator $\mathcal{L}_2$ is reversible.*

**Step 3. Time reparameterization**

We have obtained $\mathcal{L}_2$ which is an operator of the form (4.2.37) in the variables $(\varphi, y)$. In this section we want to make constant the coefficient of the highest order spatial derivative operator $\partial_{yy}$ of $\mathcal{L}_2$, by a quasi-periodic reparameterization of time. We consider a diffeomorphism of the torus $\mathbb{T}^d$ of the form

$$\theta = \varphi + \omega\alpha(\varphi), \quad \varphi \in \mathbb{T}^d, \quad \alpha(\varphi) \in \mathbb{R}, \tag{4.2.62}$$

where $\alpha$ is a small real valued function, $2\pi-$periodic in all its arguments. The induced linear operator on the space of functions is

$$(\mathcal{T}_3 h)(\varphi, y) := h(\varphi + \omega\alpha(\varphi), y), \quad \text{with} \quad (\mathcal{T}_3^{-1}v)(\theta, y) = v(\theta + \omega\widehat{\alpha}(\theta), y), \tag{4.2.63}$$

where $\varphi = \theta + \omega\widehat{\alpha}(\theta)$ is the inverse diffeomorphism of $\theta = \varphi + \omega\alpha(\varphi)$. We extend the operator

$$\mathcal{T}_3 : \mathbf{H}^s \to \mathbf{H}^s, \quad (\mathcal{T}_3\mathbf{h})(\varphi, x) = \begin{pmatrix} (\mathcal{T}_3 h)(\varphi, x) \\ (\mathcal{T}_3 \bar{h})(\varphi, x) \end{pmatrix}. \tag{4.2.64}$$

By conjugation, we have that the differential operator becomes

$$\mathcal{T}_3^{-1}\omega\cdot\partial_\varphi\mathcal{T}_3=\rho(\theta)\omega\cdot\partial_\theta,\quad \mathcal{T}_3^{-1}\partial_y\mathcal{T}_3=\partial_y,\quad \rho(\theta):=\mathcal{T}_3^{-1}(1+\omega\cdot\partial_\varphi\alpha). \tag{4.2.65}$$

We have obtained $\mathcal{T}_3{}^{-1}\mathcal{L}_2\mathcal{T}_3=\rho\mathcal{L}_3$ with $\mathcal{L}_3$ as in (4.2.37) where

$$\begin{aligned}1+a_2^{(3)}(\theta)&:=(\mathcal{T}_3^{-1}(1+a_2^{(2)}))(\theta),\\ \rho(\theta)a_j^{(3)}(\theta,y)&:=(\mathcal{T}_3^{-1}a_j^{(2)})(\theta,y),\qquad \rho(\theta)b_j^{(3)}(\theta,y):=(\mathcal{T}_3^{-1}b_j^{(2)})(\theta,y),\end{aligned} \tag{4.2.66}$$

for $j=0,1$. We look for solutions $\alpha$ such that the coefficients of the highest order derivatives ($i\omega\cdot\partial_\theta$ and $\partial_{yy}$) are proportional, namely

$$(\mathcal{T}_3^{-1}(1+a_2^{(2)}))(\theta)=m\rho(\theta)=m\mathcal{T}_3^{-1}(1+\omega\cdot\partial_\varphi\alpha) \tag{4.2.67}$$

for some constant $m$, that is equivalent to require that

$$1+a_2^{(2)}(\varphi)=m(1+\omega\cdot\partial_\varphi\alpha(\varphi)), \tag{4.2.68}$$

By setting

$$m=\frac{1}{(2\pi)^d}\int_{\mathbb{T}^d}(1+a_2^{(2)}(\varphi))d\varphi, \tag{4.2.69}$$

we can find the (unique) solution of (4.2.68) with zero average

$$\alpha(\varphi):=\frac{1}{m}(\omega\cdot\partial_\varphi)^{-1}(1+a_2^{(2)}-m)(\varphi), \tag{4.2.70}$$

where $(\omega\cdot\partial_\varphi)^{-1}$ is defined by linearity

$$(\omega\cdot\partial_\varphi)^{-1}e^{i\ell\cdot\varphi}:=\frac{e^{i\ell\cdot\varphi}}{i\omega\cdot\ell},\ \ell\neq 0,\quad (\omega\cdot\partial_\varphi)^{-1}1=0.$$

thanks to this choice of $\alpha$ we have $\mathcal{T}_3^{-1}\mathcal{L}_2\mathcal{T}_3=\rho\mathcal{L}_3$ with $1+a_2^{(3)}(\theta)=m$.
First of all, note that the bounds (4.2.35) on the coefficient $m$ in (6.4.106) follow by the (4.2.39) for $a_2^{(2)}$.
Moreover the function $\alpha(\varphi)$ defined in (6.4.107) satisfies the tame estimates:

$$|\alpha|_s^\infty\leq\varepsilon\gamma_0^{-1}C(s)(1+||\mathbf{u}||_{s+d+\mathfrak{s}_0+2}), \tag{4.2.71a}$$

$$|d_{\mathbf{u}}\alpha(\mathbf{u})[\mathbf{h}]|_s^\infty\leq\varepsilon\gamma_0^{-1}C(s)(||\mathbf{h}||_{s+d+\mathfrak{s}_0+2}+||\mathbf{u}||_{s+d+\mathfrak{s}_0+2}||\mathbf{h}||_{d+\mathfrak{s}_0+2}), \tag{4.2.71b}$$

$$|\alpha|_{s,\gamma}^\infty\leq\varepsilon\gamma_0^{-1}C(s)(1+||\mathbf{u}||_{s+d+\mathfrak{s}_0+2,\gamma}). \tag{4.2.71c}$$

Since $\omega=\lambda\bar{\omega}$ and by (1.1.2) one has $|\bar{\omega}\cdot\ell|\geq 3\gamma_0|\ell|^{-d}$, $\forall\ \ell\neq 0$, then one has the (4.2.71a). One can prove similarly the (4.2.71c) by using (4.2.39a), (4.2.35) and the fact $(\omega\cdot\ell)^{-1}=\lambda^{-1}(\bar{\omega}\cdot\ell)^{-1}$. To prove (4.2.71b) we compute

$$d_{\mathbf{u}}\alpha(\varphi;\mathbf{u})[\mathbf{h}]=(\lambda\bar{\omega}\cdot\partial_\varphi)^{-1}\left(\frac{d_{\mathbf{u}}(1+a_2^{(2)}(\mathbf{u}))[\mathbf{h}]m-(1+a_2^{(2)})d_{\mathbf{u}}m(\mathbf{u})[\mathbf{h}]}{m^2}\right) \tag{4.2.72}$$

and use the estimates (4.2.39a), (4.2.39b) and (4.2.35). Finally, the diffeomorphism (6.4.101) is well-defined if $|\alpha|_1^\infty \leq 1/2$. This is implied by (4.2.71a) and (4.2.32) for $\varepsilon$ small enough.
The inverse diffeomorphism $\theta \to \theta + \omega\widehat{\alpha}(\theta)$ of (6.4.101) satisfies the same estimates in (4.2.71) with $d + \mathfrak{s}_0 + 3$. The (4.2.71a), (4.2.71c) on $\widehat{\alpha}$ follow by the bounds (A.14), (A.15) in Lemma A.167 and (4.2.71a), (4.2.71c). As in the second step the estimate on $d_{\mathbf{u}}\widehat{\alpha}(\mathbf{u})[\mathbf{h}]$ follows by the chain rule using Lemma A.167(iii), (A.6), (4.2.71a), (4.2.71b) on $\alpha$ and (A.2) with $a = d + \mathfrak{s}_0 + 3$, $b = d + \mathfrak{s}_0 + 1$ and $p = s - 1$, $q = 2$ one has the (4.2.71b) for $\widehat{\alpha}$.
We claim that the operators $\mathcal{T}_3(\mathbf{u})$ and $\mathcal{T}_3^{-1}(\mathbf{u})$ defined in (5.1.46), satisfy for any $\mathbf{g}, \mathbf{h} \in \mathbf{H}^s$ the (4.2.38) with $\kappa_3 := d + \mathfrak{s}_0 + 3$. Indeed to prove estimates (4.2.38a), we apply Lemma A.167(ii) and the estimates (4.2.71a), (4.2.71c) on $\alpha$ and $\widehat{\alpha}$ obtained above. Now, since

$$d_{\mathbf{u}}(\mathcal{T}_3(\mathbf{u})\mathbf{g})[\mathbf{h}] = \mathcal{T}_3(\mathbf{u})(\omega \cdot \partial_\varphi \mathbf{g}) d_{\mathbf{u}}\alpha(\mathbf{u})[\mathbf{h}] \tag{4.2.73}$$

then (A.7), (4.2.71b) and (4.2.38a), imply (4.2.38b). Reasoning in the same way one has that (4.2.71a), (4.2.38b) imply (4.2.38b) on $\mathcal{T}_3^{-1}$.
By the (4.2.65) one has $\rho = 1 + \mathcal{T}_3^{-1}(\omega \cdot \partial_\varphi \alpha)$. By using the (A.17a), (A.17b), the bounds (4.2.71) on $\alpha$ and (4.2.32) one can prove

$$|\rho - 1|_{s,\gamma}^\infty \leq \varepsilon \gamma_0^{-1} C(s)(1 + ||\mathbf{u}||_{s+d+\mathfrak{s}_0+4,\gamma}) \tag{4.2.74a}$$
$$|d_{\mathbf{u}}\rho(\mathbf{u})[\mathbf{h}]|_s^\infty \leq \varepsilon \gamma_0^{-1} C(s)(||\mathbf{h}||_{s+d+\mathfrak{s}_0+3} + ||\mathbf{u}||_{s+d+\mathfrak{s}_0+4}||\mathbf{h}||_{d+\mathfrak{s}_0+3}). \tag{4.2.74b}$$

Bounds (4.2.39) on the coefficients $a_j^{(3)}, b_j^{(3)}$ follows, with $\kappa_3 := d + \mathfrak{s}_0 + 5$, by using the (4.2.74) on $\rho$, the (4.2.38) on $\mathcal{T}_3$ and $\mathcal{T}_3^{-1}$, the (A.5)-(A.7) and the condition (4.2.32).

**Remark 4.2.56.** *Note that $\alpha$ is a real function and $\alpha \in X^0$, then the operators $\mathcal{T}_3$ and $\mathcal{T}_3^{-1}$ map $X^0 \to X^0$ and $Y^0 \to Y^0$. Moreover we have that $m \in \mathbb{R}$ , $a_0^{(3)}, b_0^{(3)} \in Y^0$, while $a_1^{(3)}, b_1^{(3)} \in X^0$. Then then by Lemma 4.1.49, one has that the operator $\mathcal{L}_3$ is reversible.*

In the following we rename $y = x$ and $\theta = \varphi$.

### Step 4. Descent Method: conjugation by multiplication operator

The aim of this section is to conjugate the operator $\mathcal{L}_3$ to an operator $\mathcal{L}_4$ which has zero on the diagonal of the first order spatial differential operator.

We consider an operator of the form

$$\mathcal{T}_4 := \begin{pmatrix} 1 + z(\varphi, x) & 0 \\ 0 & 1 + \bar{z}(\varphi, x) \end{pmatrix}, \tag{4.2.75}$$

where $z : \mathbb{T}^{d+1} \to \mathbb{C}$ is small enough so that $\mathcal{T}_4$ is invertible. By a direct calculation we have that $\mathcal{L}_4$ has the form (4.2.37) where the second order coefficients are those of $\mathcal{L}_3$ while[2]

$$\begin{aligned} a_1^{(4)} &:= 2m\frac{z_x}{1+z} + a_1^{(3)}, \quad q_2 \equiv a_0^{(4)} := \frac{-i(\omega\cdot\partial_\varphi z) + m z_{xx}}{1+z} + a_0^{(3)}, \\ q_1 \equiv b_1^{(4)} &:= b_1^{(3)}\frac{1+\bar{z}}{1+z}, \qquad q_3 \equiv b_0^{(4)} := b_0^{(3)}\frac{1+\bar{z}}{1+z}. \end{aligned} \tag{4.2.76}$$

We look for $z(\varphi,x)$ such that $a_1^{(4)} \equiv 0$. If we look for solutions of the form $1 + z(\varphi,x) = \exp(s(\varphi,x))$ we have that $a_1^{(4)} = 0$ becomes

$$\mathrm{Re}(s_x) = -\frac{1}{2m}\mathrm{Re}(a_1^{(3)}), \qquad \mathrm{Im}(s_x) = -\frac{1}{2m}\mathrm{Im}(a_1^{(3)}), \tag{4.2.77}$$

that have unique (with zero average in $x$) solution

$$\mathrm{Re}(s) = -\frac{1}{2m}\partial_x^{-1}\mathrm{Re}(a_1^{(3)}), \qquad \mathrm{Im}(s) = -\frac{1}{2m}\partial_x^{-1}\mathrm{Im}(a_1^{(3)}) \tag{4.2.78}$$

where $\partial_x^{-1}$ is defined in (4.2.52).
The function $s$ defined in (4.2.78) satisfies the following tame estimates:

$$||s||_{s,\gamma} \le \varepsilon C(s)(1 + ||\mathbf{u}||_{s+d+\mathfrak{s}_0+5,\gamma}), \tag{4.2.79a}$$

$$||d_{\mathbf{u}} s(\mathbf{u})[\mathbf{h}]||_s \le \varepsilon C(s)(||\mathbf{h}||_{s+d+\mathfrak{s}_0+4} + ||\mathbf{u}||_{s+d+\mathfrak{s}_0+5}||\mathbf{h}||_{d+\mathfrak{s}_0+4}). \tag{4.2.79b}$$

The (4.2.79) follow by (4.2.35), used to estimate $m$, the estimates (4.2.39), on the coefficient of $a_1^{(3)}$, and (4.2.32). Since by definition one has

$$z(\varphi,x) = \exp(s(\varphi,x)) - 1,$$

clearly the function $z$ satisfies the same estimates (4.2.79a)-(4.2.79b).
The estimates (4.2.79a)-(4.2.79b) on the function $z(\varphi,x)$ imply directly the tame estimates in (4.2.38) on the operator $\mathcal{T}_4$ defined in (4.2.75). The bound (4.2.38a) on the operator $\mathcal{T}_4^{-1}$ follows in the same way. In order to prove the (4.2.38b) we note that

$$d_{\mathbf{u}}\mathcal{T}_4^{-1}(\mathbf{u})[\mathbf{h}] = -\mathcal{T}_4^{-1}(\mathbf{u})d_{\mathbf{u}}\mathcal{T}_4(\mathbf{u})[\mathbf{h}]\mathcal{T}_4^{-1}(\mathbf{u}),$$

then, using the (4.2.32) and the (4.2.38b) on $\mathcal{T}_4$ we get the (4.2.38b) on $\mathcal{T}_4^{-1}$. We show that the coefficients in (4.2.76), for $i = 1,2,3$ satisfy the tame estimates in (4.2.39) with $\kappa_4 = d + \mathfrak{s}_0 + 7$ that simply are the (4.2.36a), (4.2.36b). The strategy to prove the tame bounds on $q_i$ is the same used in (4.2.61) on $a_1^{(2)}$. Collecting together the loss of regularity at each step one gets $\eta_1$ as in (4.2.31).

[2]We use $\mathcal{T}_4$ to cancel $a_1^{(4)}$, then to avoid apices we rename the remaining coefficients coherently with the definition of $\mathcal{L}_4$.

**Remark 4.2.57.** *Since $a_1^{(3)} \in X^0$, then $s(\varphi, x) \in Y^0$, so that the operator $\mathcal{T}_4$ does not change the parity properties of functions. This implies that the operator $\mathcal{L}_4$, defined in (4.2.33), is reversible.*

The several steps performed in the previous sections (semi)-conjugate the linearized operator $\mathcal{L}$ to the operator $\mathcal{L}_4$ defined in (4.2.33), namely

$$\mathcal{L} = \mathcal{V}_1 \mathcal{L}_4 \mathcal{V}_2^{-1}, \quad \mathcal{V}_1 := \mathcal{T}_1 \mathcal{T}_2 \mathcal{T}_3 \rho \mathcal{T}_4, \quad \mathcal{V}_2 = \mathcal{T}_1 \mathcal{T}_2 \mathcal{T}_3 \mathcal{T}_4. \tag{4.2.80}$$

where $\rho$ is the multiplication operator by the function $\rho$ defined in (4.2.65). Now by Lemma A.168, the operators $\mathcal{V}_1$ and $\mathcal{V}_2$ defined in (4.2.80) satisfy, using (4.2.32), the (4.2.34). Note that we used that $\eta_1 > d + 2\mathfrak{s}_0 + 7$. The estimates in (ii) and (iii) have been already proved, hence the proof of Lemma 4.2.53 has been completed. □

The following Lemma is a consequence of the discussion above.

**Lemma 4.2.58.** *Under the Hypotheses of Lemma 4.2.53 possibly with smaller $\epsilon_0$, if (4.2.32) holds, one has that the$\mathcal{T}_i$ $i \neq$ identify operators $\mathcal{T}_i(\varphi)$, of the phase space $\mathbf{H}^s_x := \mathbf{H}^s(\mathbb{T})$. Moreover they are invertible and the following estimates hold for $\mathfrak{s}_0 \leq s \leq q - \eta_1$ and for $i = 1, 2, 4$:*

$$||(\mathcal{T}_i^{\pm 1}(\varphi) - \mathbb{1})\mathbf{h}||_{\mathbf{H}^s_x} \leq \varepsilon C(s)(||\mathbf{h}||_{\mathbf{H}^s_x} + ||\mathbf{u}||_{s+d+2\mathfrak{s}_0+4}||\mathbf{h}||_{\mathbf{H}^1_x}), \tag{4.2.81a}$$

*Proof.* $\mathcal{T}_1$ and $\mathcal{T}_4$ are multiplication operators then, it is enough to perform the proof on any component $(\mathcal{T}_i)_\sigma^{\sigma'}$, for $\sigma, \sigma' = \pm 1$ and $i = 1, 4$, that are simply multiplication operators from $H^s_x \to H^s_x$. One has

$$\begin{aligned} ||(\mathcal{T}_i)_\sigma^{\sigma'}(\varphi)h||_{H^s_x} &\overset{(A.5)}{\leq} C(s)(||(\mathcal{T}_i)_\sigma^{\sigma'}(\varphi)||_{H^s_x}||h||_{H^1_x} + ||(\mathcal{T}_i)_\sigma^{\sigma'}(\varphi)||_{H^1_x}||h||_{H^s_x}) \\ &\leq ||(\mathcal{T}_i)_\sigma^{\sigma'}||_{s+\mathfrak{s}_0}||h||_{H^s_x} + ||(\mathcal{T}_i)_\sigma^{\sigma'}||_{1+\mathfrak{s}_0}||h||_{H^s_x}) \\ &\overset{(4.2.38a)}{\leq} C(s)(||h||_{H^s_x} + ||\mathbf{u}||_{s+\mathfrak{s}_0+2}||h||_{H^1_x}), \end{aligned} \tag{4.2.82}$$

where we used also (4.2.32). In the same way one can show that

$$||((\mathcal{T}_i)_\sigma^{\sigma'}(\varphi, \cdot) - \mathbb{1})h||_s \leq \varepsilon C(s)(||h||_{H^s_x} + ||\mathbf{u}||_{s+\mathfrak{s}_0+2}||h||_{H^1_x}). \tag{4.2.83}$$

and hence the bound (5.1.102a) follow. Note that we used the simple fact that given a function $v \in H^s(\mathbb{T}^{d+1}; \mathbb{C})$ then $||v(\varphi)||_{H^s_x} \leq C||v||_{s+\mathfrak{s}_0}$. Now, for fixed $\varphi \in \mathbb{T}^d$ one has $\mathcal{T}_2(\varphi)h(x) := h(x + \xi(\varphi, x))$. We can bound, by using the (A.16a) on the change of variable $\mathbb{T} \to \mathbb{T}$, $x \to x + \xi(\varphi, x)$,

$$\begin{aligned} ||\mathcal{T}_2(\varphi)h||_{H^s_x} &\leq C(s)(||h||_{H^s_x} + |\xi(\varphi)|_{W^{s,\infty}(\mathbb{T})}||h||_{H^1_x}) \\ &\overset{(4.2.56a)}{\leq} C(s)(||h||_{H^s_x} + ||\mathbf{u}||_{s+\mathfrak{s}_0+2}||h||_{H^1_x}) \end{aligned} \tag{4.2.84}$$

where we have used also the fact $|\xi(\varphi)|_{W^{s,\infty}(\mathbb{T})} \leq |\xi|^\infty_{s+\mathfrak{s}_0}$. One can prove (5.1.102a) by using (A.16b), (4.2.32) and (4.2.56a). The estimates (5.1.102a) hold for $\mathcal{T}_2^{-1}(\varphi) : h(y) \to h(y + \widehat{\xi}(\varphi, y))$ thanks to the (4.2.58). □

Note that the fact that $\mathcal{T}_3$ maps $H^s_x \to H^s_x$ is trivial.

## 4.3 The diagonalization algorithm: KAM reduction

In this section we diagonalize the operator $\mathcal{L}_4$ in (4.2.33) in Section 4.2. In order to implement our procedure we pass to Fourier coefficients and introduce an "off diagonal decay norm" which is stronger that the standard operatorial one. We also define the reversibility properties of the operators, in terms of the Fourier coefficients.

Consider the bases $\{e_k = e^{i\ell\cdot\varphi}\sin jx : k = (\ell, j) \in \mathbb{Z}^d \times \mathbb{N}\}$ and $\{e_k = e^{i\ell\cdot\varphi}\cos jx : k = (\ell, j) \in \mathbb{Z}^d \times \mathbb{Z}_+\}$ for functions which are odd (resp. even) in $x$. Then any linear operator $A : \mathbf{G}^0_1 \to \mathbf{G}^0_2$, where $\mathbf{G}^0_{1,2} = \mathbf{X}^0, \mathbf{Y}^0, \mathbf{Z}^0$, can be represented by an infinite dimensional matrix

$$A := (A_i^{i'})_{i,i'\in \boldsymbol{C}\times\mathbb{Z}_+\times\mathbb{Z}^d},\ (A_\sigma^{\sigma'})_k^{k'} = (Ae_{k'}, e_k)_{L^2(\mathbb{T}^{d+1})},\ (A_\sigma^{\sigma'})u = \sum_{k,k'} (A_\sigma^{\sigma'})_k^{k'} u_{k'} e_k,$$

where $(\cdot,\cdot)_{L^2(\mathbb{T}^{d+1})}$ is the usual scalar product on $L^2$, we are denoting $i = (\sigma, k) = (\sigma, j, p) \in \boldsymbol{C} \times \mathbb{Z}_+ \times \mathbb{Z}^d$ and $\boldsymbol{C} := \{+1, -1\}$.

In the case of functions which are odd in $x$ we set the *extra* matrix coefficients (corresponding to $j = 0$) to zero.

**Definition 4.3.59.** *(**s-decay norm**). Given an infinite dimensional matrix $A := (A_i^{i'})_{i,i'\in\boldsymbol{C}\times\mathbb{Z}_+\times\mathbb{Z}^d}$ we define the norm of off-diagonal decay*

$$|A|_s^2 := \sup_{\sigma,\sigma'\in\boldsymbol{C}} |A_\sigma^{\sigma'}|_s^2 := \sup_{\sigma,\sigma'\in\boldsymbol{C}} \sum_{h\in\mathbb{Z}_+\times\mathbb{Z}^d} \langle h\rangle^{2s} \sup_{k-k'=h} |A_{\sigma,k}^{\sigma',k'}|^2 \tag{4.3.85}$$

*If one has that $A := A(\lambda)$ for $\lambda \in \Lambda \subset \mathbb{R}$, we define*

$$\begin{aligned} &|A|_s^{sup} := \sup_{\lambda\in\Lambda} |A(\lambda)|_s, \quad |A|_s^{lip} := \sup_{\lambda_1\neq\lambda_2} \frac{|A(\lambda_1) - A(\lambda_2)|_s}{|\lambda_1 - \lambda_2|}, \\ &|A|_{s,\gamma} := |A|_s^{sup} + \gamma |A|_s^{lip}. \end{aligned} \tag{4.3.86}$$

The decay norm we have introduced in (4.3.85) is suitable for the problem we are studying. Note that

$$\forall\, s \le s' \ \Rightarrow\ |A_\sigma^{\sigma'}|_s \le |A_\sigma^{\sigma'}|_{s'}.$$

Moreover norm (4.3.85) gives information on the polynomial off-diagonal decay of the matrices, indeed

$$|A_{\sigma,k}^{\sigma,k'}| \le \frac{|A_\sigma^{\sigma'}|_s}{\langle k - k'\rangle^s}, \quad \forall\, k, k' \in \mathbb{Z}_+ \times \mathbb{Z}^d,\ |A_i^i| \le |A|_0,\ |A_i^i|^{lip} \le |A|_0^{lip}. \tag{4.3.87}$$

We have the following important result:

**Theorem 4.3.60.** *Let $f \in C^q$ satisfy the Hypotheses of Proposition 4.0.46 with $q > \eta_1 + \beta + \mathfrak{s}_0$ where $\eta_1$ is defined in (4.2.31) and $\beta = 7\tau + 5$ for some $\tau > d$. Let $\gamma \in (0, \gamma_0)$, $\mathfrak{s}_0 \leq s \leq q - \eta_1 - \beta$ and $\mathbf{u}(\lambda) \in \mathbf{X}^0$ be a family of functions depending on a Lipschitz way on a parameter $\lambda \in \Lambda_o \subseteq \Lambda : [1/2, 3/2]$. Assume that*

$$||\mathbf{u}||_{\mathfrak{s}_0+\eta_1+\beta,\Lambda_o,\gamma} \leq 1. \tag{4.3.88}$$

*Then there exist constants $\epsilon_0$, $C$, depending only on the data of the problem, such that, if $\varepsilon\gamma^{-1} \leq \epsilon_0$, then there exists a sequence of purely imaginary numbers as in Proposition 4.1.51, namely $\forall\ h = (\sigma, j) \in \boldsymbol{C} \times \mathbb{N}$, and $\forall\ \lambda \in \Lambda$,*

$$\mu_h^\infty := \mu_{\sigma,j}^\infty(\lambda) := \mu_{\sigma,j}^\infty(\lambda, \mathbf{u}) = -\sigma i m j^2 + r_{\sigma,j}^\infty, \tag{4.3.89}$$

*where $m$ is defined in (6.4.106) with*

$$|r_{\sigma,j}^\infty|_\gamma \leq \varepsilon C, \quad \forall\ \sigma \in \boldsymbol{C},\ j \in \mathbb{N}. \tag{4.3.90}$$

*and such that, for any $\lambda \in \Lambda_\infty^{2\gamma}(\mathbf{u})$, defined in (4.1.23), there exists a bounded, invertible linear operator $\Phi_\infty(\lambda) : \mathbf{H}^s \to \mathbf{H}^s$, with bounded inverse $\Phi_\infty^{-1}(\lambda)$, such that*

$$\begin{aligned} \mathcal{L}_\infty(\lambda) &:= \Phi_\infty^{-1}(\lambda) \circ \mathcal{L}_4 \circ \Phi_\infty(\lambda) = \lambda\bar{\omega} \cdot \partial_\varphi \mathbb{1} + i\mathcal{D}_\infty, \\ \text{where} \qquad \mathcal{D}_\infty &:= \mathrm{diag}_{\mathrm{h} \in \mathrm{C} \times \mathbb{N}}\{\mu_h^\infty(\lambda)\}, \end{aligned} \tag{4.3.91}$$

*with $\mathcal{L}_4$ defined in* (4.2.33). *Moreover, the transformations $\Phi_\infty(\lambda)$, $\Phi_\infty^{-1}$ satisfy*

$$|\Phi_\infty(\lambda) - \mathbb{1}|_{s,\Lambda_\infty^{2\gamma},\gamma} + |\Phi_\infty^{-1}(\lambda) - \mathbb{1}|_{s,\Lambda_\infty^{2\gamma},\gamma} \leq \varepsilon\gamma^{-1} C(s)(1 + ||\mathbf{u}||_{s+\eta_1+\beta,\Lambda_o,\gamma}). \tag{4.3.92}$$

*In addition to this $\Phi_\infty$ defines, for any $\varphi \in \mathbb{T}^d$, the operator $\Phi_\infty(\varphi)$ (see Remark 2.2.10) which is an invertible operator of the phase space $\mathbf{X}_x^s := \mathbf{X}^s(\mathbb{T})$, for any $\mathfrak{s}_0 \leq s \leq q - \eta_1 - \beta$, with inverse $(\Phi_\infty(\varphi))^{-1} := \Phi_\infty^{-1}(\varphi)$ and*

$$||(\Phi_\infty^{\pm 1}(\varphi) - \mathbb{1})\mathbf{h}||_{\mathbf{H}_x^s} \leq \varepsilon\gamma^{-1} C(s)(||\mathbf{h}||_{\mathbf{H}_x^s} + ||\mathbf{u}||_{s+\eta_1+\beta+\mathfrak{s}_0}||\mathbf{h}||_{\mathbf{H}_x^1}). \tag{4.3.93}$$

**Remark 4.3.61.** *It is important to note that thanks to Reversibility Hypothesis 1, the operator $\mathcal{L}_\infty : \mathbf{X}^0 \to \mathbf{Z}^0$ i.e. it is reversible.*

The main point of the Theorem 4.3.60 is that the bound on the low norm of $u$ in (4.3.88) guarantees the bound on *higher* norms (4.3.92) for the transformations $\Phi_\infty^{\pm 1}$. This is fundamental in order to get the estimates on the inverse of $\mathcal{L}$ in high norms.

Moreover, the definition (4.1.23) of the set where the second Mel'nikov conditions hold, depends only on the final eigenvalues. Usually in KAM theorems, the non-resonance conditions have to be checked, inductively, at each step of the algorithm. This formulation, on the contrary, allow us to discuss the measure estimates only once. Indeed, the functions $\mu_h(\lambda)$ are well-defined even if $\Lambda_\infty = \emptyset$, so that, we will perform the measure estimates as the last step of the proof of Theorem 1.1.1.

### 4.3.1 Functional setting and notations

**The off-diagonal decay norm** Here we want to show some important properties of the norm $|\cdot|_s$. Clearly the same results hold for the norm $|\cdot|_{\mathbf{H}^s} := |\cdot|_{H^s\times H^s}$. Moreover we will introduce some characterization of the operators we have to deal with during the diagonalization procedure.

First of all we have following classical results.

**Lemma 4.3.62. Interpolation.** *For all $s \geq s_0 > (d+1)/2$ there are $C(s) \geq C(s_0) \geq 1$ such that if $A = A(\lambda)$ and $B = B(\lambda)$ depend on the parameter $\lambda \in \Lambda \subset \mathbb{R}$ in a Lipschitz way, then*

$$|AB|_{s,\gamma} \leq C(s)|A|_{s_0,\gamma}|B|_{s,\gamma} + C(s_0)|A|_{s,\gamma}|B|_{s_0,\gamma}, \tag{4.3.94a}$$

$$|AB|_{s,\gamma} \leq C(s)|A|_{s,\gamma}|B|_{s,\gamma}. \tag{4.3.94b}$$

$$||Ah||_{s,\gamma} \leq C(s)(|A|_{s_0,\gamma}||h||_{s,\gamma} + |A|_{s,\gamma}||h||_{s_0,\gamma}), \tag{4.3.94c}$$

Lemma 4.3.62 implies that for any $n \geq 0$ one has $\forall\ s \geq s_0$

$$|A^n|_{s_0,\gamma} \leq [C(s_0)]^{n-1}|A|^n_{s_0,\gamma}, \quad |A^n|_{s,\gamma} \leq n[C(s_0)]^{n-1}C(s)|A|_{s,\gamma}. \tag{4.3.95}$$

The following Lemma shows how to invert linear operators which are "*near*" to the identity in norm $|\cdot|_{\mathfrak{s}_0}$.

**Lemma 4.3.63.** *Let $C(s_0)$ be as in Lemma 4.3.62. Consider an operator of the form $\Phi = \mathbb{1} + \Psi$ where $\Psi = \Psi(\lambda)$ depends in a Lipschitz way on $\lambda \in \Lambda \subset \mathbb{R}$. Assume that $C(s_0)|\Psi|_{s_0,\gamma} \leq 1/2$. Then $\Phi$ is invertible and, for all $s \geq s_0 \geq (d+1)/2$,*

$$|\Phi^{-1}|_{s_0,\gamma} \leq 2, \quad |\Phi^{-1} - \mathbb{1}|_{s,\gamma} \leq C(s)|\Psi|_{s,\gamma} \tag{4.3.96}$$

*Moreover, if one has $\Phi_i = \mathbb{1} + \Psi_i$, $i = 1, 2$ such that $C(s_0)|\Psi_i|_{s_0,\gamma} \leq 1/2$, then*

$$|\Phi_2^{-1} - \Phi_1^{-1}|_{s,\gamma} \leq C(s)\left(|\Psi_2 - \Psi_1|_{s,\gamma} + (|\Psi_1|_{s,\gamma} + |\Psi_2|_{s,\gamma})|\Psi_2 - \Psi_1|_{s_0,\gamma}\right). \tag{4.3.97}$$

*Proof.* One has that $(\mathbb{1} + \Psi)^{-1} = \sum_{k\geq 0} \frac{(-1)^k}{k!}\Psi^k$, then by (4.3.95) we get bounds (4.3.96). Now, we can note that

$$\begin{aligned}
|\Phi_2^{-1} - \Phi_1^{-1}|_{s,\gamma} &= |\Phi_1^{-1}(\Psi_1 - \Psi_2)\Phi_2^{-1}|_{s,\gamma} \overset{(4.3.94a)}{\leq} C(s)|\Phi_1^{-1}|_{s_0,\gamma}|\Psi_1 - \Psi_2|_{s_0,\gamma}|\Phi_2^{-1}|_{s,\gamma} \\
&\quad + C(s)|\Phi_1^{-1}|_{s_0,\gamma}|\Psi_1 - \Psi_2|_{s,\gamma}|\Phi_2^{-1}|_{s_0,\gamma} + C(s)|\Phi_1^{-1}|_{s,\gamma}|\Psi_1 - \Psi_2|_{s_0,\gamma}|\Phi_2^{-1}|_{s_0,\gamma} \\
&\overset{(4.3.96)}{\leq} C(s)(|\Psi_1 - \Psi_2|_{s,\gamma} + (|\Psi_1|_{s,\gamma} + |\Psi_2|_{s,\gamma})|\Psi_1 - \Psi_2|_{s_0,\gamma})
\end{aligned}$$

that is the (4.3.97). □

**Töpliz-in-time matrices** We study now the special class of operators introduced in Remark 2.2.10, the so-called *Töpliz in time* matrices, i.e.

$$A_i^{i'} = A_{(\sigma,j,p)}^{(\sigma',j',p')} := A_{\sigma,j}^{\sigma' j'}(p-p'), \quad \text{for} \quad i,i' \in \boldsymbol{C} \times \mathbb{Z}_+ \times \mathbb{Z}^d. \tag{4.3.98}$$

To simplify the notation in this case, we shall write $A_i^{i'} = A_k^{k'}(\ell)$, $i = (k,p) = (\sigma,j,p) \in \boldsymbol{C} \times \mathbb{Z}_+ \times \mathbb{Z}^d$, $i' = (k',p') = (\sigma',j',p') \in \boldsymbol{C} \times \mathbb{Z}_+ \times \mathbb{Z}^d$, with $k,k' \in \boldsymbol{C} \times \mathbb{Z}_+$.

They are relevant because one can identify the matrix $A$ with a one-parameter family of operators, acting on the space $\mathbf{H}_x^s$, which depend on the time, namely

$$A(\varphi) := (A_{\sigma,j}^{\sigma',j'}(\varphi))_{\substack{\sigma,\sigma' \in C \\ j,j' \in \mathbb{Z}_+}}, \quad A_{\sigma,j}^{\sigma',j'}(\varphi) := \sum_{\ell \in \mathbb{Z}^d} A_{\sigma,j}^{\sigma',j'}(\ell) e^{i\ell\cdot\varphi}. \tag{4.3.99}$$

To obtain the stability result on the solutions we will strongly use this property.

**Lemma 4.3.64.** *If $A$ is a Töpliz in time matrix as in (4.3.98), and $\mathfrak{s}_0 := (d+2)/2$, then one has*

$$|A(\varphi)|_s \le C(\mathfrak{s}_0)|A|_{s+\mathfrak{s}_0}, \quad \forall\, \varphi \in \mathbb{T}^d. \tag{4.3.100}$$

*Proof.* We can note that, for any $\varphi \in \mathbb{T}^d$,

$$\begin{aligned} |A(\varphi)|_s^2 &:= \sup_{\sigma,\sigma' \in \mathbb{C}} \sum_{h \in \mathbb{Z}_+} \langle h \rangle^{2s} \sup_{j-j'=h} |A_{\sigma,j}^{\sigma',j'}(\varphi)|^2 \\ &\le C(\mathfrak{s}_0) \sup_{\sigma,\sigma' \in \boldsymbol{C}} \sum_{h \in \mathbb{Z}_+} \langle h \rangle^{2s} \sup_{j-j'=h} \sum_{\ell \in \mathbb{Z}^d} |A_{\sigma,j}^{\sigma',j'}(\ell)|^2 \langle \ell \rangle^{2\mathfrak{s}_0} \\ &\le C(\mathfrak{s}_0) \sup_{\sigma.\sigma' \in \boldsymbol{C}} \sum_{h \in \mathbb{Z}_+} \sup_{j-j'=h} |A_{\sigma,j}^{\sigma;j}(\ell)|^2 \langle \ell, h \rangle^{2(s+\mathfrak{s}_0)} \\ &\le C(\mathfrak{s}_0) \sup_{\sigma,\sigma' \in \boldsymbol{C}} \sum_{\substack{h \in \mathbb{Z}_+ \\ \ell \in \mathbb{Z}^d}} \sup_{j-j'=h} |A_{\sigma,j}^{\sigma',j'}(\ell)|^2 \langle \ell, h \rangle^{2(s+\mathfrak{s}_0)} \\ &\overset{(4.3.85)}{\le} C(\mathfrak{s}_0)|A|_{s+\mathfrak{s}_0}^2, \end{aligned} \tag{4.3.101}$$

that is the assertion. □

**Definition 4.3.65. (Smoothing operator)** *Given $N \in \mathbb{N}$, we the define the* smoothing operator $\Pi_N$ *as*

$$(\Pi_N A)_{\sigma,j,\ell}^{\sigma',j',\ell'} = \begin{cases} A_{\sigma,j,\ell}^{\sigma',j',\ell'}, & |\ell-\ell'| \le N, \\ 0 \quad \text{otherwise} \end{cases} \tag{4.3.102}$$

**Lemma 4.3.66.** *Let* $\Pi_N^\perp := \mathbb{1} - \Pi_N$,

*if* $A = A(\lambda)$ *is a Lipschitz family* $\lambda \in \Lambda$, *then*

$$|\Pi_N^\perp A|_{s,\gamma} \le N^{-\beta}|A|_{s+\beta,\gamma}, \quad \beta \ge 0. \tag{4.3.103}$$

*Proof.* Note that one has,

$$\begin{aligned} |\Pi_N^\perp A|_s^2 &= N^{-2\beta} \sup_{\sigma,\sigma' \in \boldsymbol{C}} \sum_{\substack{h \in \mathbb{Z}_+ \\ |\ell| > N}} \sup_{j-j'=h} |A_{\sigma,j}^{\sigma',j'}(\ell)|^2 \langle \ell, h\rangle^{2s} N^{2\beta} \\ &\le N^{-2\beta} \sup_{\sigma,\sigma' \in \boldsymbol{C}} \sum_{\substack{h \in \mathbb{Z}_+ \\ |\ell| > N}} \sup_{j-j'=h} |A_{\sigma,j}^{\sigma',j'}(\ell)|^2 \langle \ell, h\rangle^{2(s+\beta)} \\ &\le N^{-2\beta}|A|_{s+\beta}^2, \end{aligned} \tag{4.3.104}$$

The estimate on the Lipschitz norm follows similarly. □

**Remark 4.3.67. (Multiplication operator)** *We have already seen that if the decay norm is finite then the operator has a "good" off-diagonal decay. Although this property is strictly stronger than just being bounded, this class contains many useful operators in particular multiplication ones. Indeed, let* $\mathcal{T}_a : G_1^s \to G_2^s$, *where* $G_{1,2}^s = X^s, Y^s, Z^s$, *be the multiplication operator by a function* $a \in G^s$ *with* $G^s = X^s, Y^s, Z^s$, *i.e* $(\mathcal{T}_a h) = ah$. *Then one can check, in coordinates, that it is represented by the matrix* $T$ *such that*

$$|T|_s \le ||a||_s. \tag{4.3.105a}$$

*Moreover, if* $a = a(\lambda)$ *is a Lipschitz family of functions,*

$$|T|_{s,\gamma} \le ||a||_{s,\gamma}. \tag{4.3.106}$$

*At the beginning of our algorithm we actually deal with multiplication operators, so that one could try to control the operator by using only the Sobolev norms of functions. Unfortunately, it is not possible since the class of multiplication operators is not closed under our algorithm. This is the reason why we have introduced the decay norms that control decay in more general situations.*

**Matrix representation** In this paragraph we give a characterizations of reversible operators in the Fourier space. We need it to deal with a more general class of operators than the multiplication operators.

**Lemma 4.3.68.** *We have that, for* $G^s = X^s, Y^s, Z^s$,

$$R : G^s \to G^s \quad \Leftrightarrow \quad R_j^{j'}(\ell) = \overline{R_j^{j'}(\ell)},\ \forall \ell \in \mathbb{Z}^d,\ \forall\, j, j' \in \mathbb{Z}_+. \tag{4.3.107}$$

*Moreover,*

$$R : X^s \to Z^s \quad \Rightarrow R_j^{j'}(\ell) = -\overline{R_j^{j'}(\ell)},\ \forall \ell \in \mathbb{Z}^d,\ \forall\, j, j' \geq 1. \tag{4.3.108}$$

*Proof.* One can consider a function $a(\varphi, x) \in G^s$ where $G^s = X^s, Y^s, Z^s$, and develop it in a suitable basis $e_{\ell,j}$, $(\ell, j) \in \mathbb{Z}^d \times \mathbb{Z}_+$ (to fix the idea we can think $e_{\ell,j} = e^{i\ell\varphi} \sin jx$, that is the correct basis for $X^s$). One has that the coefficients of the function $a$ satisfies $a_j(\ell) = \overline{a_j(\ell)}$ for $G^s = X^s, Y^s$ while $a_j(\ell) = -\overline{a_j(\ell)}$ if $G^s = Z^s$. Then (4.3.107) and (4.3.108) follow by applying the definitions of reversibility or reversibility preserving in (4.1.7) and (4.1.8). □

**Lemma 4.3.69.** *Consider operators $A : \mathbf{G}^s \to \mathbf{G}^s$ with $G^s = X^s, Y^s, Z^s$ of the form $A := (A_\sigma^{\sigma'})_{\sigma,\sigma'=\pm 1}$, then*

$$\begin{pmatrix} A_1^1 & A_{-1}^1 \\ A_{-1}^1 & A_{-1}^{-1} \end{pmatrix} \begin{pmatrix} u \\ \bar{u} \end{pmatrix} = \begin{pmatrix} w \\ \bar{w} \end{pmatrix} \in \mathbf{G^s}, \quad \text{for any } (u, \bar{u}) \in \mathbf{G}^s \tag{4.3.109}$$

*if and only if* $\forall\ \sigma, \sigma' = \pm 1,\ \ell \in \mathbb{Z}^d,\ j, j' \in \mathbb{Z}_+$

$$A_{\sigma,j}^{\sigma',j'}(\ell) = \overline{A_{\sigma,j}^{\sigma',j'}(\ell)}, \quad \text{and} \quad \overline{A_{\sigma,j}^{\sigma',j'}(-\ell)} = A_{-\sigma,j}^{-\sigma',j'}(\ell), \tag{4.3.110}$$

*An operator $B : \mathbf{X}^s \to \mathbf{Z}^s$ if and only if* $\forall\ \sigma, \sigma' = \pm 1,\ \ell \in \mathbb{Z}^d,\ j, j' \geq 1$

$$B_{\sigma,j}^{\sigma',j'}(\ell) = -\overline{B_{\sigma,j}^{\sigma',j'}(\ell)}, \quad \text{and} \quad \overline{B_{\sigma,j}^{\sigma',j'}(-\ell)} = B_{-\sigma,j}^{-\sigma',j'}(\ell). \tag{4.3.111}$$

*Proof.* Lemma 4.3.68 implies only that $A_{\sigma,j}^{\sigma',j'}(\ell) = \overline{A_{\sigma,j}^{\sigma',j'}(\ell)}$. Since we need that the complex conjugate of the first component of $A\mathbf{u}$, with $\mathbf{u} \in \mathbf{G}^s$, is equal to the second one, the components of $A$ have to satisfy

$$\overline{A_{\sigma,j}^{\sigma',j'}(-\ell)} = A_{-\sigma,j}^{-\sigma',j'}(\ell), \quad \forall \sigma, \sigma' = \pm 1, \ell \in \mathbb{Z}^d, j, j' \in \mathbb{Z}_+. \tag{4.3.112}$$

In this case we say that the operator $A : \mathbf{G}^s \to \mathbf{G}^s$ is *reversibility-preserving.*

Following the same reasoning we have that for *reversible* operators the (4.3.111) hold. □

### 4.3.2 Reduction Algorithm

We prove Theorem 4.3.60 by means of the following Iterative Lemma on the class of linear operators

**Definition 4.3.70.**

$$\omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D} + \mathcal{R} : \ \mathbf{X}^0 \to \mathbf{Z}^0, \tag{4.3.113}$$

*where* $\omega = \lambda \bar{\omega}$, *and*

$$\mathcal{D} = (-i\sigma m(\lambda, \mathbf{u}(\lambda))D^2)_{\sigma=\pm 1}, \qquad \mathcal{R} = E_1 D + E_0 \tag{4.3.114}$$

*with* $D := \operatorname{diag}_{j\in\mathbb{N}}\{j\}$, *and where, if we write* $k = (\sigma, j, p) \in \boldsymbol{C} \times \mathbb{N} \times \mathbb{Z}^d$, $q = 0, 1$,

$$\begin{aligned} E_q = \Big((E_q)_k^{k'}\Big)_{k,k'\in \boldsymbol{C}\times\mathbb{N}\times\mathbb{Z}^d} &= \Big((E_q)_{\sigma,j}^{\sigma',j'}(p-p')\Big)_{k,k'\in \boldsymbol{C}\times\mathbb{N}\times\mathbb{Z}^d}, \\ (E_1)_{\sigma,j}^{\sigma,j'}(p-p') \equiv 0, \quad &\forall\, j, j' \in \mathbb{N}, \ \ p, p' \in \mathbb{Z}^d. \end{aligned} \tag{4.3.115}$$

Now given $\mathbf{u}(\lambda)$ defined on $\Lambda_o$, we note that the operator $\mathcal{L}_4$ constructed in Lemma 4.2.53 has the form (4.3.113) and satisfies the (4.3.114) and (4.3.115) as well as the estimates (4.2.36a) and (4.2.36b). Note that each component $(E_q)_\sigma^{\sigma'}$, $q = 0, 1$, represents the matrix of the multiplication operator by a function. This fact is not necessary for our analysis, and it cannot be preserved during the algorithm.

Define

$$N_{-1} := 1, \quad N_\nu := N_{\nu-1}^{\chi} = N_0^{\chi^\nu}, \ \ \forall\, \nu \geq 0, \ \ \chi = \frac{3}{2}. \tag{4.3.116}$$

and

$$\alpha = 7\tau + 3, \qquad \eta_2 := \eta_1 + \beta, \tag{4.3.117}$$

where $\eta_1$ is defined in (4.2.31) and $\beta = 7\tau + 5$. Fix $\mathcal{L}_4 = \mathcal{L}_0 = \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}_0 + \mathcal{R}_0$ with $\mathcal{R}_0 = E_1^0 D + E_0^0$, we define

$$\delta_s^0 := |E_1^0|_{s,\gamma} + |E_0^0|_{s,\gamma}, \quad \text{for} \quad s \geq 0. \tag{4.3.118}$$

**Lemma 4.3.71 (KAM iteration).** *Let* $q > \eta_1 + \mathfrak{s}_0 + \beta$. *There exist constant* $C_0 > 0$, $N_0 \in \mathbb{N}$ *large, such that if*

$$N_0^{C_0} \gamma^{-1} \delta^0_{\mathfrak{s}_0+\beta} \leq 1, \tag{4.3.119}$$

*then, for any* $\nu \geq 0$, *one has:*
$(\mathbf{S1}.)_\nu$ *Set* $\Lambda_0^\gamma := \Lambda_o$ *and for* $\nu \geq 1$

$$\Lambda_\nu^\gamma := \left\{ \begin{array}{cc} \lambda \in \Lambda_{\nu-1}^\gamma : & |\omega\cdot\ell + \mu_h^{\nu-1}(\lambda) - \mu_{h'}^{\nu-1}(\lambda)| \geq \frac{\gamma|\sigma j^2 - \sigma' j'^2|}{\langle \ell\rangle^\tau}, \\ & \forall |\ell| \leq N_{\nu-1}, h, h' \in \boldsymbol{C}\times\mathbb{N} \end{array} \right\}, \tag{4.3.120}$$

*For any* $\lambda \in \Lambda_\nu^\gamma := \Lambda_\nu^\gamma(\mathbf{u})$, *there exists an invertible map* $\Phi_{\nu-1}$ *of the form* $\Phi_{-1} = \mathbb{1}$ *and for* $\nu \geq 1$, $\Phi_{\nu-1} := \mathbb{1} + \Psi_{\nu-1} : \mathbf{H}^s \to \mathbf{H}^s$, *with the following properties.*

*The maps* $\Phi_{\nu-1}$, $\Phi_{\nu-1}^{-1}$ *are reversibility-preserving according to Definition 4.1.47, moreover* $\Psi_{\nu-1}$ *is Töplitz in time,* $\Psi_{\nu-1} := \Psi_{\nu-1}(\varphi)$ *(see (4.3.98)) and satisfies the bounds:*

$$|\Psi_{\nu-1}|_{s,\gamma} \leq \delta^0_{s+\beta} N_{\nu-1}^{2\tau+1} N_{\nu-2}^{-\alpha}, \tag{4.3.121}$$

*Setting, for* $\nu \geq 1$, $\mathcal{L}_\nu := \Phi_{\nu-1}^{-1}\mathcal{L}_{\nu-1}\Phi_{\nu-1}$, *we have:*

$$\begin{aligned} \mathcal{L}_\nu &= \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}_\nu + \mathcal{R}_\nu, \qquad \mathcal{D}_\nu = \mathrm{diag}_{h\in \boldsymbol{C}\times\mathbb{N}}\{\mu_h^\nu\}, \\ \mu_h^\nu(\lambda) &= \mu_{\sigma,j}^\nu = \mu_{\sigma,j}^0(\lambda) + r_{\sigma,j}^\nu(\lambda), \quad \mu_{\sigma,j}^0(0) = -\sigma i\, m(\lambda, \mathbf{u}(\lambda)) j^2, \end{aligned} \tag{4.3.122}$$

*and*

$$\mathcal{R}_\nu = E_1^\nu(\lambda) D + E_0^\nu(\lambda), \tag{4.3.123}$$

*where* $\mathcal{R}_\nu$ *is reversible and the matrices* $E_q^\nu$ *satisfy (4.3.115) for* $q = 1, 2$. *For* $\nu \geq 0$ *one has* $r_h^\nu \in i\mathbb{R}$, $r_{\sigma,j}^\nu = -r_{-\sigma,j}^\nu$ *and the following bound holds:*

$$|r_h^\nu|_\gamma := |r_h^\nu|_{\Lambda_\nu^\gamma,\gamma} \leq \varepsilon C. \tag{4.3.124}$$

*Finally, if we define*

$$\delta_s^\nu := |E_1^\nu|_{s,\gamma} + |E_0^\nu|_{s,\gamma}, \quad \forall s \geq 0, \tag{4.3.125}$$

*one has* $\forall\, s \in [\mathfrak{s}_0, q - \eta_1 - \beta]$ *(*$\alpha$ *is defined in* (4.3.117)*) and* $\nu \geq 0$

$$\begin{aligned} \delta_s^\nu &\leq \delta_{s+\beta}^0 N_{\nu-1}^{-\alpha}, \\ \delta_{s+\beta}^\nu &\leq \delta_{s+\beta}^0 N_{\nu-1}. \end{aligned} \tag{4.3.126}$$

$(\mathbf{S2})_\nu$ *For all* $j \in \mathbb{N}$ *there exists Lipschitz extensions* $\widetilde{\mu}_h^\nu(\cdot) : \Lambda \to \mathbb{R}$ *of* $\mu_h^\nu(\cdot) : \Lambda_\nu^\gamma \to \mathbb{R}$, *such that for* $\nu \geq 1$,

$$|\widetilde{\mu}_h^\nu - \widetilde{\mu}_h^{\nu-1}|_\gamma \leq \delta_{\mathfrak{s}_0}^{\nu-1}, \qquad \forall\, k \in \boldsymbol{C} \times \mathbb{N}. \tag{4.3.127}$$

$(\mathbf{S3})_\nu$ *Let* $\mathbf{u}_1(\lambda)$, $\mathbf{u}_2(\lambda)$ *be Lipschitz families of Sobolev functions, defined for* $\lambda \in \Lambda_o$ *such that (4.3.88), (4.3.119) hold with* $\mathcal{R}_0 = \mathcal{R}_0(\mathbf{u}_i)$ *with* $i = 1, 2$. *Then for* $\nu \geq 0$, *for any* $\lambda \in \Lambda_\nu^{\gamma_1}(\mathbf{u}_1) \cap \Lambda_\nu^{\gamma_2}(\mathbf{u}_2)$, *with* $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$, *one has*

$$|E_1^\nu(\mathbf{u}_1) - E_1^\nu(\mathbf{u}_2)|_{\mathfrak{s}_0} + |E_0^\nu(\mathbf{u}_1) - E_0^\nu(\mathbf{u}_2)|_{\mathfrak{s}_0} \leq \varepsilon N_{\nu-1}^{-\alpha} ||\mathbf{u}_1 - \mathbf{u}_2||_{\mathfrak{s}_0+\eta_2}, \tag{4.3.128a}$$

$$|E_1^\nu(\mathbf{u}_1) - E_1^\nu(\mathbf{u}_2)|_{\mathfrak{s}_0+\beta} + |E_0^\nu(\mathbf{u}_1) - E_0^\nu(\mathbf{u}_2)|_{\mathfrak{s}_0+\beta} \leq \varepsilon N_{\nu-1} ||\mathbf{u}_1 - \mathbf{u}_2||_{\mathfrak{s}_0+\eta_2} \tag{4.3.128b}$$

*and moreover, for* $\nu \geq 1$, *for any* $s \in [\mathfrak{s}_0, \mathfrak{s}_0 + \beta]$, *for any* $h \in \boldsymbol{C} \times \mathbb{N}$ *and for any* $\lambda \in \Lambda_\nu^{\gamma_1}(\mathbf{u}_1) \cap \Lambda_\nu^{\gamma_2}(\mathbf{u}_2)$,

$$|(r_h^\nu(\mathbf{u}_2) - r_h^\nu(\mathbf{u}_1)) - (r_h^{\nu-1}(\mathbf{u}_2) - r_h^{\nu-1}(\mathbf{u}_1))| \leq |E_0^{\nu-1}(\mathbf{u}_1) - E_0^{\nu-1}(\mathbf{u}_2)|_{\mathfrak{s}_0}, \tag{4.3.129a}$$

$$|(r_h^\nu(\mathbf{u}_2) - r_h^\nu(\mathbf{u}_1))| \leq \varepsilon C ||\mathbf{u}_1 - \mathbf{u}_2||_{\mathfrak{s}_0+\eta_2}. \tag{4.3.129b}$$

$(\mathbf{S4})_\nu$ *Let* $u_1, u_2$ *be as in* $(\mathbf{S3})_\nu$ *and* $0 < \rho < \gamma/2$. *For any* $\nu \geq 0$ *one has*

$$\varepsilon C N_{\nu-1}^\tau \sup_{\lambda\in\Lambda_o} ||\mathbf{u}_1 - \mathbf{u}_2||_{\mathfrak{s}_0+\eta_2} \leq \rho \quad \Rightarrow \quad \Lambda_\nu^\gamma(\mathbf{u}_1) \subset \Lambda_\nu^{\gamma-\rho}(\mathbf{u}_2), \tag{4.3.130}$$

*Proof.* We start by proving that $(\mathbf{Si})_{\mathbf{0}}$ hold for $i = 1, \dots, 4$.

$(\mathbf{S1})_{\mathbf{0}}$. Clearly the properties (4.3.124)-(4.3.126) hold by (4.3.113), (4.3.114) and the form of $\mu_k^0$ in (4.3.122), recall that $r_k^0 = 0$ . Moreover, $m$ real implies that $\mu_k^0$ are imaginary. In addition to this, our hypotheses guarantee that $\mathcal{R}_0 = E_1^0 \partial_x + E_0^0$ and $\mathcal{L}_0$ are reversible operators.

$(\mathbf{S2})_{\mathbf{0}}$. We have to extend the eigenvalues $\mu_k^0$ from the set $\Lambda_0^\gamma$ to the entire $\Lambda$. Namely we extend the function $m(\lambda)$ to a $\widetilde{m}(\lambda)$ that is Lipschitz in $\Lambda$, with the same sup norm and Lipschitz semi-norm, by Kirszbraun theorem.

$(\mathbf{S3})_{\mathbf{0}}$. It holds by (4.2.36b) for $\mathfrak{s}_0$, $\mathfrak{s}_0 + \beta$ using (4.3.88) and (4.3.117).

$(\mathbf{S4})_{\mathbf{0}}$. By definition one has $\Lambda_0^\gamma(\mathbf{u}_1) = \Lambda_o = \Lambda_0^{\gamma-\rho}(\mathbf{u}_2)$, then the (4.3.130) follows trivially.

**KAM step** In this Section we show in detail one step of the KAM iteration. In other words we show how to define the transformation $\Phi_\nu$ which conjugates $\mathcal{L}_\nu$ to $\mathcal{L}_{\nu+1}$. For simplicity we shall avoid to write the index, but we will only write $+$ instead of $\nu + 1$.

We consider a transformation of the form $\Phi = \mathbb{1} + \Psi$, with $\Psi := (\Psi_\sigma^{\sigma'})_{\sigma,\sigma'=\pm 1}$, acting on the operator

$$\mathcal{L} = \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D} + \mathcal{R}$$

with $\mathcal{D}$ and $\mathcal{R}$ as in (4.3.122), (4.3.123). Then, $\forall\, \mathbf{h} \in \mathbf{H}^s$, one has

$$\begin{aligned} \mathcal{L}\Phi\mathbf{h} &= \omega \cdot \partial_\varphi(\Phi(\mathbf{h})) + \mathcal{D}\Phi\mathbf{h} + \mathcal{R}\Phi\mathbf{h} \\ &= \Phi\left(\omega \cdot \partial_\varphi \mathbf{h} + \mathcal{D}\mathbf{h}\right) + \left(\omega \cdot \partial_\varphi \Psi + [\mathcal{D}, \Psi] + \Pi_N \mathcal{R}\right)\mathbf{h} + \Big(\Pi_N^\perp \mathcal{R} + \mathcal{R}\Psi\Big)\,\mathbf{h}, \end{aligned} \tag{4.3.131}$$

where $[\mathcal{D}, \Phi] := \mathcal{D}\Phi - \Phi\mathcal{D}$, and $\Pi_N$ is defined in (4.3.102). The smoothing operator $\Pi_N$ is necessary for technical reasons: it will be used in order to obtain suitable estimates on the high norms of the transformation $\Phi$.

In the following Lemma we show how to solve the *homological* equation

$$\omega \cdot \partial_\varphi \Psi + [\mathcal{D}, \Psi] + \Pi_N \mathcal{R} = [\mathcal{R}], \quad [\mathcal{R}]_k^{k'} := \begin{cases} (E_0)_k^k = (E_0)_{\sigma,j}^{\sigma,j}(0), & k = k', \\ 0 \qquad k \neq k', \end{cases} \tag{4.3.132}$$

for $k, k' \in \boldsymbol{C} \times \mathbb{N} \times \mathbb{Z}^d$.

**Lemma 4.3.72 (Homological equation).** *For any $\lambda \in \Lambda_+^\gamma$ there exists a unique solution $\Psi = \Psi(\varphi)$ of the homological equation (4.3.132), such that*

$$|\Psi|_{s,\gamma} \leq C N^{2\tau+1} \gamma^{-1} \delta_s \tag{4.3.133}$$

*Moreover, for $\gamma/2 \leq \gamma_1, \gamma_2 \leq 2\gamma$, and if $\mathbf{u}_1(\lambda), \mathbf{u}_2(\lambda)$ are Lipschitz functions, then $\forall\ s \in [\mathfrak{s}_0, \mathfrak{s}_0 + \beta]$, $\lambda \in \Lambda_+^{\gamma_1}(\mathbf{u}_1) \cap \Lambda_+^{\gamma_2}(\mathbf{u}_2)$, one has*

$$|\Delta_{12}\Psi|_s \leq CN^{2\tau+1}\gamma^{-1}\Big((|E_1(\mathbf{u}_2)|_s + |E_0(\mathbf{u}_2)|_s)||\mathbf{u}_1 - \mathbf{u}_2||_{\mathfrak{s}_0+\eta_2} + \\ +|\Delta_{12}E_1|_s + |\Delta_{12}E_0|_s\Big), \tag{4.3.134}$$

*where we define $\Delta_{12}\Psi = \Psi(\mathbf{u}_1) - \Psi(\mathbf{u}_2)$.*
*Finally, one has $\Psi : \mathbf{X}^s \to \mathbf{X}^s$, i.e. the operator $\Psi$ is reversible preserving.*

*Proof.* On each component $k = (\sigma, j, p), k' = (\sigma', j', p') \in \boldsymbol{C} \times \mathbb{N} \times \mathbb{Z}^d$, the equation (4.3.132) reads

$$i\omega \cdot (p - p')\Psi_k^{k'} + \mathcal{D}_k^k \Psi_k^{k'} - \Psi_k^{k'}\mathcal{D}_{k'}^{k'} + \mathcal{R}_k^{k'} = [\mathcal{R}]_k^{k'}. \tag{4.3.135}$$

then, by defining

$$d_k^{k'} := i\omega \cdot (p - p') + \mu_{\sigma,j} - \mu_{\sigma',j'} \tag{4.3.136}$$

we get

$$\Psi_k^{k'} = \frac{-\mathcal{R}_k^{k'}}{d_k^{k'}}, \qquad k \neq k', \ \ |p - p'| \leq N, \tag{4.3.137}$$

and $\Psi_k^{k'} \equiv 0$ otherwise. Clearly the solution has the form $\Psi_{\sigma,j,p}^{\sigma',j',p'} = \Psi_{\sigma,j}^{\sigma',j'}(p - p')$ and hence we can define a time-dependent change of variables as $\Psi_{\sigma,j}^{\sigma',j'}(\varphi) = \sum_{\ell \in \mathbb{Z}^d} \Psi_{\sigma,j}^{\sigma',j'}(\ell) e^{i\ell\cdot\varphi}$.

Note that, by (4.3.120) and (1.1.2) one has for all $k \neq k' \in \boldsymbol{C} \times \mathbb{N} \times \mathbb{Z}^d$, setting $k = (\sigma, j, p)$, $k' = (\sigma', j', p')$ and $\ell = p - p'$

$$|d_k^{k'}| \geq \begin{cases} \dfrac{\gamma(j^2 + j'^2)}{\langle \ell \rangle^\tau}, & \sigma = -\sigma', \\ \dfrac{\gamma(j + j')}{\langle \ell \rangle^\tau}, & \text{if} \quad \sigma = \sigma' \ j \neq j', \\ \dfrac{\gamma}{\langle \ell \rangle^\tau}, & \text{if} \quad \sigma = \sigma' \ j = j' \ p \neq p' \end{cases} \tag{4.3.138}$$

This implies that, for $\sigma \neq \sigma'$, we have

$$|\Psi_k^{k'}| \leq \gamma^{-1}|\ell|^\tau \left(|(E_1)_k^{k'}| + |(E_0)_k^{k'}|\right) \frac{j}{j^2 + j'^2}, \tag{4.3.139}$$

while, for $\sigma = \sigma'$,

$$|\Psi_k^{k'}| \leq \begin{cases} \gamma^{-1}\langle \ell \rangle^\tau |(E_0)_k^{k'}| \dfrac{1}{j + j'}, & j \neq j', \\ \gamma^{-1}\langle \ell \rangle^\tau |(E_0)_k^{k'}|, & j = j', \end{cases} \tag{4.3.140}$$

and we can estimate the divisors $d_k^{k'}$ from below, hence, by the definition of the $s-$norm in (4.3.85) in any case we obtain the estimate

$$|\Psi|_s \leq \gamma^{-1} N^{\tau} \delta_s. \tag{4.3.141}$$

If we define the operator $A$ as

$$A_k^{k'} = A_{\sigma,j}^{\sigma',j'}(\ell) := \begin{cases} \Psi_{\sigma,j}^{\sigma,j}(\ell), & (\sigma,j) = (\sigma',j') \in \boldsymbol{C} \times \mathbb{N}, \;\; \ell \in \mathbb{Z}^d \backslash \{0\}, \\ 0, & \text{otherwise}, \end{cases} \tag{4.3.142}$$

we have proved the following Lemma.

**Lemma 4.3.73.** *The operator $\Psi - A$ is regularizing, indeed,*

$$|D(\Psi - A)|_s^2 := \sup_{\sigma,\sigma' \in \boldsymbol{C}} \sum_{\substack{k \in \mathbb{N}, \\ \ell \in \mathbb{Z}^d}} \sup_{\substack{j-j'=k \\ j \neq j'}} |\Psi_{\sigma,j}^{\sigma',j'}(\ell)|^2 \langle \ell, k \rangle^{2s}, \overset{(4.3.120)}{\leq_s} \gamma^{-2} N^{2\tau} \delta_s \tag{4.3.143}$$

*where $D$ is defined in the line above (4.3.115).*

This Lemma will be used in the study of the remainder of the conjugate operator. In particular we will use it to prove that the reminder is still in the class of operators described in (4.3.114).

Now we need a bound on the Lipschitz semi-norm of the transformation. Then, given $\lambda_1, \lambda_2 \in \Lambda_+^{\gamma}$, one has, for $k = (\sigma, j, p), k' = (\sigma', j', p') \in \boldsymbol{C} \times \mathbb{N} \times \mathbb{Z}^d$, and $\ell := p - p'$,

$$|\Psi_k^{k'}(\lambda_1) - \Psi_k^{k'}(\lambda_2)| \leq \frac{|\mathcal{R}_k^{k'}(\lambda_1) - \mathcal{R}_k^{k'}(\lambda_2)|}{|d_k^{k'}(\lambda_1)|} + |\mathcal{R}_k^{k'}(\lambda_2)| \frac{|d_k^{k'}(\lambda_1) - d_k^{k'}(\lambda_2)|}{|d_k^{k'}(\lambda_1)||d_k^{k'}(\lambda_2)|}. \tag{4.3.144}$$

Now, recall that $\omega = \lambda \bar{\omega}$, by using that $\gamma |m|^{lip} = \gamma |m-1|^{lip} \overset{(4.2.35a)}{\leq} \varepsilon C$, and by (4.3.124), we obtain

$$|d_k^{k'}(\lambda_1) - d_k^{k'}(\lambda_2)| \overset{(4.3.136),(4.3.122)}{\lessdot} |\lambda_1 - \lambda_2| \cdot \left(|\ell| + \varepsilon \gamma^{-1} |\sigma j^2 - \sigma' j'^2| + \varepsilon \gamma^{-1}\right). \tag{4.3.145}$$

Then, for $\sigma, \sigma' = \pm 1$, $j \neq j'$ and $\varepsilon \gamma^{-1} \leq 1$,

$$\frac{|d_k^{k'}(\lambda_1) - d_k^{k'}(\lambda_2)|}{|d_k^{k'}(\lambda_1)||d_k^{k'}(\lambda_2)|} \overset{(4.3.145),(4.3.120)}{\lessdot} |\lambda_1 - \lambda_2| \frac{N^{2\tau+1}\gamma^{-2}}{|\sigma j^2 - \sigma' j'^2|} \tag{4.3.146}$$

for $|\ell| \leq N$. Summarizing we have proved that, for any $|\ell| \leq N$, $j, j' \geq 1$, $j \neq j'$, $\sigma, \sigma' = \pm 1$

$$|\mathcal{R}_k^{k'}| / |\sigma j^2 - \sigma' j'^2| \leq |(E_1)_k^{k'}| + |(E_0)_k^{k'}|$$

for any $\lambda \in \Lambda_+^{\gamma}$. Moreover $|d_k^{k'}| \geq \gamma \cdot \langle \ell \rangle^{-\tau}$ for $\sigma = \sigma'$ and $j = j'$. We apply this two bounds in (4.3.144) together with (4.3.146) and (4.3.141). We get

$$|\Psi|_{s,\gamma} \leq \gamma^{-1} C N^{2\tau+1} \left(|E_1|_{s,\gamma} + |E_0|_{s,\gamma}\right), \tag{4.3.147}$$

and hence the (4.3.133) is proved.

Let us check the (4.3.134). For $\lambda \in \Lambda_+^{\gamma_1} \cap \Lambda_+^{\gamma_2}$, if $k = (\sigma, j, p) \neq (\sigma', j', p') = k'$, one has

$$\begin{aligned} |\Delta_{12}\Psi_k^{k'}| &\leq \frac{|\Delta_{12}\mathcal{R}_k^{k'}|}{|d_k^{k'}(\mathbf{u}_1)|} + |\mathcal{R}_k^{k'}(\mathbf{u}_2)| \frac{|\Delta_{12}d_k^{k'}|}{|d_k^{k'}(\mathbf{u}_1)||d_k^{k'}(\mathbf{u}_2)|} \\ &\overset{(4.2.35b),(4.3.129b)}{\leq} N^{2\tau}\gamma^{-1} \Big( |(\Delta_{12}E_1)_k^{k'}| + |(\Delta_{12}E_0)_k^{k'}| \\ &+ \Big( |(E_1)_k^{k'}(\mathbf{u}_2)| + |(E_0)_k^{k'}(\mathbf{u}_2)| \Big) ||\mathbf{u}_1 - \mathbf{u}_2||_{\mathfrak{s}_0+\eta_2} \Big) \end{aligned} \tag{4.3.148}$$

where we used $\varepsilon\gamma^{-1} \leq 1$ $\gamma_1^{-1}, \gamma_2^{-1} \leq \gamma^{-1}$, hence (4.3.148) implies the (4.3.134).

Since $\overline{\mu_{\sigma,j}} = -\mu_{\sigma,j}$ and the operator $\mathcal{R}$ is reversible (see (4.3.111)), by (4.3.137), we have that

$$\overline{\Psi_{\sigma,j}^{\sigma',j'}(\ell)} = \frac{-\overline{\mathcal{R}_{\sigma,j}^{\sigma',j'}(\ell)}}{-i\omega\cdot\ell + \overline{\mu_{\sigma',j'}} - \overline{\mu_{\sigma,j}}} = \frac{\mathcal{R}_{\sigma,j}^{\sigma',j'}(\ell)}{-(i\omega\cdot\ell + \mu_{\sigma',j'} - \mu_{\sigma,j})} = \Psi_{\sigma,j}^{\sigma',j'}(\ell), \tag{4.3.149}$$

so that, by Lemma 4.3.68, for any $\sigma, \sigma' = \pm 1$, the operators $\Psi_\sigma^{\sigma'}$ are reversibility preserving. In the same way, again thanks to the reversibility of $\mathcal{R}$, one can check $\overline{\Psi_{\sigma,j}^{\sigma',j'}(-\ell)} = \Psi_{-\sigma,j}^{-\sigma',j'}(\ell)$ which implies $\Psi : \mathbf{X}^s \to \mathbf{X}^s$, i.e. $\Psi$ is reversibility preserving. $\square$

By Lemma 4.3.63, for $\delta_{\mathfrak{s}_0}$ small enough, we have by (4.3.133) for $s = \mathfrak{s}_0$

$$C(\mathfrak{s}_0)|\Psi|_{\mathfrak{s}_0} \leq \frac{1}{2}, \tag{4.3.150}$$

then, the operator $\Phi = \mathcal{I} + \Psi$ is invertible. In this case we can conjugate the operator $\mathcal{L}$ to an operator $\mathcal{L}_+$ as shown in the next Lemma.

**Lemma 4.3.74 (The new operator $\mathcal{L}_+$).** *Consider the operator $\Phi = \mathcal{I} + \Psi$ with $\Psi$ defined in Lemma 4.3.72. Then, one has*

$$\mathcal{L}_+ := \Phi^{-1}\mathcal{L}\Phi := \omega\cdot\partial_\varphi\mathbb{1} + \mathcal{D}_+ + \mathcal{R}_+, \tag{4.3.151}$$

*where the diagonal operator $\mathcal{D}_+$ has the form*

$$\begin{aligned} &\mathcal{D}_+ := diag_{h\in C\times\mathbb{N}}\{\mu_h^+\}, \\ &\mu_h^+ := \mu_h + (E_0)_h^h(0) = \mu_h^0 + r_h + (E_0)_h^h =: \mu_h^0 + r_h^+, \end{aligned} \tag{4.3.152}$$

*with $h := (\sigma, j) \in \boldsymbol{C} \times \mathbb{N}$. The remainder has the form*

$$\mathcal{R}_+ := E_1^+ D + E_0^+, \tag{4.3.153}$$

*where $E_i^+$ are linear bounded operators of the form (4.3.115) for $i = 0, 1$. Moreover, the eigenvalues $\mu_h^+$ satisfy*

$$|\mu_h^+ - \mu_h|^{lip} = |r_h^+ - r_h|^{lip} = |(E_0)_h^h(0)|^{lip} \leq |E_0|_{\mathfrak{s}_0}^{lip}, \quad h \in \Sigma\times\mathbb{N}, \tag{4.3.154}$$

*while the remainder $\mathcal{R}_+$ satisfies*

$$\begin{aligned} \delta_s^+ &:= |E_1^+|_{s,\gamma} + |E_0^+|_{s,\gamma} \leq_s N^{-\beta}\delta_{s+\beta} + N^{2\tau+1}\gamma^{-1}\delta_s\delta_{\mathfrak{s}_0}, \\ \delta_{s+\beta}^+ &\leq_{s+\beta} \delta_{s+\beta} + N^{2\tau+1}\gamma^{-1}\delta_{s+\beta}\delta_{\mathfrak{s}_0}. \end{aligned} \tag{4.3.155}$$

*Finally, for $\gamma/2 \leq \gamma_1, \gamma_2 \leq 2\gamma$, and if $u_1(\lambda), u_2(\lambda)$ are Lipschitz functions, then $\forall\ s \in [\mathfrak{s}_0, \mathfrak{s}_0 + \beta]$, $\lambda \in \Lambda_+^{\gamma_1}(u_1) \cap \Lambda_+^{\gamma_2}(u_2)$, setting $|\Delta_{12}E_1|_s + |\Delta_{12}E_0|_s = \Delta_s$, we have:*

$$\begin{aligned} \Delta_s^+ &\leq |\Pi_N^{\perp}\Delta_{12}E_0|_s + |\Pi_N^{\perp}\Delta_{12}E_1|_s \\ &+ N^{2\tau+1}\gamma^{-1}(\delta_s(\mathbf{u}_1) + \delta_s(\mathbf{u}_2))(\delta_{\mathfrak{s}_0}(\mathbf{u}_1) + \delta_{\mathfrak{s}_0}(\mathbf{u}_2))||\mathbf{u}_1 - \mathbf{u}_2||_{s+\eta_2} \\ &+ N^{2\tau+1}\gamma^{-1}(\delta_s(\mathbf{u}_1) + \delta_s(\mathbf{u}_2))\Delta_{\mathfrak{s}_0} + N^{2\tau+1}\gamma^{-1}(\delta_{\mathfrak{s}_0}(\mathbf{u}_1) + \delta_{\mathfrak{s}_0}(\mathbf{u}_2))\Delta_s. \end{aligned} \tag{4.3.156}$$

*Proof.* The expression (4.3.152) follows by (4.3.132), the bound (4.3.154) follows by (4.3.87).

The bound (4.3.155) is more complicated. First of all we note that, by (4.3.131) and (4.3.132), we have

$$\mathcal{R}_+ := \Phi^{-1}\left(\Pi_N^{\perp}\mathcal{R} + \mathcal{R}\Psi - \Psi[\mathcal{R}]\right) := E_1^+ D + E_0^+, \tag{4.3.157}$$

where

$$\begin{aligned} E_1^+ &:= \Phi^{-1}\left(\Pi_N^{\perp}E_1 + E_1 A\right), \\ E_0^+ &:= \Phi^{-1}\left(\Pi_N^{\perp}E_0 + E_0\Psi - \Psi[\mathcal{R}] + E_1 D(\Psi - A)\right), \end{aligned} \tag{4.3.158}$$

where $A$ is defined in (4.3.142).

We can estimate the first of the (4.3.158) by

$$\begin{aligned} |E_1^+|_{s,\gamma} &\overset{(4.3.94a),(4.3.96)}{\leq_s} 2|\Pi_N^{\perp}E_1|_{s,\gamma} + (1 + |\Psi|_{s,\gamma})\left(|\Pi_N^{\perp}E_1|_{\mathfrak{s}_0,\gamma} + |E_1|_{\mathfrak{s}_0,\gamma}|A|_{\mathfrak{s}_0,\gamma}\right) \\ &\quad + 2(|E_1|_{s,\gamma}|A|_{\mathfrak{s}_0,\gamma} + |E_1|_{\mathfrak{s}_0,\gamma}|A|_{s,\gamma}) \\ &\overset{(4.3.133),(4.3.103)}{\leq_s} N^{-\beta}|E_1|_{s+\beta,\gamma} + N^{2\tau+1}\gamma^{-1}\delta_{\mathfrak{s}_0}\delta_s. \end{aligned} \tag{4.3.159}$$

The bound on $E_0^+$ is obtained in the same way by recalling that, by Lemma 4.3.73,

$$|D(\Psi - A)|_{s,\gamma} \leq \gamma^{-1}N^{2\tau+1}\delta_s. \tag{4.3.160}$$

The second bound in (4.3.155) follows exactly in the same way.

Now, consider $\Delta_{12}E_1 + \Delta_{12}E_0$, that is defined for $\lambda \in \Lambda^{\gamma_1}(\mathbf{u}_1) \cap \Lambda^{\gamma_2}(\mathbf{u}_2)$. Define also $E_{1,i} := E_1(\mathbf{u}_i)$ and $E_{0,i} := E_0(\mathbf{u}_i)$, for $i = 1, 2$. We prove the bounds only for $E_0^{+1}$, which is the hardest case, the bounds on $E_1^+$ follow in the same way. By Lemma 4.3.62 and the definition of $E_0^+$ (see (4.3.158)) one has

$$
\begin{aligned}
|\Delta_{12}E_0^+|_s &\overset{(4.3.133),(4.3.134)}{\leq_s} |\Pi_N^\perp \Delta_{12}E_0|_s + N^{2\tau+1}\gamma^{-1}(\delta_s(\mathbf{u}_1)+\delta_s(\mathbf{u}_2))|\Delta_{12}E_0|_{\mathfrak{s}_0} \\
&+N^{2\tau+1}\gamma^{-1}(\delta_{\mathfrak{s}_0}(\mathbf{u}_1)+\delta_{\mathfrak{s}_0}(\mathbf{u}_2))(\delta_s(\mathbf{u}_1)+\delta_s(\mathbf{u}_2))||\mathbf{u}_1-\mathbf{u}_2||_{s+\eta_2} \\
&+N^{2\tau+1}\gamma^{-1}(\delta_{\mathfrak{s}_0}(\mathbf{u}_1)+\delta_{\mathfrak{s}_0}(\mathbf{u}_2))|\Delta_{12}E_0|_s, \qquad (4.3.161)
\end{aligned}
$$

We prove equivalent bounds for $E_1^+$; then we obtain (4.3.156) by using the bounds given in Lemmata 4.3.72 and 4.3.63 to estimate the norms of the transformation $\Phi$. □

In the next Section we will show that it is possible to iterate the procedure described above infinitely many times.

**The iterative Scheme** Here we complete the proof of the Lemma 4.3.71 by induction on $\nu \geq 0$. Hence, assume that $(\mathbf{Si})_\nu$ hold. Then we prove $(\mathbf{Si})_{\nu+1}$ for $i = 1, 2, 3, 4$. We will use the estimates obtained in the previous Section.
$(\mathbf{S1})_{\nu+1}$ The eigenvalues $\mu_h^\nu$ of $\mathcal{D}_\nu$ are defined on $\Lambda_\nu^\gamma$, and identify the set $\Lambda_{\nu+1}^\gamma$ . Then, by Lemma 4.3.72, for any $\lambda \in \Lambda_{\nu+1}^\gamma$ there exists a unique solution $\Psi_\nu$ of the equation (4.3.132) such that, by inductive hypothesis $(\mathbf{S1})_\nu$, one has

$$
|\Psi_\nu|_{s,\gamma} \overset{(4.3.133)}{\lessdot} \gamma^{-1}N_\nu^{2\tau+1}\delta_s^\nu \overset{(4.3.126)}{\leq} \gamma^{-1}N_\nu^{2\tau+1}N_{\nu-1}^{-\alpha}\delta_{s+\beta}^0. \qquad (4.3.162)
$$

Hence the (4.3.121) holds at the step $\nu+1$. Moreover, by (4.3.162) and hypothesis (4.3.119), one has for $s = \mathfrak{s}_0$

$$
C(\mathfrak{s}_0)|\Psi_\nu|_{\mathfrak{s}_0,\gamma} \leq C(\mathfrak{s}_0)\gamma^{-1}N_\nu^{2\tau+1}N_{\nu-1}^{-\alpha}\delta_{\mathfrak{s}_0+\beta}^0 \leq \frac{1}{2}, \qquad (4.3.163)
$$

for $N_0$ large enough and using for $\nu = 1$ the smallness condition (4.3.119). By Lemma 4.3.63 we have that the transformation $\Phi_\nu := \mathcal{I} + \Psi_\nu$ is invertible with

$$
|\Phi_\nu^{-1}|_{\mathfrak{s}_0,\gamma} \leq 2, \quad |\Phi_\nu^{-1}|_{s,\gamma} \leq 1 + C(s)|\Psi_\nu|_{s,\gamma}. \qquad (4.3.164)
$$

Now, by Lemma 5.4.120, we have $\mathcal{L}_{\nu+1} := \Phi_\nu^{-1}\mathcal{L}_\nu\Phi_\nu = \omega\cdot\partial_\varphi \mathbb{1} + \mathcal{D}_{\nu+1} + \mathcal{R}_{\nu+1}$, where

$$
\begin{aligned}
&\mathcal{D}_{\nu+1} := \mathrm{diag}_{h\in \boldsymbol{C}\times\mathbb{N}}\{\mu_h^{\nu+1}\}, \quad \mu_h^{\nu+1} := \mu_h^\nu + (E_0^\nu)_h^h(0) = \mu_h^0 + r_h^{\nu+1}, \\
&\mathcal{R}_{\nu+1} = \Phi_\nu^{-1}\left(\Pi_{N_\nu}^\perp\mathcal{R}_\nu + \mathcal{R}_\nu\Psi_\nu - \Psi_\nu[\mathcal{R}_\nu]\right) = E_1^{\nu+1}D + E_0^{\nu+1},
\end{aligned} \qquad (4.3.165)
$$

where $E_i^{\nu+1} \rightsquigarrow E_i^+$, see (4.3.158). Let us check the (4.3.126) on the remainder $\mathcal{R}_{\nu+1}$. By (4.3.155) in

Lemma 5.4.120, we have

$$\begin{aligned}\delta_s^{\nu+1} &\leq_s N_\nu^{-\beta}\delta_{s+\beta}^{\nu}+\gamma^{-1}N_\nu^{2\tau+1}\delta_{\mathfrak{s}_0}^{\nu}\delta_s^{\nu}\\ &\overset{(4.3.126)}{\leq_s} N_\nu^{-\beta}N_{\nu-1}\delta_{s+\beta}^{0}+\gamma^{-1}N_\nu^{2\tau+1}N_{\nu-1}^{-2\alpha}\delta_{\mathfrak{s}_0+\beta}^{0}\delta_{s+\beta}^{0}\\ &\overset{(4.3.117),(6.4.92),(4.3.119)}{\leq_s} \delta_{s+\beta}^{0}N_\nu^{-\alpha},\end{aligned} \tag{4.3.166}$$

that is the first of the (4.3.126) for $\nu \rightsquigarrow \nu+1$. In the last inequality we used that $\chi = 3\backslash 2$, $\beta > \alpha+1$ and $\chi(2\tau+1+\alpha) < 2\alpha$, and this justifies the choices of $\beta$ and $\alpha$ in (4.3.117). By using the (4.3.155) we have

$$\delta_{s+\beta}^{\nu+1} \leq_{s+\beta} \delta_{s+\beta}^{\nu}+\gamma^{-1}N_\nu^{2\tau+1}\delta_{\mathfrak{s}_0}\delta_{s+\beta}^{\nu} \leq_{s+\beta} \delta_{s+\beta}^{0}N_\nu, \tag{4.3.167}$$

for $N_0 = N_0(s,\beta)$ large enough. This completes the proof of the (4.3.126).

By using (4.3.154) in Lemma 5.4.120 we have, $\forall\, h \in \boldsymbol{C}\times\mathbb{N}$,

$$|\mu_h^{\nu+1}-\mu_h^{\nu}|_\gamma = |r_h^{\nu+1}-r_h^{\nu}|_\gamma \leq \delta_{\mathfrak{s}_0}^{\nu} \overset{(4.3.126)}{\leq} \delta_{\mathfrak{s}_0+\beta}^{0}N_{\nu-1}^{-\alpha}, \tag{4.3.168}$$

hence we get the (4.3.124) since $|r_h^{\nu+1}|_\gamma \leq \sum_{i=0}^{\nu}|r_h^{\nu+1}-r_h^{\nu}|_\gamma \overset{(4.3.168)}{\leq} \delta_{\mathfrak{s}_0+\beta}^{0}K$.

Finally, we have to check that $\overline{\mu_{\sigma,j}^{\nu+1}} = -\mu_{\sigma,j}^{\nu+1} = \mu_{-\sigma,j}^{\nu+1}$. It follows by the inductive hypotheses since, by (4.3.111), one has

$$\overline{(E_0^{\nu})_{\sigma,j}^{\sigma,j}(0)} = -(E_0^{\nu})_{\sigma,j}^{\sigma,j}(0) = (E_0^{\nu})_{-\sigma,j}^{-\sigma,j}(0)$$

$(\mathbf{S2})_{\nu+1}$ Thanks to (4.3.168) we can extend, by Kirszbraun theorem, the function $\mu_h^{\nu+1}-\mu_h^{\nu}$ to a Lipschitz function on $\Lambda$. Defining $\widetilde{\mu}_k^{\nu+1}$ in this way, this extension has the same Lipschitz norm, so that the bound (4.3.127) hold.

$(\mathbf{S3})_{\nu+1}$. Let $\lambda \in \Lambda_\nu^{\gamma_1}(\mathbf{u}_1)\cap\Lambda_\nu^{\gamma_2}(\mathbf{u}_2)$, by Lemma 4.3.72 we can construct operators $\Psi_\nu^i := \Psi_\nu(\mathbf{u}_i)$ and $\Phi_\nu^i = \Phi_\nu(\mathbf{u}_i)$ for $i=1,2$. Using the (4.3.134) we have that

$$\begin{aligned}|\Delta_{12}\Psi_\nu|_{\mathfrak{s}_0} &\overset{(4.3.126),(4.3.128)}{\lessdot} N_\nu^{2\tau+1}N_{\nu-1}^{-\alpha}\gamma^{-1}\left(\delta_{\mathfrak{s}_0+\beta}^{0}+\varepsilon\right)||\mathbf{u}_1-\mathbf{u}_2||_{\mathfrak{s}_0+\eta_2}\\ &\overset{(4.3.119)}{\lessdot} N_\nu^{2\tau+1}N_{\nu-1}^{-\alpha}||\mathbf{u}_1-\mathbf{u}_2||_{\mathfrak{s}_0+\eta_2} \overset{(4.3.117)}{\leq} ||\mathbf{u}_1-\mathbf{u}_2||_{\mathfrak{s}_0+\eta_2},\end{aligned} \tag{4.3.169}$$

where we used the fact that $\varepsilon\gamma^{-1}$ is small. Moreover one can note that

$$|\Delta_{12}\Phi_\nu^{-1}|_s \overset{(4.3.97),(4.3.169)}{\leq_s} \left(|\Psi_\nu^1|_s+|\Psi_\nu^2|_s\right)||\mathbf{u}_1-\mathbf{u}_2||_{\mathfrak{s}_0+\eta_2}+|\Delta_{12}\Psi_\nu|_s, \tag{4.3.170}$$

then, by using the inductive hypothesis (4.3.121), the (4.3.119) and the (4.3.170) for $s=\mathfrak{s}_0$, one obtains

$$|\Delta_{12}\Phi_\nu^{-1}|_{\mathfrak{s}_0} \lessdot ||\mathbf{u}_1-\mathbf{u}_2||_{\mathfrak{s}_0+\eta_2}. \tag{4.3.171}$$

The (4.3.156) with $s = \mathfrak{s}_0$ together with (4.3.119), (4.3.126) and (4.3.128) implies

$$\begin{aligned}|\Delta_{12}E_1^{\nu+1}|_{\mathfrak{s}_0}+|\Delta_{12}E_0^{\nu+1}|_{\mathfrak{s}_0}&\leq_{\mathfrak{s}_0}\left(\varepsilon N_{\nu-1}N_\nu^{-\beta}+N_\nu^{2\tau+1}N_{\nu-1}^{-2\alpha}\varepsilon^2\gamma^{-1}\right)||\mathbf{u}_1-\mathbf{u}_2||_{\mathfrak{s}_0+\eta_2}\\ &\leq \varepsilon N_\nu^{-\alpha}||\mathbf{u}_1-\mathbf{u}_2||_{\mathfrak{s}_0+\eta_2},\end{aligned} \tag{4.3.172}$$

for $N_0$ large enough and $\varepsilon\gamma^{-1}$ small. Moreover consider the (4.3.156) with $s = \mathfrak{s}_0 + \beta$, then by (4.3.119), (4.3.128) and (4.3.126), we obtain for $N_0$ large enough

$$\begin{aligned}&|\Delta_{12}E_1^{\nu+1}|_{\mathfrak{s}_0+\beta}+|\Delta_{12}E_0^{\nu+1}|_{\mathfrak{s}_0+\beta}\\ \leq_{\mathfrak{s}_0+\beta}&(\delta^\nu_{\mathfrak{s}_0+\beta}(\mathbf{u}_1)+\delta^\nu_{\mathfrak{s}_0+\beta}(\mathbf{u}_2))||\mathbf{u}_1-\mathbf{u}_2||_{s+\eta_2|}+|\Delta_{12}E_1^\nu|_{\mathfrak{s}_0+\beta}+|\Delta_{12}E_0^\nu|_{\mathfrak{s}_0+\beta}\\ \leq\quad& C(\mathfrak{s}_0+\beta)\varepsilon N_{\nu-1}||\mathbf{u}_1-\mathbf{u}_2||_{\mathfrak{s}_0+\eta_2}\leq\varepsilon N_\nu||\mathbf{u}_1-\mathbf{u}_2||_{\mathfrak{s}_0+\eta_2}.\end{aligned} \tag{4.3.173}$$

Finally note that the (4.3.129) is implied by (4.3.154) that has been proved in Lemma 5.4.120.

$(\mathbf{S4})_{\nu+1}$. Let $\lambda \in \Lambda^\gamma_{\nu+1}$, by (4.3.120) and the inductive hypothesis $(\mathbf{S4})_\nu$ one has that $\Lambda^\gamma_{\nu+1}(\mathbf{u}_1) \subseteq \Lambda^\gamma_\nu(\mathbf{u}_1) \subseteq \Lambda^{\gamma-\rho}_\nu(\mathbf{u}_2) \subseteq \Lambda^{\gamma/2}_\nu(\mathbf{u}_2)$. Hence the eigenvalues $\mu^\nu_h(\lambda, \mathbf{u}_2(\lambda))$ are well defined by the $(\mathbf{S1})_\nu$. Now, since $\lambda \in \Lambda^\gamma_\nu(\mathbf{u}_1) \cap \Lambda^{\gamma/2}_\nu(\mathbf{u}_2)$, we have for $h = (\sigma, j) \in \boldsymbol{C} \times \mathbb{N}$ and setting $h' = (\sigma', j') \in \boldsymbol{C} \times \mathbb{N}$

$$\begin{aligned}&|(\mu^\nu_h-\mu^\nu_{h'})(\lambda,\mathbf{u}_2(\lambda))-(\mu^\nu_h-\mu^\nu_{h'})(\lambda,\mathbf{u}_1(\lambda))|\\ \overset{(4.2.35)}{\leq}\;&|(\mu^0_h-\mu^0_{h'})(\lambda,\mathbf{u}_2(\lambda))-(\mu^0_h-\mu^0_{h'})(\lambda,\mathbf{u}_1(\lambda))|\\ +\;&2\sup_{h\in\boldsymbol{C}\times\mathbb{N}}|r^\nu_h(\lambda,\mathbf{u}_2(\lambda))-r^\nu_h(\lambda,\mathbf{u}_1(\lambda))|\\ \overset{(4.3.129)}{\leq}\;&\varepsilon C|\sigma j^2-\sigma' j'^2|||\mathbf{u}_2-\mathbf{u}_1||_{\mathfrak{s}_0+\eta_2}.\end{aligned} \tag{4.3.174}$$

The (4.3.174) implies that, for any $|\ell| \leq N_\nu$ and $j \neq j'$,

$$\begin{aligned}&|i\omega\cdot\ell+\mu^\nu_h(\mathbf{u}_2)-\mu^\nu_{h'}(\mathbf{u}_2)|\\ \overset{(4.3.120),(4.3.174)}{\geq}\;&\gamma|\sigma j^2-\sigma' j'^2|\langle\ell\rangle^{-\tau}-C|\sigma j^2-\sigma' j'^2|||\mathbf{u}_2-\mathbf{u}_1||_{\mathfrak{s}_0+\eta_2}\\ \overset{(\mathbf{S4})_\nu}{\geq}\;&(\gamma-\rho)|\sigma j^2-\sigma' j'^2|\langle\ell\rangle^{-\tau},\end{aligned} \tag{4.3.175}$$

where we used that, for any $\lambda \in \Lambda_0$, one has $C\varepsilon N_\nu^\tau||\mathbf{u}_1 - \mathbf{u}_2||_{\mathfrak{s}_0+\eta_2} \leq \rho$. Now the (4.3.175) imply that if $\lambda \in \Lambda^\gamma_{\nu+1}(\mathbf{u}_1)$ then $\lambda \in \Lambda^{\gamma-\rho}_{\nu+1}(\mathbf{u}_2)$, that is the $(\mathbf{S4})_{\nu+\mathbf{1}}$.

**Proof of Theorem 4.3.60** We want apply Lemma 4.3.71 to the linear operator $\mathcal{L}_0 = \mathcal{L}_4$ defined in (4.2.33) where $\mathcal{R}_0 := E^0_1 D + E^0_0$ defined in (4.3.115), and we have defined for $s \in [\mathfrak{s}_0, q - \eta_1 - \beta]$, $\delta^0_s := |E^0_1|_{s,\gamma} + |E^0_0|_{s,\gamma}$, then

$$\delta^0_{\mathfrak{s}_0+\beta}\overset{(4.2.36)}{\leq}\varepsilon C(\mathfrak{s}_0+\beta)(1+||\mathbf{u}||_{\beta+\mathfrak{s}_0+\eta_1,\gamma})\overset{(4.3.88)}{\leq}2\varepsilon C(\mathfrak{s}_0+\beta) \tag{4.3.176}$$

which implies $N_0^{C_0}\delta^0_{\mathfrak{s}_0+\beta}\gamma^{-1}\le 1$ if $\varepsilon\gamma^{-1}\le\epsilon_0$ is small enough, that is the (4.3.119). We first prove that there exists a final transformation $\Phi_\infty$. For any $\lambda\in\cap_{\nu\ge0}\Lambda^\gamma_\nu$ we define

$$\widetilde{\Phi}_\nu := \Phi_0\circ\Phi_1\circ\ldots\circ\Phi_\nu. \tag{4.3.177}$$

One can note that $\widetilde{\Phi}_{\nu+1}=\widetilde{\Phi}_\nu\circ\Phi_{\nu+1}=\widetilde{\Phi}_\nu+\widetilde{\Phi}_\nu\Psi_{\nu+1}$. Then, one has

$$|\widetilde{\Phi}_{\nu+1}|_{\mathfrak{s}_0,\gamma}\overset{(4.3.94b)}{\le}|\widetilde{\Phi}_\nu|_{\mathfrak{s}_0,\gamma}+C|\widetilde{\Phi}_\nu|_{\mathfrak{s}_0,\gamma}|\Psi_{\nu+1}|_{\mathfrak{s}_0,\gamma}\overset{(4.3.121)}{\le}|\widetilde{\Phi}_\nu|_{\mathfrak{s}_0,\gamma}(1+\varepsilon_\nu^{(\mathfrak{s}_0)}), \tag{4.3.178}$$

where we have defined for $s\ge\mathfrak{s}_0$,

$$\varepsilon_\nu^{(s)}:=K\gamma^{-1}N_{\nu+1}^{2\tau+1}N_\nu^{-\alpha}\delta^0_s, \tag{4.3.179}$$

for some constant $K>0$. Now, by iterating (4.3.178) and using (4.3.119) and (4.3.121), we obtain

$$|\widetilde{\Phi}_{\nu+1}|_{\mathfrak{s}_0,\gamma}\le|\widetilde{\Phi}_0|_{\mathfrak{s}_0,\gamma}\prod_{\nu\ge0}(1+\varepsilon_\nu^{(\mathfrak{s}_0)})\le 2 \tag{4.3.180}$$

The estimate on the high norm follows by

$$\begin{aligned}|\widetilde{\Phi}_{\nu+1}|_{s,\gamma}&\overset{(4.3.94a),(4.3.180)}{\le}|\widetilde{\Phi}_\nu|_{s,\gamma}(1+C(\mathfrak{s}_0)|\Psi_{\nu+1}|_{\mathfrak{s}_0,\gamma})+C(s)|\widetilde{\Phi}_\nu|_{\mathfrak{s}_0,\gamma}|\Psi_{\nu+1}|_{s,\gamma}\\&\overset{(4.3.121),(6.4.92)}{\le}|\widetilde{\Phi}_\nu|_{s,\gamma}(1+\varepsilon_\nu^{(\mathfrak{s}_0)})+\varepsilon_\nu^{(s)}\le C\left(\sum_{j=0}^\infty\varepsilon_j^{(s)}+|\widetilde{\Phi}_0|_{s,\gamma}\right)\\&\overset{(4.3.121)}{\le}C(s)\left(1+\delta^0_{s+\beta}\gamma^{-1}\right),\end{aligned} \tag{4.3.181}$$

where we used the inequality $\prod_{j\ge0}(1+\varepsilon_j^{(\mathfrak{s}_0)})\le2$. Thanks to (4.3.181) we can prove that the sequence $\widetilde{\Phi}_\nu$ is a Cauchy sequence in norm $|\cdot|_{s,\gamma}$. Indeed one has

$$\begin{aligned}|\widetilde{\Phi}_{\nu+m}-\widetilde{\Phi}_\nu|_{s,\gamma}&\le\sum_{j=\nu}^{\nu+m-1}|\widetilde{\Phi}_{j+1}-\widetilde{\Phi}_j|_{s,\gamma}\\&\overset{(4.3.94a)}{\le}C(s)\sum_{j=\nu}^{\nu+m-1}(|\widetilde{\Phi}_j|_{s,\gamma}|\Psi_{j+1}|_{\mathfrak{s}_0,\gamma}+|\widetilde{\Phi}_j|_{\mathfrak{s}_0,\gamma}|\Psi_{j+1}|_{s,\gamma})\\&\overset{(4.3.121),(4.3.180),(4.3.181),(4.3.119)}{\le}C(s)\sum_{j\ge\nu}\delta^0_{s+\beta}\gamma^{-1}N_j^{-1}\\&\le C(s)\delta^0_{s+\beta}\gamma^{-1}N_\nu^{-1}.\end{aligned} \tag{4.3.182}$$

As consequence one has that $\widetilde{\Phi}_\nu\overset{|\cdot|_{s,\gamma}}{\to}\Phi_\infty$. Moreover (4.3.182) used with $m=\infty$ and $\nu=0$ and $|\widetilde{\Phi}_0-\mathbb{1}|_{s,\gamma}=|\Psi_0|_{s,\gamma}\le\gamma^{-1}\delta^0_{s+\beta}$ imply

$$|\Phi_\infty-\mathbb{1}|_{s,\gamma}\le C(s)\gamma^{-1}\delta^0_{s+\beta},\qquad|\Phi_\infty^{-1}-\mathbb{1}|_{s,\gamma}\overset{(4.3.96)}{\le}C(s)\gamma^{-1}\delta^0_{s+\beta}. \tag{4.3.183}$$

Hence the (4.3.92) is verified. Let us now define for $k = (\sigma, j) \in \boldsymbol{C} \times \mathbb{N}$,

$$\mu_k^\infty := \mu_{\sigma,j}^\infty(\lambda) = \lim_{\nu\to+\infty} \widetilde{\mu}_{\sigma,j}^\nu(\lambda) = \widetilde{\mu}_{\sigma,j}^0(\lambda) + \lim_{\nu\to+\infty} \widetilde{r}_{\sigma,j}^\nu. \tag{4.3.184}$$

We can note that, for any $\nu, j \in \mathbb{N}$, the following estimates on the eigenvalues hold:

$$|\mu_k^\infty - \widetilde{\mu}_k^\nu|_{\Lambda,\gamma} \le \sum_{m=\nu}^\infty |\widetilde{\mu}_k^{m+1} - \widetilde{\mu}_k^m|_{\Lambda,\gamma} \overset{(4.3.127),(4.3.126)}{\le} C\delta_{\mathfrak{s}_0+\beta}^0 N_{\nu-1}^{-\alpha}, \tag{4.3.185}$$

and moreover,

$$|\mu_k^\infty - \widetilde{\mu}_k^0|_{\Lambda,\gamma} \le C\delta_{\mathfrak{s}_0+\beta}^0. \tag{4.3.186}$$

As seen in Lemma 4.3.71 the corrections are $r_{\sigma,j}^\nu = r_{\sigma,j}^{\nu-1} + (E_0^\nu)_{\sigma,j}^{\sigma,j}(0)$.
The following Lemma gives us a connection between the Cantor sets defined in Lemma 4.3.71 and Theorem 4.3.60.

**Lemma 4.3.75.** *One has that*

$$\Lambda_\infty^{2\gamma} \subset \cap_{\nu\ge 0}\Lambda_\nu^\gamma. \tag{4.3.187}$$

*Proof.* Consider $\lambda \in \Lambda_\infty^{2\gamma}$. We show by induction that $\lambda \in \Lambda_\nu^\gamma$ for $\nu > 0$, since by definition we have $\Lambda_\infty^{2\gamma} \subset \Lambda_0^\gamma := \Lambda_o$. Assume that $\Lambda_\infty^{2\gamma} \subset \Lambda_{\nu-1}^\gamma$. Hence $\mu_h^\nu$ are well defined and coincide, in $\Lambda_\infty^{2\gamma}$, with their extension. Then, for any fixed $k = (\sigma, j, p), k' = (\sigma', j', p') \in \boldsymbol{C} \times \mathbb{N} \times \mathbb{Z}^d$, we have (recall $\ell = p - p'$)

$$|\omega \cdot \ell + \mu_{\sigma,j}^\nu - \mu_{\sigma',j'}^\nu| \overset{(4.1.23),(4.3.185)}{\ge} \frac{2\gamma|\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^\tau} - 2C\delta_{\mathfrak{s}_0+\beta}^0 N_{\nu-1}^{-\alpha}. \tag{4.3.188}$$

Now, by the smallness hypothesis (4.3.119), we can estimate for $|p - p'| = |\ell| \le N_\nu$,

$$|\omega \cdot \ell + \mu_{\sigma,j}^\nu - \mu_{\sigma',j'}^\nu| \ge \frac{\gamma|\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^\tau}, \tag{4.3.189}$$

that implies $\lambda \in \Lambda_\nu^\gamma$. □

Now, for any $\lambda \in \Lambda_\infty^{2\gamma} \subset \cap_{\nu\ge 0}\Lambda_\nu^\gamma$ (see (4.3.187)), one has

$$\begin{aligned} |\mathcal{D}_\nu - \mathcal{D}_\infty|_{s,\gamma} &= \sup_{k\in \boldsymbol{C}\times\mathbb{N}\times\mathbb{Z}^d} |\mu_{\sigma,j}^\nu - \mu_{\sigma',j'}^\infty|_\gamma \overset{(4.3.185),(4.3.186)}{\le} K\delta_{\mathfrak{s}_0+\beta}^0 N_{\nu-1}^{-\alpha}, \\ \delta_s^\nu &\overset{(4.3.126)}{\le} \delta_{s+\beta}^\beta N_{\nu-1}^{-\alpha}, \end{aligned} \tag{4.3.190}$$

that implies

$$\mathcal{L}_\nu \overset{(4.3.122)}{=} \mathcal{D}_\nu + \mathcal{R}_\nu \overset{|\cdot|_{s,\gamma}}{\to} \mathcal{D}_\infty =: \mathcal{L}_\infty, \quad \mathcal{D}_\infty := \mathrm{diag}_{k\in\boldsymbol{C}\times\mathbb{N}\times\mathbb{Z}^n}\mu_k^\infty. \tag{4.3.191}$$

By applying iteratively the (4.3.122) we obtain $\mathcal{L}_\nu = \widetilde{\Phi}_{\nu-1}^{-1}\mathcal{L}_0\widetilde{\Phi}_{\nu-1}$ where $\widetilde{\Phi}_{\nu-1}$ is defined in (4.3.177) and, by (4.3.182), $\widetilde{\Phi}_{\nu-1} \to \Phi_\infty$ in norm $|\cdot|_{s,\gamma}$. Passing to the limit we get

$$\mathcal{L}_\infty = \Phi_\infty^{-1} \circ \mathcal{L}_0 \circ \Phi_\infty, \tag{4.3.192}$$

that is the (4.3.91), while the (4.3.90) follows by (4.3.176), (4.3.185) and (4.3.186). Finally, (4.3.94a), (4.3.94c), Lemma 4.3.64 and (4.3.92) implies the bounds (4.3.93). This concludes the proof. □

## 4.4 Conclusion of the diagonalization algorithm

In the previous Section we have conjugated the operator $\mathcal{L}_4$ (see (4.2.33)) to a diagonal operator $\mathcal{L}_\infty$. In conclusion, we have that

$$\mathcal{L} = W_1\mathcal{L}_\infty W_2^{-1}, \quad W_i = \mathcal{V}_i\Phi_\infty, \quad \mathcal{V}_1 := \mathcal{T}_1\mathcal{T}_2\mathcal{T}_3\rho\mathcal{T}_4, \quad \mathcal{V}_2 = \mathcal{T}_1\mathcal{T}_2\mathcal{T}_3\mathcal{T}_4. \tag{4.4.193}$$

We have the following result

**Lemma 4.4.76.** *Let $\mathfrak{s}_0 \le s \le q - \beta - \eta_1 - 2$, with $\eta_1$ define in (4.2.31) and $\beta$ in Theorem (4.3.60). Then, for $\varepsilon\gamma^{-1}$ small enough, and*

$$||\mathbf{u}||_{\mathfrak{s}_0+\beta+\eta_1+2,\gamma} \le 1, \tag{4.4.194}$$

*one has for any $\lambda \in \Lambda_\infty^{2\gamma}$,*

$$||W_i\mathbf{h}||_{s,\gamma} + ||W_i^{-1}\mathbf{h}||_{s,\gamma} \le C(s)\left(||\mathbf{h}||_{s+2,\gamma} + ||\mathbf{u}||_{s+\beta+\eta_1+4,\gamma}||\mathbf{h}||_{\mathfrak{s}_0,\gamma}\right), \tag{4.4.195}$$

*for $i = 0, 1$. Moreover $W_i$ and $W_i^{-1}$ are reversibility preserving.*

*Proof.* Each $W_i$ is composition of two operators, the $\mathcal{V}_i$ satisfy the (4.2.34) while $\Phi_\infty$ satisfies (4.3.92). We use (4.3.94c) in order to pass to the operator norm. Then Lemma A.168 and (A.2) with $p = s - \mathfrak{s}_0$, $q = 2$ implies the bounds (4.4.195). Moreover the transformations $W_i$ and $W_i^{-1}$ are reversibility preserving because each of the transformations $\mathcal{V}_i, \mathcal{V}_i^{-1}$ and $\Phi_\infty$, $\Phi_\infty^{-1}$ is reversibility preserving. □

### 4.4.1 Proof of Proposition 4.1.51

We fix $\eta = \eta_1 + \beta + 2$ and $q > \mathfrak{s}_0 + \eta$. Let $\mu_h^\infty$ be the functions defined in (4.3.184). Then by Theorem 4.3.60 and Lemma 4.4.76 for $\lambda \in \Lambda_\infty^{2\gamma}$ we have the (4.1.24). Hence item (i) is proved.

Item (ii) follows by applying the dynamical system point of view. We have already proved that

$$\mathcal{L} = \mathcal{T}_1\mathcal{T}_2\mathcal{T}_3\rho\mathcal{T}_4\Phi_\infty\mathcal{L}_\infty\Phi_\infty^{-1}\mathcal{T}_4^{-1}\mathcal{T}_3^{-1}\mathcal{T}_2^{-1}\mathcal{T}_1^{-1}. \tag{4.4.196}$$

By Lemma 4.2.58 all the changes of variables in (4.4.196) can be seen as transformations of the phase space $\mathbf{H}^s_x$ depending in a quasi-periodic way on time plus quasi periodic reparametrization of time ($\mathcal{T}_3$). With this point of view, consider a dynamical system of the form

$$\partial_t \mathbf{u} = L(\omega t)\mathbf{u}. \tag{4.4.197}$$

Under a transformation of the form $\mathbf{u} = A(\omega t)\mathbf{v}$, one has that the system (4.4.197) become

$$\partial_t \mathbf{v} = L_+(\omega t)\mathbf{v}, \qquad L_+(\omega t) = A(\omega t)^{-1}L(\omega t)A(\omega t) - A(\omega t)^{-1}\partial_t A(\omega t) \tag{4.4.198}$$

The transformation $A(\omega t)$ acts on the functions $\mathbf{u}(\varphi, x)$ as

$$\begin{aligned} (A\mathbf{u})(\varphi, x) &:= (A(\varphi)\mathbf{u}(\varphi, \cdot))(x) := A(\varphi)\mathbf{u}(\varphi, x), \\ (A^{-1}\mathbf{u})(\varphi, x) &= A^{-1}(\varphi)\mathbf{u}(\varphi, x). \end{aligned} \tag{4.4.199}$$

Then the operator on the quasi-periodic functions

$$\mathcal{L} := \omega \cdot \partial_\varphi - L(\varphi), \tag{4.4.200}$$

associated to the system (4.4.197), is transformed by $A$ into

$$A^{-1}\mathcal{L}A = \omega \cdot \partial_\varphi - L_+(\varphi), \tag{4.4.201}$$

that represent the system in (4.4.198) acting on quasi-periodic functions. The same considerations hold for transformations of the type

$$\begin{aligned} &\tau := \psi(t) := t + \alpha(\omega t), \quad t = \psi^{-1}(\tau) := \tau + \widetilde{\alpha}(\omega\tau), \\ &(B\mathbf{u})(t) := \mathbf{u}(t + \alpha(\omega t)), \qquad (B^{-1}\mathbf{v})(\tau) = v(\tau + \widetilde{\alpha}(\omega\tau)). \end{aligned} \tag{4.4.202}$$

with $\alpha(\varphi)$, $\varphi \in \mathbb{T}^d$ is $2\pi-$periodic in all the $d$ variables. The operator $B$ is nothing but the operator on the functions induced by the diffeomorfism of the torus $t \to t + \alpha(\omega t)$. The transformation $\mathbf{u} = B\mathbf{v}$ transform the system (4.4.197) into

$$\partial_t \mathbf{v} = L_+(\omega t)\mathbf{v}, \qquad L_+(\omega\tau) := \left(\frac{L(\omega t)}{1 + (\omega \cdot \partial_\varphi \alpha)(\omega t)}\right)_{|t = \widetilde{\psi}(\tau)} \tag{4.4.203}$$

If we consider the operator $B$ acting on the quasi-periodic functions as $(B\mathbf{u})(\varphi, x) = \mathbf{u}(\varphi + \omega\alpha(\varphi), x)$ and $(B^{-1}\mathbf{u})(\varphi, x) := \mathbf{u}(\varphi + \omega\widetilde{\alpha}(\varphi), x)$, we have that

$$\begin{aligned} B^{-1}\mathcal{L}B &= \rho(\varphi)\mathcal{L}_+ = \rho(\varphi)\left(\omega \cdot \partial_\varphi - L_+(\varphi)\right) \\ &= \rho(\varphi)\left(\omega \cdot \partial_\varphi - \frac{1}{\rho(\varphi)}L(\varphi + \omega\widetilde{\alpha}(\varphi))\right), \end{aligned} \tag{4.4.204}$$

and $\rho(\varphi) := B^{-1}(1+\omega\cdot\partial_\varphi\alpha)$, that means that $\mathcal{L}_+$ is the linear system (4.4.203) acting on quasi-periodic functions.

By these arguments, we have simply that a curve $\mathbf{u}(t)$ in the phase space of functions of $x$, i.e. $\mathbf{H}^s_x$, solves the linear dynamical system (4.1.29) if and only if the curve

$$\mathbf{v}(t) := \Phi_\infty^{-1}\mathcal{T}_4^{-1}\mathcal{T}_3^{-1}\mathcal{T}_2^{-1}\mathcal{T}_1^{-1}(\omega t)\mathbf{h}(t) \tag{4.4.205}$$

solves the system (4.1.27). This completely justify Remark 5.0.86. In Lemma 4.2.58 and the (4.3.93) we have checked that these transformations are well defined. □

### 4.4.2 Inversion of $\mathcal{L}$

We chose to completely reduce to constant coefficients the operator $\mathcal{L}(\mathbf{u})$ because once it is diagonal it is trivial to invert it in an explicit Cantor like set. The following Lemma concludes the inversion of the linearized operator $\mathcal{L}$.

**Lemma 4.4.77 (Right inverse of $\mathcal{L}$).** *Under the hypotheses of Proposition 4.1.51, let us set*

$$\zeta := 4\tau + \eta + 8 \tag{4.4.206}$$

*where $\eta$ is fixed in Proposition 4.1.51. Consider a Lipschitz family $\mathbf{u}(\lambda)$ with $\lambda \in \Lambda_o \subseteq \Lambda \subseteq \mathbb{R}$ such that*

$$||\mathbf{u}||_{\mathfrak{s}_0+\zeta,\gamma} \leq 1. \tag{4.4.207}$$

*Define the set*

$$P_\infty^{2\gamma}(\mathbf{u}) := \left\{ \begin{array}{cc} \lambda\in\Lambda_o : & |\lambda\bar{\omega}\cdot\ell + \mu^\infty_{\sigma,j}(\lambda)| \geq \frac{2\gamma j^2}{\langle\ell\rangle^\tau}, \\ & \forall\ell\in\mathbb{Z}^d, \ \forall\,(\sigma,j)\in\boldsymbol{C}\times\mathbb{N} \end{array} \right\}. \tag{4.4.208}$$

*There exists $\epsilon_0$, depending only on the data of the problem, such that if $\varepsilon\gamma^{-1} < \epsilon_0$ then, for any $\lambda \in \Lambda^{2\gamma}_\infty(\mathbf{u})\cap P^{2\gamma}_\infty(\mathbf{u})$ (see (4.1.23)), and for any Lipschitz family $\mathbf{g}(\lambda) \in \mathbf{Z}^0$, the equation $\mathcal{L}\mathbf{h} := \mathcal{L}(\lambda, \mathbf{u}(\lambda))\mathbf{h} = \mathbf{g}$, where $\mathcal{L}$ is the linearized operator in (4.2.30), admits a solution*

$$\mathbf{h} := \mathcal{L}^{-1}\mathbf{g} := W_2\mathcal{L}_\infty^{-1}W_1^{-1}\mathbf{g} \in \mathbf{X}^0 \tag{4.4.209}$$

*such that*

$$||\mathbf{h}||_{s,\gamma} \leq C(s)\gamma^{-1}\left(||\mathbf{g}||_{s+2\tau+5,\gamma} + ||\mathbf{u}||_{s+\zeta,\gamma}||\mathbf{g}||_{\mathfrak{s}_0,\gamma}\right), \quad \mathfrak{s}_0 \leq s \leq q-\zeta. \tag{4.4.210}$$

*Proof.* As explained in the Introduction, we now study the invertibility of

$$\mathcal{L}_\infty := \mathrm{diag}_{k\in\boldsymbol{C}\times\mathbb{N}\times\mathbb{Z}^d}\{i\omega\cdot\ell + \mu^\infty_{\sigma,j}\}, \quad \mu^\infty_{\sigma,j}(\lambda) = -i\sigma m(\lambda)j^2 + r^\infty_{\sigma,j}(\lambda) \tag{4.4.211}$$

in order to obtain a better understanding of the set $\mathcal{G}_\infty$ of the Nash-Moser Proposition 4.0.46.

**Lemma 4.4.78.** *For $\mathbf{g} \in \mathbf{Z}^s$, consider the equation*

$$\mathcal{L}_\infty(\mathbf{u})\mathbf{h} = \mathbf{g}. \tag{4.4.212}$$

*If $\lambda \in \Lambda_\infty^{2\gamma}(\mathbf{u}) \cap P_\infty^{2\gamma}(\mathbf{u})$ (defined respectively in (4.1.23) and (4.4.208)), then there exists a unique solution $\mathcal{L}_\infty^{-1}\mathbf{g} := \mathbf{h} = (h, \bar{h}) \in \mathbf{X}^s$. Moreover, for any Lipschitz family $\mathbf{g} := \mathbf{g}(\lambda) \in \mathbf{Z}^s$ one has*

$$||\mathcal{L}_\infty^{-1}\mathbf{g}||_{s,\gamma} \leq C\gamma^{-1}||\mathbf{g}||_{s+2\tau+1,\gamma}. \tag{4.4.213}$$

*Proof.* By hypothesis $\mathbf{g} = (g, \bar{g})$. By solving the (6.4.208) one obtains the solution $\mathbf{h} := (h_+, h_-)$ of the form

$$\begin{aligned} h_+(\varphi, x) &:= \sum_{\ell \in \mathbb{Z}^d, j \geq 1} \frac{g_j(\ell)}{i\omega \cdot \ell + \mu_{1,j}^\infty} e^{i\ell\cdot\varphi} \sin jx, \\ h_-(\varphi, x) &:= \sum_{\ell \in \mathbb{Z}^d, j \geq 1} \frac{\overline{g_j(\ell)}}{-i\omega \cdot \ell + \mu_{-1,j}^\infty} e^{-i\ell\cdot\varphi} \sin jx. \end{aligned} \tag{4.4.214}$$

By the hypothesis of reversibility, we have already seen that $\mu_{1,j}^\infty = -\overline{\mu_{1,j}^\infty}$ and $\mu_{-1,j}^\infty = -\mu_{1,j}^\infty$, then one has that $\overline{h_-} = h_+ := h$. Moreover, one has

$$\overline{h_j(\ell)} = \overline{\frac{g_j(\ell)}{i\omega \cdot \ell + \mu_{1,j}^\infty}} = \frac{-g_j(\ell)}{-(i\omega \cdot \ell + \mu_{1,j}^\infty)} = h_j(\ell) \tag{4.4.215}$$

then the Lemma (4.3.68) implies that $\mathbf{h} \in \mathbf{X}^s$.

Now, since $\lambda \in \Lambda_\infty^{2\gamma}(\mathbf{u}) \cap P_\infty^{2\gamma}(\mathbf{u})$ then, by (4.4.208), we can estimate the (4.4.214)

$$||\mathbf{h}||_s \leq C\gamma^{-1}||\mathbf{g}||_{s+\tau}. \tag{4.4.216}$$

The Lipschitz bound on $\mathbf{h}$ follow exactly as in formulæ(4.3.144)-(4.3.146) and we obtain

$$||\mathbf{h}||_{s,\gamma} = ||\mathbf{h}||_s^{sup} + \gamma||\mathbf{h}||_s^{lip} \lessdot \gamma^{-1}||\mathbf{g}||_{s+2\tau+1,\gamma}, \tag{4.4.217}$$

that is the (4.4.213). □

We show in the next Lemma how to solve the equation $\mathcal{L}\mathbf{h} = \mathbf{g}$ for $\mathbf{g} \in \mathbf{Z}^s$:

By (4.4.193) one has that the equation $\mathcal{L}\mathbf{h} = \mathbf{g}$ is equivalent to $\mathcal{L}_\infty W_2^{-1}\mathbf{h} = W_1^{-1}\mathbf{g}$. By Lemma 4.4.78 this second equation has a unique solution $W_2^{-1}\mathbf{h} \in \mathbf{X}^s$. Note that this is true because $W_1^{-1}$ is reversibility-preserving, so that $W_1^{-1}\mathbf{g} \in \mathbf{Z}^s$ if $\mathbf{g} \in \mathbf{Z}^s$. Hence the solution with zero average of $\mathcal{L}\mathbf{h} = \mathbf{g}$ is of the form

$$\mathbf{h} := W_2\mathcal{L}_\infty^{-1}W_1^{-1}\mathbf{g}, \tag{4.4.218}$$

Since $W_2$ is reversibility-preserving and, by Lemma 4.4.78, one has that $\mathcal{L}_\infty^{-1} : \mathbf{Z}^0 \to \mathbf{X}^0$ then $\mathbf{h} \in \mathbf{X}^s$. Now we have

$$\begin{aligned} ||\mathbf{h}||_{s,\gamma} &\overset{(4.4.195)}{\leq} C(s)\left(||\mathcal{L}_\infty^{-1} W_1^{-1}\mathbf{g}||_{s+2,\gamma} + ||\mathbf{u}||_{s+\beta+\eta_1+4,\gamma}||\mathcal{L}_\infty^{-1} W_1^{-1}\mathbf{g}||_{\mathfrak{s}_0,\gamma}\right) \\ &\leq C(s)\gamma^{-1}\left(||\mathbf{g}||_{s+2\tau+5,\gamma} + ||\mathbf{u}||_{s+4\tau+\beta+10+\eta_1,\gamma}||\mathbf{g}||_{\mathfrak{s}_0,\gamma}\right), \end{aligned} \tag{4.4.219}$$

where, in the second inequality, we used (4.4.213) on $\mathcal{L}_\infty^{-1}$, again the (4.4.195) for $W_1^{-1}$ and (4.4.207). Finally we used the (A.2) with $a = \mathfrak{s}_0 + 2\tau + \eta_1 + \beta + 7, b = \mathfrak{s}_0$ and $p = s - \mathfrak{s}_0, q = 2\tau + 3$. Formula (4.4.219) implies (4.4.210) with $\zeta$ defined in (4.4.206) where we already fixed $\eta := \eta_1 + \beta + 2$ in the proof of Proposition 4.1.51. $\square$

## 4.5 Measure estimates and conclusions

The aim of this Section is to use the information obtained in Sections 4.2 and 4.3, in order to apply Theorem 3.1.18 to our problem and prove Theorem 1.1.1.

By formula (4.4.210) we have good bounds on the inverse of $\mathcal{L}(\mathbf{u}_n)$ in the set $\Lambda_\infty^{2\gamma}(\mathbf{u}_n) \cap P_\infty^{2\gamma}(\mathbf{u}_n)$. It is easy to see that this sets have positive measure for all $n \geq 0$. Now in the Nash-Moser proposition 4.0.46 we defined the sets $\mathcal{G}_n$ in order to ensure bounds on the inverse of $\mathcal{L}(\mathbf{u}_n)$, thus we have the following Proposition on the measure of such set.

**Proposition 4.5.79 (Measure estimates).** *Set $\gamma_n := (1+2^{-n})\gamma$ and consider the set $\mathcal{G}_\infty$ of Proposition 4.0.46 with $\mu = \zeta$ defined in Lemma 4.4.77 . We have*

$$\cap_{n\geq 0} \Lambda_\infty^{2\gamma_n}(\mathbf{u}_n) \cap P_\infty^{2\gamma_n}(\mathbf{u}_n) \subseteq \mathcal{G}_\infty, \tag{4.5.220a}$$

$$|\Lambda \backslash \mathcal{G}_\infty| \to 0, \quad \text{as} \quad \gamma \to 0. \tag{4.5.220b}$$

Formula (4.5.220a) is essentially trivial. One just needs to look at Definition 3.1.17 and item $(N1)_n$ of Theorem 3.1.18, which fix the sets $\mathcal{G}_n$. The (4.5.220b) is more delicate. The first point is that we reduce to computing the measure of the left hand side of (4.5.220a). It is simple to show that each $\Lambda_\infty^{2\gamma_n}(\mathbf{u}_n) \cap P_\infty^{2\gamma_n}(\mathbf{u}_n)$ has measure $1 - O(\gamma)$, however in principle as $n$ varies this sets are unrelated and then the intersection might be empty. We need to study the dependence of the Cantor sets on the function $\mathbf{u}_n$. Indeed $\Lambda_\infty^{2\gamma}(\mathbf{u})$ is constructed by imposing infinitely many *second Mel'nikov conditions.* We show that this conditions imply a **finitely** many second Mel'nikov conditions on a whole neighborhood of $\mathbf{u}$.

We first prove the approximate reducibility

**Lemma 4.5.80.** *Under the hypotheses of Proposition 4.1.51, for $N$ sufficiently large, for any $0 < \rho < \gamma/2$ and for any Lipschitz family $\mathbf{v}(\lambda) \in \mathbf{X}^0$ with $\lambda \in \Lambda_o$ such that*

$$\sup_{\lambda \in \Lambda_o} ||\mathbf{u} - \mathbf{v}||_{\mathfrak{s}_0+\eta} \le \varepsilon C \rho N^{-\tau}, \tag{4.5.221}$$

*we have the following. For all $\lambda \in \Lambda_\infty^{2\gamma}(\mathbf{u})$ there exist invertible and reversibility-preserving (see Section 4.1 for a precise definition) transformations $V_i$ for $i = 1, 2$ such that*

$$V_1^{-1} \mathcal{L}(\mathbf{v}) V_2 = \omega \cdot \partial_\varphi \mathbb{1} + diag_{h \in \boldsymbol{C} \times \mathbb{N}} \{\mu_h^{(N)}\} + E_1 \partial_x + E_0 : \mathbf{X}^0 \to \mathbf{Z}^0, \tag{4.5.222}$$

*where $\mu_h^{(N)}$ have the same form of $\mu_h^\infty$ in (4.1.24) with bounds*

$$|r_h^\infty - r_h^{(N)}|_\gamma \le \varepsilon C ||\mathbf{u} - \mathbf{v}||_{\mathfrak{s}_0+\eta,\gamma} + C \varepsilon N^{-\kappa}, \tag{4.5.223}$$

*for an appropriate $\kappa$ depending only on $\tau$. More precisely $\Lambda_\infty^{2\gamma}(\mathbf{u}) \subset \Lambda_N^{\gamma-\rho}(\mathbf{v})$ with*

$$\Lambda_N^{\gamma-\rho}(\mathbf{v}) := \left\{ \begin{array}{c} \lambda \in \Lambda_o : \quad |\lambda \bar{\omega} \cdot \ell + \mu_{\sigma,j}^{(N)}(\lambda) - \mu_{\sigma',j'}^{(N)}(\lambda)| \ge \frac{(\gamma-\rho)|\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^\tau}, \\ \forall\, |\ell| < N, \ \forall (\sigma, j), (\sigma', j') \in \boldsymbol{C} \times \mathbb{N} \end{array} \right\}.$$

*Finally the $V_i$ satisfy bounds like (5.0.7) and the remainders satisfy*

$$||E_0 \mathbf{h}||_s + ||E_1 \mathbf{h}||_s \le \varepsilon C N^{-\kappa} (||\mathbf{h}||_s + ||\mathbf{v}||_{s+\eta} ||\mathbf{h}||_{\mathfrak{s}_0}). \tag{4.5.224}$$

*Proof.* We first apply the change of variables defined in (4.2.80) to $\mathcal{L}(\mathbf{v})$ in order to reduce to $\mathcal{L}_4(\mathbf{v})$. We know that Lemma 4.3.71 holds for $\mathcal{L}_4(\mathbf{u})$, now we fix $\nu$ such that $N_{\nu-1} \le N \le N_\nu$ and apply $(\mathbf{S3})_\nu - (\mathbf{S4})_\nu$ with $\mathbf{u}_1 = \mathbf{u}$, $\mathbf{u}_2 = \mathbf{v}$. This implies our claim since, by Lemma 4.3.75, we have $\Lambda_\infty^{2\gamma}(\mathbf{u}) \subseteq \Lambda_\nu^{\gamma}(\mathbf{u}) \subseteq \Lambda_\nu^{\gamma-\rho}(\mathbf{v})$. Finally for all $\lambda \in \Lambda_{\nu+1}^{\gamma-\rho}(\mathbf{v})$ we can perform $\nu + 1$ steps in Lemma 4.3.71. Fixing $\kappa = 2\alpha/3$ we obtain the bounds on the changes of variables and remainders, using formulæ (4.3.164) and (4.3.168). □

**Proof of Proposition 4.5.79.** Recall that we have set $\gamma_n := \gamma\,(1 + 2^{-n})$, $(\mathbf{u}_n)_{\ge 0}$ is the sequence of approximate solutions introduced in Theorem 3.1.18 which is well defined in $\mathcal{G}_n$ and satisfies the hypotheses of Proposition 4.1.51. $\mathcal{G}_n$ in turn is defined in $(N1)_n$ and Definition 3.1.17. For notational convenience we extend the eigenvalues $\mu_{\sigma,j}^\infty(\mathbf{u}_n)$ introduced in Proposition 4.1.51, which are defined only for $j \in \mathbb{N}$, to a function defined for $j \in \mathbb{Z}_+$ in the following way:

$$\Omega_{\sigma,j}(\mathbf{u}_n) := \mu_{\sigma,j}^\infty(\mathbf{u}_n), \ \ (\sigma, j) \in \boldsymbol{C} \times \mathbb{N}, \ \ \Omega_{\sigma,j}(\mathbf{u}_n) \equiv 0, \quad \sigma \in \boldsymbol{C}, \ \ j = 0. \tag{4.5.225}$$

**Definition 4.5.81.** *We now define inductively a sequence of nested sets $G_n$ for $n \geq 0$. Set $G_0 = \Lambda$ and*

$$G_{n+1} := \left\{ \begin{array}{c} \lambda \in G_n \cap \mathcal{G}_n : |i\omega \cdot \ell + \Omega_{\sigma,j}(\mathbf{u}_n) - \Omega_{\sigma',j'}(\mathbf{u}_n)| \geq \dfrac{2\gamma_n |\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^{\tau}}, \\ \forall \ell \in \mathbb{Z}^n, \ \ \sigma, \sigma' \in \boldsymbol{C}, \ \ j, j' \in \mathbb{Z}_+ \end{array} \right\}.$$

The following Lemma implies (4.5.220a).

**Lemma 4.5.82.** *Under the Hypotheses of Proposition 4.5.79, for any $n \geq 0$, one has*

$$G_{n+1} \subseteq \mathcal{G}_{n+1}. \tag{4.5.226}$$

*Proof.* For any $n \geq 0$ and if $\lambda \in G_{n+1}$ one has, by Lemmata 4.4.78-4.4.77 and recalling that $\gamma \leq \gamma_n \leq 2\gamma$ and $2\tau + 5 < \zeta$,

$$\begin{aligned} &||\mathcal{L}^{-1}(\mathbf{u}_n)\mathbf{g}||_{s,\gamma} \leq C(s)\gamma^{-1} \left( ||\mathbf{g}||_{s+\zeta,\gamma} + ||\mathbf{u}_n||_{s+\zeta,\gamma} ||\mathbf{g}||_{\mathfrak{s}_0,\gamma} \right), \\ &||\mathcal{L}^{-1}(\mathbf{u}_n)||_{\mathfrak{s}_0,\gamma} \leq C(\mathfrak{s}_0)\gamma^{-1} N_n^{\zeta} ||\mathbf{g}||_{\mathfrak{s}_0,\gamma}, \end{aligned} \tag{4.5.227}$$

for $\mathfrak{s}_0 \leq s \leq q - \mu$, for any $\mathbf{g}(\lambda)$ Lipschitz family. The (5.4.213) are nothing but the (3.1.4) in Definition 3.1.17 with $\mu = \zeta$ . $\mu$ represents the loss of regularity that you have when you perform the regularization procedure in Section 4.2 and during the diagonalization algorithm in Section 4.3. This justifies our choice of $\mu$ in Proposition 4.5.79. □

By Lemma 5.4.113, in order to obtain the bound (4.5.220b), it is enough to prove that

$$|\Lambda \backslash \cap_{n \geq 0} G_n| \to 0, \quad \text{as} \quad \gamma \to 0. \tag{4.5.228}$$

We will prove by induction that, for any $n \geq 0$, one has

$$|G_0 \backslash G_1| \leq C_{\star}\gamma, \quad |G_n \backslash G_{n+1}| \leq C_{\star}\gamma N_n^{-1}, \quad n \geq 1. \tag{4.5.229}$$

First of all write

$$\begin{aligned} G_n \backslash G_{n+1} &:= \bigcup_{\substack{\sigma,\sigma' \in \boldsymbol{C}, j,j' \in \mathbb{Z}_+ \\ \ell \in \mathbb{Z}^n}} R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_n) \\ R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_n) &:= \left\{ \lambda \in G_n : |i\lambda\bar{\omega} \cdot \ell + \Omega_{\sigma,j}(\mathbf{u}_n) - \Omega_{\sigma',j'}(\mathbf{u}_n)| < \tfrac{2\gamma_n |\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^{\tau}} \right\}. \end{aligned} \tag{4.5.230}$$

By (1.1.2) we have $R^{\sigma,\sigma}_{\ell jj}(\mathbf{u_n}) = \emptyset$. In the following we assume that if $\sigma = \sigma'$, then $j \neq j'$. Important properties of the sets $R^{\sigma,\sigma'}_{\ell jj'}$ are the following. The proofs are quite standard and follow very closely Lemmata 5.2 and 5.3 in [4]. For completeness we give a proof in the Appendix C.1.

**Lemma 4.5.83.** *For any $n \geq 0$, $|\ell| \leq N_n$, one has, for $\varepsilon\gamma^{-1}$ small enough,*

$$R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_n) \subseteq R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_{n-1}). \tag{4.5.231}$$

*Moreover,*

$$\text{if} \quad R^{\sigma,\sigma'}_{\ell j j'} \neq \emptyset, \quad \text{then} \quad |\sigma j^2 - \sigma' j'^2| \leq 8|\bar{\omega}\cdot\ell|. \tag{4.5.232}$$

**Lemma 4.5.84.** *For all $n \geq 0$, one has*

$$|R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_n)| \leq C\gamma\langle\ell\rangle^{-\tau}. \tag{4.5.233}$$

We now prove (4.5.229)-(4.5.228) by assuming Lemmata 4.5.83 and 4.5.84. By (4.5.230) one has $R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_n) \subset G_n$, and at the same time for all $|\ell| \leq N_n$ one has $R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_n) \subseteq R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_{n-1})$ by (4.5.231). Hence, if $|\ell| \leq N_n$, one has $R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_n) = \emptyset$ since $R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_{n-1}) \cap G_n = \emptyset$ by Definition 5.4.211. This implies that

$$G_n \backslash G_{n+1} \subseteq \bigcup_{\substack{\sigma,\sigma' \in \boldsymbol{C}, j,j' \in \mathbb{Z}_+ \\ |\ell| > N_n}} R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_n) \tag{4.5.234}$$

Now, consider the sets $R^{\sigma,\sigma'}_{\ell j j'}(0)$. By (4.5.232), we know that if $R^{\sigma,\sigma'}_{\ell j j'}(0) \neq \emptyset$ then we must have $j + j' \leq 16|\bar{\omega}||\ell|$. Indeed, if $\sigma = \sigma'$, then

$$|j^2 - j'^2| = |j - j'|(j + j') \geq \frac{1}{2}(j + j'), \quad \forall j, j' \in \mathbb{Z}_+, \;\; j \neq j', \tag{4.5.235}$$

while, if $\sigma \neq \sigma'$, one has $(j + j')/2 \leq (j^2 + j'^2) \leq 8|\bar{\omega}||\ell|$ see (4.5.232). Then, for $\tau > d + 2$, we obtain the first of (4.5.229), by

$$|G_0 \backslash G_1| \leq \sum_{\substack{\sigma,\sigma' \in \boldsymbol{C}, \\ j,j' \in \mathbb{Z}_+ \\ \ell \in \mathbb{Z}^d}} |R^{\sigma,\sigma'}_{\ell j j'}(0)| \leq \sum_{\substack{\sigma,\sigma' \in \boldsymbol{C}, \\ (j+j') \leq 16|\bar{\omega}||\ell| \\ \ell \in \mathbb{Z}^d}} |R^{\sigma,\sigma'}_{\ell j j'}(0)| \overset{(4.5.233)}{\leq} C\gamma \sum_{\ell \in \mathbb{Z}^d} \langle\ell\rangle^{-(\tau-1)} \leq C\gamma.$$

Finally, we have for any $n \geq 1$,

$$|G_n \backslash G_{n+1}| \overset{(4.5.234)}{\leq} \sum_{\substack{\sigma,\sigma' \in \boldsymbol{C}, \\ (j+j') \leq 16|\bar{\omega}||\ell| \\ |\ell| > N_n}} |R^{\sigma,\sigma'}_{\ell j j'}(\mathbf{u}_n)| \overset{(4.5.233)}{\leq} \sum_{|\ell| > N_n} \frac{C\gamma}{\langle\ell\rangle^{(\tau-1)}} \leq C\gamma N_n^{-1}, \tag{4.5.236}$$

since $\tau \geq d + 2$; we have obtained the (4.5.229). Now we have

$$|\Lambda \backslash \cap_{n \geq 0} G_n| \leq \sum_{n \geq 0} |G_n \backslash G_{n+1}| \leq C\gamma + C\gamma \sum_{n \geq 1} N_n^{-1} \leq C\gamma \to 0, \tag{4.5.237}$$

as $g \to 0$. By (5.4.212), we have that $\cap_{n \geq 0} G_n \subseteq \mathcal{G}_\infty$. Then, by (4.5.237), we obtain (4.5.220b). $\square$

### 4.5.1 Proof of Theorem 1.1.1

Fix $\gamma := \varepsilon^a$, $a \in (0,1)$. Then the smallness condition $\varepsilon\gamma^{-1} = \varepsilon^{1-a} < \epsilon_0$ of Theorem 3.1.18 is satisfied for $\varepsilon$ small. Then we can apply it with $\mu = \zeta$ in (4.4.206) (see Lemma 4.5.79). Hence by (3.1.9) we have that the function $\mathbf{u}_\infty$ in $\mathbf{X}^{\mathfrak{s}_0+\zeta}$ is a solution of the perturbed NLS with $\omega = \lambda\bar{\omega}$. Moreover, one has

$$|\Lambda\backslash\mathcal{G}_\infty| \stackrel{(4.5.220b)}{\to} 0, \tag{4.5.238}$$

as $\varepsilon$ tends to zero. To complete the proof of the theorem, it remains to prove the linear stability of the solution.

Since the eigenvalues $\mu^\infty_{\sigma,j}$ are purely imaginary, we know that the Sobolev norm of the solution $\mathbf{v}(t)$ of (4.1.27) is constant in time. We show that the Sobolev norm of $\mathbf{h}(t) = W_2\mathbf{v}(t)$, solution of (4.1.29) does not grow in time. To do this we first note that, by (5.1.102a) and (4.3.93), one has $\forall t \in \mathbb{R}$, $\forall \mathbf{g} = \mathbf{g}(x) \in \mathbf{H}^s_x$

$$\begin{aligned} ||\mathcal{T}_i^{\pm1}(\omega t)\mathbf{g}||_{H^s_x} + ||(\mathcal{T}_4\Phi_\infty)^{\pm1}(\omega t)\mathbf{g}||_{H^s_x} \le C(s)||\mathbf{g}||_{H^s_x}, \\ ||(\mathcal{T}_i^{\pm1}(\omega t) - \mathbb{1})\mathbf{g}||_{H^s_x} + ||((\mathcal{T}_4\Phi_\infty)^{\pm1}(\omega t) - \mathbb{1})\mathbf{g}||_{H^s_x} \le \varepsilon\gamma^{-1}C(s)||\mathbf{g}||_{H^{s+1}_x}. \end{aligned} \tag{4.5.239}$$

with $i = 1, 2$. In both cases, the constant $C(s)$ depends on $||\mathbf{u}||_{s+\mathfrak{s}_0+\beta+\eta_1}$. We claim that there exists a constant $K > 0$ such that the following bounds hold:

$$\begin{aligned} ||\mathbf{h}(t)||_{H^s_x} \le K||\mathbf{h}(0)||_{H^s_x}, \\ ||\mathbf{h}(0)||_{H^s_x} - \varepsilon^b K||\mathbf{h}(0)||_{H^{s+1}_x} \le ||\mathbf{h}(t)||_{H^s_x} \le ||\mathbf{h}(0)||_{H^s_x} + \varepsilon^b K||\mathbf{h}(0)||_{H^{s+1}_x}, \end{aligned} \tag{4.5.240}$$

for some $b \in (0,1)$. The (4.5.240) imply the linear stability of the solution.

Recalling that $\mathcal{T}_3 f(t) := f(t + \alpha(\omega t)) = f(\mathfrak{t})$ and $\mathcal{T}_3^{-1} f(\mathfrak{t}) = f(\mathfrak{t} + \widehat{\alpha}(\omega \mathfrak{t})) = f(t)$, fixing $\mathfrak{t}_0 = \alpha(0)$, one has,

$$\begin{aligned} ||\mathbf{h}(t)||_{H^s_x} &\stackrel{(4.4.205)}{=} ||\mathcal{T}_1\mathcal{T}_2\mathcal{T}_3\mathcal{T}_4\Phi_\infty\mathbf{v}(t)||_{H^s_x} \stackrel{(4.5.239)}{\le} C(s)||\mathcal{T}_3\mathcal{T}_4\Phi_\infty\mathbf{v}(t)||_{H^s_x} \\ &= ||\mathcal{T}_4\Phi_\infty\mathbf{v}(\mathfrak{t})||_{H^s_x} \stackrel{(4.5.239)}{\le} C(s)||\mathbf{v}(\mathfrak{t})||_{H^s_x} \stackrel{(4.1.28)}{=} C(s)||\mathbf{v}(\mathfrak{t}_0)||_{H^s_x} \\ &\stackrel{(4.4.205)}{=} C(s)||\Phi_\infty^{-1}\mathcal{T}_4^{-1}\mathcal{T}_3^{-1}\mathcal{T}_2^{-1}\mathcal{T}_1^{-1}\mathbf{h}(\mathfrak{t}_0)||_{H^s_x} \\ &\stackrel{(4.5.239)}{\le} C(s)||\mathcal{T}_3^{-1}\mathcal{T}_2^{-1}\mathcal{T}_1^{-1}\mathbf{h}(\mathfrak{t}_0)||_{H^s_x} = C(s)||\mathcal{T}_2^{-1}\mathcal{T}_1^{-1}\mathbf{h}(0)||_{H^s_x} \\ &\stackrel{(4.5.239)}{\le} C(s)||\mathbf{h}(0)||_{H^s_x}, \end{aligned} \tag{4.5.241}$$

Then the first of (4.5.240) is proved. Following the same procedure, we obtain

$$
\begin{aligned}
||\mathbf{h}(t)||_{H^s_x} &\overset{(4.4.205)}{=} ||\mathcal{T}_1\mathcal{T}_2\mathcal{T}_3\mathcal{T}_4\Phi_\infty\mathbf{v}(t)||_{H^s_x} \leq ||\mathcal{T}_3\mathcal{T}_4\Phi_\infty\mathbf{v}(t)||_{H^s_x} \\
&+ ||(\mathcal{T}_1\mathcal{T}_2 - \mathbb{1})\mathcal{T}_3\mathcal{T}_4\Phi_\infty\mathbf{v}(t)||_{H^s_x} \\
&\overset{(4.5.239)}{\leq} ||\mathbf{v}(\mathfrak{t})||_{H^s_x} + ||(\mathcal{T}_4\Phi_\infty - \mathbb{1})\mathbf{v}(\mathfrak{t})||_{H^s_x} \\
&+ \varepsilon\gamma^{-1}C(s)||\mathcal{T}_4\Phi_\infty\mathbf{v}(\mathfrak{t})||_{H^{s+1}_x} \\
&\overset{(4.1.28),(4.5.239)}{\leq} ||\mathbf{v}(\mathfrak{t}_0)||_{H^s_x} + \varepsilon\gamma^{-1}C(s)||\mathbf{v}(\mathfrak{t}_0)||_{H^{s+1}_x}, \\
&\overset{(4.4.205),(4.5.239)}{\leq} ||\mathbf{h}(0)||_{H^s_x} + \varepsilon\gamma^{-1}C(s)||\mathbf{h}(0)||_{H^{s+1}_x},
\end{aligned}
\tag{4.5.242}
$$

where we used $\mathfrak{t}_0 = \alpha(0)$ and in the last inequality we have performed the same triangular inequalities used in the first two lines only with the $\mathcal{T}_i^{-1}$. Then, using that $\gamma = \varepsilon^a$, with $a \in (0,1)$, we get the second of (4.5.240) with $b = 1 - a$. The first is obtained in the same way. This concludes the proof of Theorem 1.1.1. □

# 5. Forced NLS: Hamiltonian case

In this Chapter we prove Theorem 1.1.2. We already said that the strategy we follow is the same used in Chapter 4. We will see that Proposition 4.0.46 holds also in this case. Here we focus on the study of the linearized operator. Indeed all the differences between the reversible and the Hamiltonian case stand in the inversion of $\mathcal{L}$. In particular we see how the multiplicity of the eigenvalues of NLS on $\mathbb{T}$ complicate the inversion procedure.

The aim of the next sections is to prove the invertibility of the linearized operator. It is more convinient to work on $\mathcal{L}$ since its main part is diagonal. In this contest we prove a analogous result of Proposition 4.1.51 but as we will see it is slightly weaker that the result on the reversible case.

**Proposition 5.0.85 (Diagonalization: Hamiltonian case).** *Fix $\gamma \leq \gamma_0$ and $\tau > d$ and consider any $\mathtt{f} \in C^q$ that satisfies Hypotheses 2 and 3. Then there exist $\eta, q \in \mathbb{N}$, depending only on d, such that for $0 \leq \varepsilon \leq \varepsilon_0$ with $\varepsilon_0$ small enough the following holds. Consider any subset $\Lambda_o \subseteq \Lambda \subseteq \mathbb{R}^d$ and any Lipschitz families $u(\omega) : \Lambda_o \to \mathbf{H}^0$ with $||u||_{\mathfrak{s}_0+\eta,\gamma} \leq 1$. Consider the linear operator $\mathcal{L} : \mathbf{H}^s \to \mathbf{H}^s$ in (2.3.36) computed at u. then for all $\sigma = \pm 2, j \in \mathbb{N}$ there exist Lipschitz map $\Omega_{\sigma,\underline{j}} : \Lambda \to \mathrm{Mat}(2\times 2, \mathbb{C})$ of the form*

$$\Omega_{\sigma,\underline{j}} = -i\sigma(m_2 j^2 + m_0)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} - i\sigma|m_1|j\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} + i\sigma R_{\sigma,\underline{j}}, \tag{5.0.1}$$

*where $R_{\sigma,\underline{j}}$ is a self-adjoint matrix and*

$$\begin{aligned} &|m_2 - 1|_\gamma + |m_0 - m|_\gamma \leq \varepsilon C, \quad |R_j^k|_\gamma \leq \frac{\varepsilon C}{\langle j\rangle}, \quad k = \pm j,\ j \in \mathbb{Z},\\ &\varepsilon c \leq |m_1|^{sup} \leq \varepsilon C, \quad |m_1|^{lip} \leq \varepsilon^2\gamma^{-1}C. \end{aligned} \tag{5.0.2}$$

*for any $\sigma \in \boldsymbol{C}$, $j \in \mathbb{N}\cup\{0\}$, here and in the following $\boldsymbol{C} := \{+1, -1\}$. Set*

$$\Omega_{\sigma,\underline{j}} := \begin{pmatrix}\Omega_{\sigma,j}^{\ j} & \Omega_{\sigma,j}^{-j}\\ \Omega_{\sigma,-j}^{j} & \Omega_{\sigma,-j}^{-j}\end{pmatrix}, \tag{5.0.3}$$

*Define* $\mu_{\sigma,j}$ *and* $\mu_{\sigma,-j}$ *to be the eigenvalues of* $\Omega_{\sigma,\underline{j}}$ *Define* $\Lambda^{2\gamma}_{\infty}(u) := \mathcal{S}^{2\gamma}_{\infty}(u) \cap \mathcal{O}^{2\gamma}_{\infty}(u)$ *with*

$$\begin{aligned} \mathcal{S}^{2\gamma}_{\infty}(u) &:= \left\{ \begin{array}{ll} \omega \in \Lambda_o : & |\omega \cdot \ell + \mu_{\sigma,j}(\omega) - \mu_{\sigma',j'}(\omega)| \geq \frac{2\gamma|\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^{\tau}}, \\ & \ell \in \mathbb{Z}^d, \sigma, \sigma' \in \boldsymbol{C}, j, j' \in \mathbb{Z} \end{array} \right\}, \\ \mathcal{O}^{2\gamma}_{\infty}(u) &:= \left\{ \begin{array}{ll} \omega \in \Lambda_o : & |\omega \cdot \ell + \mu_{\sigma,j} - \mu_{\sigma,k}| \geq \frac{2\gamma}{\langle \ell \rangle^{\tau} \langle j \rangle}, \\ & \ell \in \mathbb{Z}^d \backslash \{0\}, j \in \mathbb{Z}, k = \pm j, \sigma \in \boldsymbol{C} \end{array} \right\}, \end{aligned} \tag{5.0.4}$$

*then we have:*

*(i) for any* $s \in (\mathfrak{s}_0, q - \eta)$*, if* $||z||_{\mathfrak{s}_0 + \eta} < +\infty$ *there exist linear bounded operators* $W_1, W_2 : \mathbf{H}^s(\mathbb{T}^{d+1}) \to \mathbf{H}^s(\mathbb{T}^{d+1})$ *with bounded inverse, such that* $\mathcal{L}(u)$ *satisfies*

$$\mathcal{L}(\mathbf{u}) = W_1 \mathcal{L}_{\infty} W_2^{-1}, \; \mathcal{L}_{\infty} = \omega \cdot \partial_{\varphi} \mathbb{1} + \mathcal{D}_{\infty}, \; \mathcal{D}_{\infty} = diag_{(\sigma,j) \in \boldsymbol{C} \times \mathbb{Z}} \{\Omega_{\sigma,\underline{j}}\}, \tag{5.0.5}$$

*(ii) for any* $\varphi \in \mathbb{T}^d$ *one has*

$$W_i(\varphi), W_i^{-1}(\varphi) : \mathbf{H}^s_x \to \mathbf{H}^s_x, \quad i = 1, 2. \tag{5.0.6}$$

*with* $\mathbf{H}^s_x := H^s(\mathbb{T}; \mathbb{C}) \times H^s(\mathbb{T}; \mathbb{C}) \cap \mathcal{U}$ *and such that*

$$||(W_i^{\pm 1}(\varphi) - \mathbb{1})h||_{\mathbf{H}^s_x} \leq \varepsilon \gamma^{-1} C(s)(||h||_{\mathbf{H}^s_x} + ||u||_{s+\eta+\mathfrak{s}_0} ||h||_{\mathbf{H}^1_x}). \tag{5.0.7}$$

**Remark 5.0.86.** *Note that function* $h(t) \in \mathbf{H}^s_x$ *is a solution of the forced NLS*

$$\mathcal{L}(z)h = 0 \tag{5.0.8}$$

*if and only if the function* $v(t) := (v_1, v_{-1}) := W_2^{-1}(\omega t)[h(t)] \in \mathbf{H}^s_x$ *solves the constant coefficients dynamical system*

$$\begin{pmatrix} \partial_t v_1 \\ \partial_t v_{-1} \end{pmatrix} + \mathcal{D}_{\infty} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \dot{v}_{\sigma,\underline{j}} = -\Omega_{\sigma,\underline{j}} v_{\sigma,\underline{j}}, \quad (\sigma, j) \in \boldsymbol{C} \times \mathbb{Z}, \tag{5.0.9}$$

*where all the eigenvalues of the matrices* $\Omega_{\sigma,\underline{j}}$ *are purely imaginary. Moreover, since* $\overline{\Omega^j_{\sigma,j}} = -\Omega^j_{\sigma,j}$ *and* $\overline{\Omega^{\;j}_{\sigma,-j}} = -\Omega^{\;j}_{\sigma,j}$ *then one has*

$$\frac{d}{dt}(|v_{1,j}(t)|^2 + |v_{1,-j}(t)|^2) = 0, \qquad |v_{\sigma,0}(t)|^2 = constant$$

*and hence*

$$\begin{aligned} ||v_1(t)||^2_{H^s_x} &= \sum_{j \in \mathbb{Z}} |v_{1,j}(t)|^2 \langle j \rangle^{2s} \\ &= |v_{1,0}(t)|^2 + \sum_{j \in \mathbb{N}} (|v_{1,j}(t)|^2 + |v_{1,-j}(t)|^2) \langle j \rangle^{2s} \\ &= |v_{1,0}(0)|^2 + \sum_{j \in \mathbb{N}} (|v_{1,j}(0)|^2 + |v_{1,-j}(0)|^2) \langle j \rangle^{2s} = ||v_1(0)||^2_{H^s_x}. \end{aligned} \tag{5.0.10}$$

*Eq.* (5.0.10) *means that the Sobolev norm in the space of functions depending on* $x$*, is constant in time.*

Proposition 5.0.85 is fundamental in order to prove Theorem 1.1.2. Of course one can try to invert the linearized operator without diagonalize it. In addiction to this we are not able to completely diagonalize it due to the multiplicity of the eigenvalues. This is one of the main difference with respect to the reversible case. Anyway the result in Proposition 5.0.85 is enough to prove the stability of the possible solution. What we obtain is a block-diagonal operator with constant coefficients while in [46] the authors obtain a normal form depending on time. Here most of the problems appear because we want to obtain a constant coefficient linear operator. Another important difference between Proposition 4.1.51 and 5.0.85 stands in the set $\mathcal{O}^{2\gamma}_{\infty}$ in (5.0.4). Indeed, as one can see in (5.0.4), due to the multiplicity of the eigenvalues, we must impose a very weak non degeneracy condition on the eigenvalues. Moreover, as we will see in Section 5.4, the measure estimates in the Hamiltonian case are more difficult with respect to the reversible one, and most of the problems appear due to the presence of the set $\mathcal{O}^{2\gamma}_{\infty}$. In order to overcame such problems we will use the additional Hypotheses 3. As done in Sections 4.2 and 4.3, we first conjugate $\mathcal{L}$ to a differential linear operator with constant coefficients plus a *bounded* remainder, in Section 5.1 and then we complete block-diagonalize the operator in Section 5.2.

## 5.1 Regularization of the linearized operator

In this section and in Section 5.2 we apply a reducibility scheme in order to conjugate the linearized operator to a linear, constant coefficients differential operator. Here we consider the linearized operator $\mathcal{L}$ in (2.3.36) and we construct two operators $\mathcal{V}_1$ and $\mathcal{V}_2$ in order to semi-conjugate $\mathcal{L}$ to an operator $\mathcal{L}_c$ of the second order with constant coefficients plus a remainder of order $O(\partial_x^{-1})$. We look for such transformations because, in order to apply a KAM-type algorithm to diagonalize $\mathcal{L}$, we need first a precise control of the asymptotics of the eigenvalues, and also some estimates of the transformations $\mathcal{V}_i$ with $i = 1, 2$ and their inverse.

The principal result we prove is the following.

**Lemma 5.1.87.** *Let $\mathtt{f} \in C^q$ satisfy the Hypotheses of Proposition 4.0.46 and assume $q > \eta_1 + \mathfrak{s}_0$ where*

$$\eta_1 := d + 2\mathfrak{s}_0 + 10. \tag{5.1.11}$$

*There exists $\epsilon_0 > 0$ such that, if $\varepsilon\gamma_0^{-1} \leq \epsilon_0$ (see (1.1.2 for the definition of $\gamma_0$) then, for any $\gamma \leq \gamma_0$ and for all $u \in \mathbf{H}^0$ depending in a Lipschitz way on $\lambda \in \Lambda$, if*

$$||u||_{\mathfrak{s}_0+\eta_1,\gamma} \leq \varepsilon\gamma^{-1}, \tag{5.1.12}$$

*then, for $\mathfrak{s}_0 \leq s \leq q - \eta_1$, the following holds.*

*(i) There exist invertible maps* $\mathcal{V}_1, \mathcal{V}_2 : \mathbf{H}^0 \to \mathbf{H}^0$ *such that* $\mathcal{L}_7 := \mathcal{V}_1^{-1}\mathcal{L}\mathcal{V}_2 =$

$$\omega \cdot \partial_\varphi \mathbb{1} + i \begin{pmatrix} m_2 & 0 \\ 0 & -m_2 \end{pmatrix} \partial_{xx} + i \begin{pmatrix} m_1 & 0 \\ 0 & -\bar{m}_1 \end{pmatrix} \partial_x + i \begin{pmatrix} m_0 & q_0(\varphi, x) \\ -\bar{q}_0(\varphi, x) & -m_0 \end{pmatrix} + \mathcal{R} \tag{5.1.13}$$

*with* $m_2, m_0 \in \mathbb{R}$, $m_1 \in i\mathbb{R}$ *and* $\mathcal{R}$ *is a pseudo-differential operator of order* $O(\partial_x^{-1})$ *(see* (2.1.7)*). The* $\mathcal{V}_i$ *are symplectic maps and moreover for all* $h \in \mathbf{H}^0$

$$||\mathcal{V}_i h||_{s,\gamma} + ||\mathcal{V}_i^{-1} h||_{s,\gamma} \le C(s)(||h||_{s+2,\gamma} + ||u||_{s+\eta_1,\gamma}||h||_{\mathfrak{s}_0+2,\gamma}), \quad i = 1, 2. \tag{5.1.14}$$

*(ii) The coefficient* $m_i := m_i(u)$ *for* $i = 0, 1, 2$ *of* $\mathcal{L}_7$ *satisfies*

$$\begin{aligned} |m_2(u) - 1|_\gamma, |m_0(u) - \mathtt{m}|_\gamma \le \varepsilon C, \ \ |d_u m_i(u)[h]| \le \varepsilon C ||h||_{\eta_1}, \ \ i = 0, 2, \\ |m_1(u)| \le \varepsilon C, |d_u m_1(u)[h]| \le \varepsilon C ||h||_{\eta_1}, \end{aligned} \tag{5.1.15}$$

*and moreover the constant* $m_1 := m_1(\omega, u(\omega))$ *satisfies*

$$\varepsilon c \le |m_1(u)|, \tag{5.1.16a}$$

$$\sup_{\omega_1 \neq \omega_2} \frac{|m_1(\omega_1, u(\omega)) - m_1(\omega_2, u(\omega))|}{|\lambda_1 - \lambda_2|} \le \varepsilon^2 C \gamma^{-1} \tag{5.1.16b}$$

*for some* $C > 0$.

*(iii) The operator* $\mathcal{R} := \mathcal{R}(u)$ *is such that*

$$\|\mathcal{R}(u)h\|_{s,\gamma} \le \varepsilon C(s)(\|h\|_{s,\gamma} + \|u\|_{s+\eta_1,\gamma}\|h\|_{\mathfrak{s}_0}), \tag{5.1.17}$$

$$\begin{aligned} \|d_u\mathcal{R}(u)[h]g\|_s \ \le\ & \varepsilon C(s)\big(\|g\|_{s+1}\|h\|_{\mathfrak{s}_0+\eta_1} + \|g\|_2\|h\|_{s+\eta_1} \\ +\ & \|u\|_{s+\eta_1}\|g\|_2\|h\|_{\mathfrak{s}_0}\big), \end{aligned} \tag{5.1.18}$$

*and moreover*

$$\|q_0\|_{s,\gamma} \le \varepsilon C(s)(1 + \|u\|_{s+\eta_1,\gamma}), \tag{5.1.19a}$$

$$\|d_u q_0(u)[h]\|_s \le \varepsilon C(s)(\|h\|_{s+\eta_1} + \|u\|_{s+\eta_1} + \|h\|_{\mathfrak{s}_0+\eta_1}), \tag{5.1.19b}$$

*Finally* $\mathcal{L}_7$ *is Hamiltonian.*

**Remark 5.1.88.** *The estimate in* (5.1.16) *is different from that in* (5.1.15)*. As we will see, it is very important to estimate the Lipschitz norm of the constant* $m_1$ *in order to get the measure estimates in Section 6. The constant* $m_1$ *depends in* $\lambda$ *in two way: the first is trough the dependence on* $\omega$ *of the function* $u$*; secondly it presents also an explicit dependence on the external parameters. Clearly by* (5.1.15) *we can get a bound only on the variation* $|m_1(\omega, u(\omega_1)) - m_1(\omega, u(\omega_2))|$*. To estimate the* $|\cdot|^{lip}$ *seminorm we need also the* (5.1.16)*.*

We do not give the proof of Lemma 5.1.87 apart form (5.1.16) because the strategy is very similar to the ones used in Lemma 4.2.53 in Section 4.2. At each step we construct a transformation $\mathcal{T}_i$ that conjugates $\mathcal{L}_i$ to $\mathcal{L}_{i+1}$. We fix $\mathcal{L}_0 = \mathcal{L}$. Moreover the $\mathcal{T}_i$ are symplectic, hence $\mathcal{L}_i$ is Hamiltonian and has the form

$$\mathcal{L}_i := \omega \cdot \partial_\varphi \mathbb{1} + i(E + A_2^{(i)})\partial_{xx} + iA_1^{(i)}\partial_x + i(mE + A_0^{(i)}) + \mathcal{R}_i, \tag{5.1.20}$$

with $E$ defined in (2.3.37),

$$A_j^{(i)} = A_j^{(i)}(\varphi, x) := \begin{pmatrix} a_j^{(i)} & b_j^{(i)} \\ -\bar{b}_j^{(i)} & -\bar{a}_j^{(i)} \end{pmatrix}, \quad j = 0, 1, 2 \tag{5.1.21}$$

and $\mathcal{R}_i$ is a pseudo-differential operator of order $\partial_x^{-1}$. Essentially we need to prove bounds like

$$\|(\mathcal{T}_i^{\pm 1}(u) - \mathbb{1})h\|_{s,\gamma} \le \varepsilon C(s)(\|h\|_{s,\gamma} + \|u\|_{s+\kappa_i,\gamma}\|h\|_{\mathfrak{s}_0}), \tag{5.1.22}$$

$$\begin{aligned} \|d_u(\mathcal{T}_i^{\pm 1})(u)[h]g\|_s \le\ & \varepsilon C(s)\big(\|g\|_{s+1}\|h\|_{\mathfrak{s}_0+\kappa_i} + \|g\|_2\|h\|_{s+\kappa_i} \\ & + \|u\|_{s+\kappa_i}\|g\|_2\|h\|_{\mathfrak{s}_0}\big), \end{aligned} \tag{5.1.23}$$

for suitable $\kappa_i$ and on the coefficients in (5.1.20) we need

$$\|a_j^{(i)}(u)\|_{s,\gamma}, \|b_j^{(i)}(u)\|_{s,\gamma} \le \varepsilon C(s)(1 + \|u\|_{s+\kappa_i,\gamma}), \tag{5.1.24a}$$

$$\|d_u a_j^{(i)}(u)[h]\|_s, \|d_u b_j^{(i)}(u)[h]\|_s \le \varepsilon C(s)(\|h\|_{s+\kappa_i} + \|u\|_{s+\kappa_i} + \|h\|_{\mathfrak{s}_0+\kappa_i}), \tag{5.1.24b}$$

for $j = 0, 1, 2$ and $i = 1, \dots, 7$ and on $\mathcal{R}_i$ bounds like (5.1.17) with $\kappa_i$ instead of $\eta_1$.

The bounds are based on repeated use of classical tame bounds and interpolation estimates of the Sobolev norms. The proof of such properties of the norm can be found in [5] in Appendix A. To conclude one combine the bounds of each transformation to obtain estimates on the compositions. It turn out that the constant $\eta_1$ contains all the loos of regularity of each step. We present only the construction of the transformation that, in the Hamiltonian case, are more involved. Moreover the difference between Lemma 5.1.87 and Lemma 4.2.53 (see [31]) is also in equation (5.1.16). Indeed, in this case we need to prove that non degeneracy hypothesis 3 persists during the steps in order to obtain the same lower bound (possibly with a worse constant) for the constant $m_1$ in (5.1.15). This fact will be used in Section 6 in order to perform measure estimates.

**Step 1. Diagonalization of the second order coefficient** In this section we want to diagonalize the second order term $(E + A_2)$ in (2.3.36). By a direct calculation one can see that the matrix $(E + A_2)$ has eigenvalues $\lambda_{1,2} := \sqrt{(1 + a_2)^2 - |b_2|^2}$. if we set $a_2^{(1)} := \lambda_1 - 1$ we have that $a_2^{(1)} \in \mathbb{R}$ since $a_2 \in \mathbb{R}$

for any $(\varphi, x) \in \mathbb{T}^{d+1}$ and $a_i, b_i$ are small. We define the transformation $\mathcal{T}_1^{-1} : \mathbf{H}^0 \to \mathbf{H}^0$ as the matrix $\mathcal{T}_1^{-1} = \big((\mathcal{T}_1^{-1})_\sigma^{\sigma'}\big)_{\sigma,\sigma'=\pm 1}$ with

$$\mathcal{T}_1^{-1} := \begin{pmatrix} (2 + a_2 + a_2^{(1)})(i\lambda_0)^{-1} & b_2(i\lambda_0)^{-1} \\ -\bar{b}_2(i\lambda_0)^{-1} & -(2 + a_2 + a_2^{(1)})(i\lambda_0)^{-1} \end{pmatrix}, \tag{5.1.25}$$

where $\lambda_0 := i\sqrt{2\lambda_1(1 + a_2 + \lambda_1)}$. Note that $\det\mathcal{T}_1^{-1} = 1$. One has that

$$\mathcal{T}_1^{-1}(E + A_2)\mathcal{T}_1 = \begin{pmatrix} 1 + a_2^{(1)}(\varphi, x) & 0 \\ 0 & -1 - a_2^{(1)}(\varphi, x) \end{pmatrix}. \tag{5.1.26}$$

Moreover, we have that the transformation is symplectic. We can think that $\mathcal{T}_1$ act on the function of $H^s(\mathbb{T}^{d+1}; \mathbb{C})$ is the following way. Set $U = (u, \bar{u}), V = (v, \bar{v}) \in \mathrm{H}^s$ and let $(MZ)_\sigma$ for $\sigma \in \{+1, -1\}$ be the first or the second (respectively) component. Given a function $u \in H^s(\mathbb{T}^{d+1}; \mathbb{C})$ we define, with abuse of notation, $\mathcal{T}_1^{-1}u := (\mathcal{T}_1^{-1}U)_{+1} := ((\mathcal{T}_1^{-1})_1^1)u + ((\mathcal{T}_1^{-1})_1^{-1})\bar{u}$. With this notation one has that

$$\begin{aligned}
\Omega\left(\mathcal{T}_1^{-1}u, \mathcal{T}_1^{-1}v\right) := \mathrm{Re}\int_{\mathbb{T}} & i\left((\mathcal{T}_1^{-1})_1^1(\mathcal{T}_1^{-1})_{-1}^1 uv + (\mathcal{T}_1^{-1})_1^{-1}(\mathcal{T}_1^{-1})_{-1}^{-1}\bar{u}\bar{v}\right) \\
& + i\left((\mathcal{T}_1^{-1})_1^1(\mathcal{T}_1^{-1})_{-1}^{-1}u\bar{v} + (\mathcal{T}_1^{-1})_1^{-1}(\mathcal{T}_1^{-1})_{-1}^1\bar{u}v\right)\mathrm{d}x \\
= \mathrm{Re}\int_{\mathbb{T}} & i\mathrm{Re}((\mathcal{T}_1^{-1})_1^1(\mathcal{T}_1^{-1})_{-1}^1 uv) \\
& + i\left((\mathcal{T}_1^{-1})_1^{-1}(\mathcal{T}_1^{-1})_{-1}^1(u\bar{v} + \bar{u}v)\right) \\
& + i\left((\mathcal{T}_1^{-1})_1^1(\mathcal{T}_1^{-1})_{-1}^{-1} - (\mathcal{T}_1^{-1})_1^{-1}(\mathcal{T}_1^{-1})_{-1}^1\right)u\bar{v}\mathrm{d}x \\
= \mathrm{Re}\int_{\mathbb{T}} & iu\bar{v}\mathrm{d}x =: \Omega(u, v).
\end{aligned}$$

which implies that $\mathcal{T}_1^{-1}$ is symplectic.

Now we can conjugate the operator $\mathcal{L}$ to an operator $\mathcal{L}_1$ with a diagonal coefficient of the second order spatial differential operator. Indeed, one has

$$\begin{aligned}
\mathcal{L}_1 := \mathcal{T}_1^{-1}\mathcal{L}\mathcal{T}_1 &= \omega \cdot \partial_\varphi \mathbb{1} + i\mathcal{T}_1^{-1}(E + A_2)\mathcal{T}_1\partial_{xx} \\
&+ i(2\mathcal{T}_1^{-1}(E + A_2)\partial_x\mathcal{T}_1 + \mathcal{T}_1^{-1}A_1\mathcal{T}_1)\partial_x \\
&+ i\left[-i\mathcal{T}_1^{-1}(\omega \cdot \partial_\varphi\mathcal{T}_1) + \mathcal{T}_1^{-1}(E + A_2)\partial_{xx}\mathcal{T}_1\right. \\
&\left.+\mathcal{T}_1^{-1}A_1\partial_x\mathcal{T}_1 + \mathcal{T}_1^{-1}(mE + A_0)\mathcal{T}_1\right];
\end{aligned} \tag{5.1.27}$$

the (5.1.27) has the form (5.1.20). This identify uniquely the coefficients $a_j^{(1)}, b_j^{(1)}$ for $j = 0, 1, 2$ and $\mathcal{R}_1$. In particular we have that $b_2^{(1)} \equiv 0$ and $\mathcal{R}_1 \equiv 0$. Moreover, since the transformation is symplectic, then the new operator $\mathcal{L}_1$ is Hamiltonian, with an Hamiltonian function

$$\begin{aligned} H_1(u,\bar{u}) = &\int_{\mathbb{T}} (1+a_2^{(1)})|u_x|^2 - \frac{i}{2}\mathrm{Im}(a_1^{(1)})(u_x\bar{u} - u\bar{u}_x) - \mathrm{Re}(a_0^{(1)})|u|^2 \mathrm{d}x \\ &+ \int_{\mathbb{T}} -m|u|^2 - \frac{1}{2}(b_0^{(1)}\bar{u}^2 + \bar{b}_0^{(1)}u^2)\mathrm{d}x := \int_{\mathbb{T}} f_1(\varphi,x,u,\bar{u},u_x,\bar{u}_x)\mathrm{d}x, \end{aligned} \tag{5.1.28}$$

hence, since $f_1$ depends only linearly on $\bar{u}_x$, one has by (2.1.3),

$$b_1^{(1)}(\varphi,x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\partial_{\bar{z}_1\bar{z}_1} f_1\right) \equiv 0. \tag{5.1.29}$$

This means that we have diagonalized also the matrix of the first order spatial differential operator.

**Remark 5.1.89.** *It is important to note that $a_1^{(1)}(\varphi,x)$ as the form*

$$a_1^{(1)}(\varphi,x) = \frac{\mathrm{d}}{\mathrm{d}x} a_2^{(1)}(\varphi,x) + \partial_{z_0\bar{z}_1} f_1 - \partial_{z_1\bar{z}_0} f_1$$

*so that the real part of $a_1^{(1)}$ depends only on the spatial derivative of $a_2^{(1)}$.*

**Step 2. Change of the space variable** We consider a $\varphi-$dependent family of diffeomorphisms of the $1-$dimensional torus $\mathbb{T}$ of the form

$$y = x + \xi(\varphi,x), \tag{5.1.30}$$

where $\xi$ is as small real-valued funtion, $2\pi$ periodic in all its arguments. We define the change of variables on the space of functions as

$$\begin{aligned} (\mathcal{T}_2 h)(\varphi,x) &:= \sqrt{1+\xi_x(\varphi,x)}\, h(\varphi, x+\xi(\varphi,x)), \text{ with inverse} \\ (\mathcal{T}_2^{-1} v)(\varphi,y) &:= \sqrt{1+\widehat{\xi}_x(\varphi,y)}\, v(\varphi, y+\widehat{\xi}(\varphi,y)) \end{aligned} \tag{5.1.31}$$

where

$$x = y + \widehat{\xi}(\varphi,y), \tag{5.1.32}$$

is the inverse diffeomorphism of (5.1.30). With a slight abuse of notation we extend the operator to $\mathbf{H}^s$:

$$\mathcal{T}_2 : \mathbf{H}^s \to \mathbf{H}^s, \quad \mathcal{T}_2 \begin{pmatrix} h \\ \bar{h} \end{pmatrix} = \begin{pmatrix} (\mathcal{T}_2 h)(\varphi,x) \\ (\mathcal{T}_2 \bar{h})(\varphi,x) \end{pmatrix}. \tag{5.1.33}$$

Now we have to calculate the conjugate $\mathcal{T}_2{}^{-1}\mathcal{L}_1\mathcal{T}_2$ of the operator $\mathcal{L}_1$ in (5.1.27).
The conjugate $\mathcal{T}_2^{-1} a \mathcal{T}_2$ of any multiplication operator $a : h(\varphi,x) \to a(\varphi,x)h(\varphi,x)$ is the multiplication operator

$$v(\varphi,y) \mapsto (\mathcal{T}_2^{-1} a\sqrt{1+\xi_x})(\varphi,y)v(\varphi,y) = a(\varphi, y+\widehat{\xi}(\varphi,y))v(\varphi,y). \tag{5.1.34}$$

In (5.1.34) we have used the relation

$$0 \equiv \xi_x(\varphi, x) + \widehat{\xi}_y(\varphi, y) + \xi_x(\varphi, x)\widehat{\xi}_y(\varphi, y), \tag{5.1.35}$$

that follow by (5.1.30) and (5.1.32). The conjugate of the differential operators will be

$$\begin{aligned} \mathcal{T}_2^{-1}\omega \cdot \partial_\varphi \mathcal{T}_2 &= \omega \cdot \partial_\varphi + [\mathcal{T}_2^{-1}(\omega \cdot \partial_\varphi \xi)]\partial_y - \mathcal{T}_2^{-1}\left(\frac{\omega \cdot \partial_\varphi \xi_x}{2(1+\xi_x)}\right), \\ \mathcal{T}_2^{-1}\partial_x \mathcal{T}_2 &= [\mathcal{T}_2^{-1}(1+\xi_x)]\partial_y - \mathcal{T}_2^{-1}\left(\frac{\xi_{xx}}{2(1+\xi_x)}\right), \\ \mathcal{T}_2^{-1}\partial_{xx} \mathcal{T}_2 &= [\mathcal{T}_2^{-1}(1+\xi_x)^2]\partial_{yy} - \mathcal{T}_2^{-1}\left(\frac{2\xi_{xxx}+\xi_{xx}^2}{4(1+\xi_x)^2}\right), \end{aligned} \tag{5.1.36}$$

where all the coefficients are periodic functions of $(\varphi, x)$. Thus, by conjugation, we have that $\mathcal{L}_2 = \mathcal{T}_2^{-1}\mathcal{L}_1\mathcal{T}_2$ has the form (5.1.20) with

$$\begin{aligned} &1 + a_2^{(2)}(\varphi, y) = \mathcal{T}_2^{-1}[(1 + a_2^{(1)})(1+\xi_x)^2], \\ &a_1^{(2)}(\varphi, y) = \mathcal{T}_2^{-1}(a_1^{(1)}(1+\xi_x)) - i\mathcal{T}_2^{-1}(\omega \cdot \partial_\varphi \xi), \\ &a_0^{(2)}(\varphi, y) = i\mathcal{T}_2^{-1}\left(\frac{\omega \cdot \partial_\varphi \xi_x}{2(1+\xi_x)}\right) - \mathcal{T}_2^{-1}\left(\frac{\xi_{xx}}{2(1+\xi_x)}\right) - \mathcal{T}_2^{-1}\left(\frac{2\xi_{xxx}+\xi_{xx}^2}{4(1+\xi_x)^2}\right), \\ &b_0^{(2)}(\varphi, y) = \mathcal{T}_2^{-1}(b_0^{(1)}), \end{aligned} \tag{5.1.37}$$

and $b_2^{(2)} = b_1^{(2)} = 0$. We are looking for $\xi(\varphi, x)$ such that the coefficient $a_2^{(2)}(\varphi, y)$ does not depend on $y$, namely

$$1 + a_2^{(2)}(\varphi, y) = \mathcal{T}_2^{-1}[(1 + a_2^{(1)})(1+\xi_x)^2] = 1 + a_2^{(2)}(\varphi), \tag{5.1.38}$$

for some function $a_2^{(2)}(\varphi)$. Since $\mathcal{T}_2$ operates only on the space variables, the (5.1.38) is equivalent to

$$(1 + a_2^{(1)}(\varphi, x))(1 + \xi_x(\varphi, x))^2 = 1 + a_2^{(2)}(\varphi). \tag{5.1.39}$$

Hence we have to set

$$\xi_x(\varphi, x) = \rho_0, \qquad \rho_0(\varphi, x) := (1 + a_2^{(2)}(\varphi))^{\frac{1}{2}}(\varphi)(1 + a_2^{(1)}(\varphi, x))^{-\frac{1}{2}} - 1, \tag{5.1.40}$$

that has solution $\gamma$ periodic in $x$ if and only if $\int_{\mathbb{T}} \rho_0 dy = 0$. This condition implies

$$a_2^{(2)}(\varphi) = \left(\frac{1}{2\pi}\int_{\mathbb{T}} (1 + a_2^{(1)}(\varphi, x))^{-\frac{1}{2}}\right)^{-2} - 1. \tag{5.1.41}$$

Then we have the solution (with zero average) of (5.1.40)

$$\xi(\varphi, x) := (\partial_x^{-1}\rho_0)(\varphi, x), \tag{5.1.42}$$

where $\partial_x^{-1}$ is defined by linearity as

$$\partial_x^{-1} e^{ikx} := \frac{e^{ikx}}{ik}, \quad \forall\, k \in \mathbb{Z}\backslash\{0\}, \quad \partial_x^{-1} = 0. \tag{5.1.43}$$

In other word $\partial_x^{-1} h$ is the primitive of $h$ with zero average in $x$. Moreover, the map $\mathcal{T}_2$ is canonical with respect to the $NLS-$symplectic form, indeed, for any $u, v \in H^s(\mathbb{T}^{d+1};\mathbb{C})$,

$$\begin{aligned}\Omega(\mathcal{T}_2 u, \mathcal{T}_2 v) =& \mathrm{Re}\int_{\mathbb{T}} (i\sqrt{1+\xi_x}u(\varphi, x+\varphi(\varphi,x)))\sqrt{1+\xi_x}\bar{v}(\varphi, x+\varphi(\varphi,x))\mathrm{d}x\\ =& \mathrm{Re}\int_{\mathbb{T}} (1+\xi_x(\varphi,x))(iu(\varphi, x+\xi(\varphi,x)))\bar{v}(\varphi, x+\xi(\varphi,x))\mathrm{d}x\\ =& \mathrm{Re}\int_{\mathbb{T}} (iu(\varphi,y))\bar{v}(\varphi,y)\mathrm{d}y =: \Omega(u,v).\end{aligned}$$

Thus, conjugating $\mathcal{L}_1$ through the operator $\mathcal{T}_2$ in (5.1.33) we obtain the Hamiltonian operator $\mathcal{L}_2 = \mathcal{T}_2^{-1}\mathcal{L}_1\mathcal{T}_2$ with Hamiltonian function given by

$$\begin{aligned}H_2(u,\bar{u}) =& \int_{\mathbb{T}} (1+a_2^{(2)}(\varphi))|u_x|^2 - \frac{i}{2}\mathrm{Im}(a_1^{(2)})(u_x\bar{u} - u\bar{u}_x) - \mathrm{Re}(a_0^{(2)})|u|^2\mathrm{d}x\\ &+\int_{\mathbb{T}} -m|u|^2 - \frac{1}{2}(b_0^{(2)}\bar{u}^2 + \bar{b}_0^{(2)}u^2)\mathrm{d}x := \int_{\mathbb{T}} f_2(\varphi, x, u, \bar{u}, u_x, \bar{u}_x)\mathrm{d}x,\end{aligned} \tag{5.1.44}$$

**Remark 5.1.90.** *As in Remark 5.1.89, the real part of coefficients $a_1^{(2)}$ depends on the spatial derivatives of $a_2^{(2)}$, then in this case, again thanks the Hamiltonian structure of the problem, one has that $a_1^{(2)}(\varphi,y) = i\mathrm{Im}(a_1^{(2)})(\varphi,y)$, i.e. it is purely imaginary. Moreover $b_2^{(2)} = b_1^{(2)} \equiv 0$ and $\mathcal{R}_2 \equiv 0$.*

**Step 3: Time reparametrization** In this section we want to make constant the coefficient of the highest order spatial derivative operator $\partial_{yy}$ of $\mathcal{L}_2$, by a quasi-periodic reparametrization of time. We consider a diffeomorphism of the torus $\mathbb{T}^d$ of the form

$$\theta = \varphi + \omega\alpha(\varphi), \quad \varphi \in \mathbb{T}^d, \quad \alpha(\varphi) \in \mathbb{R}, \tag{5.1.45}$$

where $\alpha$ is a small real valued function, $2\pi-$periodic in all its arguments. The induced linear operator on the space of functions is

$$(\mathcal{T}_3 h)(\varphi, y) := h(\varphi + \omega\alpha(\varphi), y), \tag{5.1.46}$$

whose inverse is

$$(\mathcal{T}_3^{-1} v)(\theta, y) = v(\theta + \omega\widetilde{\alpha}(\theta), y), \tag{5.1.47}$$

where $\varphi = \theta + \omega\widetilde{\alpha}(\theta)$ is the inverse diffeomorphism of $\theta = \varphi + \omega\alpha(\varphi)$. We extend the operator to $\mathbf{H}^s$:

$$\mathcal{T}_3 : \mathbf{H}^s \to \mathbf{H}^s, \quad \mathcal{T}_3\begin{pmatrix} h \\ \bar{h}\end{pmatrix} = \begin{pmatrix} (\mathcal{T}h)(\varphi,x) \\ (\mathcal{T}_3\bar{h})(\varphi,x)\end{pmatrix}. \tag{5.1.48}$$

By conjugation, we have that the differential operator become

$$\mathcal{T}_3^{-1}\omega\cdot\partial_\varphi\mathcal{T}_3=\rho(\theta)\omega\cdot\partial_\theta,\ \ \mathcal{T}_3^{-1}\partial_y\mathcal{T}_3=\partial_y,\ \ \rho(\theta):=\mathcal{T}_3^{-1}(1+\omega\partial_\varphi\alpha). \tag{5.1.49}$$

Hence we have $\mathcal{T}_3^{-1}\mathcal{L}_2\mathcal{T}_3=\rho\mathcal{L}_3$ where $\mathcal{L}_3$ has the form (5.1.20) and

$$\begin{aligned}1+a_i^{(3)}(\theta)&:=\frac{(\mathcal{T}_3^{-1}(1+a_i^{(2)}))(\theta)}{\rho(\theta)},\quad a_i^{(3)}(\theta):=\frac{(\mathcal{T}_3^{-1}a_i^{(2)})(\theta)}{\rho(\theta)},\ \ i=0,1,\\ b_0^{(3)}(\theta,y)&:=\frac{(\mathcal{T}_3^{-1}b_0^{(2)})(\theta,y)}{\rho(\theta)},\end{aligned} \tag{5.1.50}$$

We look for solution $\alpha$ such that the coefficient $a_2^{(3)}$ is constant in time, namely

$$(\mathcal{T}_3^{-1}(1+a_2^{(2)}))(\theta)=m_2\rho(\theta)=m_2\mathcal{T}_3^{-1}(1+\omega\cdot\partial_\varphi\alpha) \tag{5.1.51}$$

for some constant $m_2$, that is equivalent to require that

$$1+a_2^{(2)}(\varphi)=m_2(1+\omega\cdot\partial_\varphi\alpha(\varphi)), \tag{5.1.52}$$

By setting

$$m_2=\frac{1}{(2\pi)^d}\int_{\mathbb{T}^d}(1+a_2^2(\varphi))d\varphi, \tag{5.1.53}$$

we can find the (unique) solution of (5.1.52) with zero average

$$\alpha(\varphi):=\frac{1}{m_2}(\omega\cdot\partial_\varphi)^{-1}(1+a_2^{(2)}-m_2)(\varphi), \tag{5.1.54}$$

where $(\omega\cdot\partial_\varphi)^{-1}$ is defined by linearity

$$(\omega\cdot\partial_\varphi)^{-1}e^{i\ell\cdot\varphi}:=\frac{e^{i\ell\cdot\varphi}}{i\omega\cdot\ell},\ \ell\neq0,\quad(\omega\cdot\partial_\varphi)^{-1}1=0.$$

Moreover, the operator $\mathcal{T}_3$ acts only on the time variables, then it is clearly symplectic, since

$$\Omega(\mathcal{T}_3u,\mathcal{T}_3v)=\Omega(u,v).$$

Then the operator $\mathcal{L}_3$ is Hamiltonian with hamiltonian function $H_3$

$$\begin{aligned}H_3(u,\bar u)&=\int_{\mathbb{T}}m_2|u_x|^2-\frac{i}{2}\mathrm{Im}(a_1^{(3)})(u_x\bar u-u\bar u_x)-\mathrm{Re}(a_0^{(3)})|u|^2\mathrm{d}x\\&+\int_{\mathbb{T}}-\frac12(b_0^{(3)}\bar u^2+\bar b_0^{(3)}u^2)\mathrm{d}x:=\int_{\mathbb{T}}f_3(\varphi,x,u,\bar u,u_x,\bar u_x)\mathrm{d}x,\end{aligned} \tag{5.1.55}$$

**Remark 5.1.91.** *Also in this case, thanks to the hamiltonian structure of the operator, we have that the coefficient* $a_1^{(3)}\in i\mathbb{R}$, $b_2^{(3)}=b_1^{(3)}\equiv0$ *and* $\mathcal{R}_3\equiv0$.

**Step 4. Change of space variable (translation)** The goal of this section, is to conjugate $\mathcal{L}_3$ in (5.1.20) with coefficients in (5.1.50) to an operator in which the coefficients of the first order spatial derivative operator, has zero average in $y$.

Consider the change of the space variable

$$z = y + \beta(\theta) \tag{5.1.56}$$

which induces the operators on functions

$$\mathcal{T}_4 h(\theta, y) := h(\theta, y + \beta(\theta)), \quad \mathcal{T}_4^{-1} v(\theta, z - \beta(\theta)). \tag{5.1.57}$$

We extend the operator $\mathcal{T}_4$ to $\mathbf{H}^s$ as

$$\mathcal{T}_4 \begin{pmatrix} h \\ \bar{h} \end{pmatrix} = \begin{pmatrix} (\mathcal{T}_4 h)(\theta, y) \\ (\mathcal{T}_4 \bar{h})(\theta, y) \end{pmatrix}. \tag{5.1.58}$$

By conjugation, the differential operators become

$$\mathcal{T}_4^{-1} \omega \cdot \partial_\theta \mathcal{T}_4 = \omega \cdot \partial_\theta + (\omega \cdot \partial_\theta \beta(\theta)) \partial_z, \qquad \mathcal{T}_4^{-1} \partial_y \mathcal{T}_4 = \partial_z. \tag{5.1.59}$$

Hence one has that $\mathcal{L}_4 := \mathcal{T}_4^{-1} \mathcal{L}_3 \mathcal{T}_4$ has the form (5.1.20) where

$$\begin{aligned} a_1^{(4)}(\theta, z) &:= -i\omega \cdot \partial_\theta \beta(\theta) + (\mathcal{T}_4^{-1} a_1^{(3)})(\theta, z), \\ a_0^{(4)}(\theta, z) &:= (\mathcal{T}_4^{-1} a_0^{(3)})(\theta, z), \quad b_0^{(4)}(\theta, z) := (\mathcal{T}_4^{-1} b_0^{(4)})(\theta, z). \end{aligned} \tag{5.1.60}$$

The aim is to find a function $\beta(\theta)$ such that

$$\frac{1}{2\pi} \int_{\mathbb{T}} a_1^{(4)}(\theta, z) \mathrm{d}z = m_1, \quad \forall\, \theta \in \mathbb{T}^d, \tag{5.1.61}$$

for some constant $m_1 \in \mathbb{C}$, independent on $\theta$. By using the (5.1.60) we have that the (5.1.61) become

$$- i\omega \cdot \partial_\theta \beta(\theta) = m_1 - \int_{\mathbb{T}} a_1^{(3)}(\theta, y) \mathrm{d}y =: V(\theta). \tag{5.1.62}$$

This equation has a solution periodic in $\theta$ if and only if $V(\theta)$ has zero average in $\theta$. So that we have to define

$$m_1 := \frac{1}{(2\pi)^{d+1}} \int_{\mathbb{T}^{d+1}} a_1^{(3)}(\theta, y) \mathrm{d}\theta \mathrm{d}y. \tag{5.1.63}$$

Note also that $m_1 \in i\mathbb{R}$ (see Remark 5.1.91). Then the function $V$ is purely imaginary. Now we can set

$$\beta(\theta) := i(\omega \cdot \partial_\theta)^{-1} V(\theta), \tag{5.1.64}$$

to obtain a real diffeomorphism of the torus $y+\beta(\theta)$. Morover one has, for any $u,v\in H^s(\mathbb{T}^{d+1};\mathbb{C})$

$$\Omega(\mathcal{T}_4 u,\mathcal{T}_4 v)=\mathrm{Re}\int_{\mathbb{T}} iu(\varphi,x+\beta(\varphi))\bar{v}(\varphi,x+\beta(\varphi))=\Omega(u,v), \tag{5.1.65}$$

hence $\mathcal{T}_4$ is symplectic. This implies that $\mathcal{L}_4$ is Hamiltonian with hamiltonian function of the form

$$\begin{aligned} H_4(u,\bar{u})=&\int_{\mathbb{T}} m_2|u_x|^2-\frac{i}{2}\mathrm{Im}(a_1^{(4)})(u_x\bar{u}-u\bar{u}_x)-\mathrm{Re}(a_0^{(4)})|u|^2-m|u|^2\mathrm{d}x\\ &+\int_{\mathbb{T}}-\frac{1}{2}(b_0^{(4)}\bar{u}^2+\bar{b}_0^{(4)}u^2)\mathrm{d}x:=\int_{\mathbb{T}} f_4(\varphi,x,u,\bar{u},u_x,\bar{u}_x)\mathrm{d}x, \end{aligned} \tag{5.1.66}$$

**Remark 5.1.92.** *Again one has $b_2^{(4)}=b_1^{(4)}\equiv 0$ and $\mathcal{R}_4\equiv 0$.*

For simplicity we rename the variables $z=x$ and $\theta=\varphi$.

**Step 5. Descent Method: conjugation by multiplication operator** In this section we want to eliminate the dependance on $\varphi$ and $x$ on the coefficient $c_9$ of the operator $\mathcal{L}_4$. To do this, we consider an operator of the form

$$\mathcal{T}_5:=\begin{pmatrix}1+z(\varphi,x) & 0\\ 0 & 1+\bar{z}(\varphi,x)\end{pmatrix}, \tag{5.1.67}$$

where $z:\mathbb{T}^{d+1}\to\mathbb{C}$. By a direct calculation we have that

$$\begin{aligned} \mathcal{L}_4\mathcal{T}_5-\mathcal{T}_5&\left[\omega\cdot\partial_\varphi\mathbb{1}+i\begin{pmatrix}m_2 & 0\\ 0 & -m_2\end{pmatrix}\partial_{xx}+i\begin{pmatrix}m_1 & 0\\ 0 & -\overline{m}_1\end{pmatrix}\partial_x\right]=\\ &=i\begin{pmatrix}r_1(\varphi,x) & 0\\ 0 & -\bar{r}_1(\varphi,x)\end{pmatrix}\partial_x+i\begin{pmatrix}m+c(\varphi,x) & d(\varphi,x)\\ -\bar{d}(\varphi,x) & -m-\bar{c}(\varphi,x)\end{pmatrix} \end{aligned} \tag{5.1.68}$$

where

$$\begin{aligned} r_1(\varphi,x)&:=2mz_x(\varphi,x)+(a_1^{(4)}(\varphi,x)-m_1)(1+z(\varphi,x)),\\ c(\varphi,x)&:=-i(\omega\cdot\partial_\varphi z)(\varphi,x)+a_0^{(4)}(\varphi,x)(1+z(\varphi,x)),\\ d(\varphi,x)&:=b_0^{(4)}(\varphi,x)(1+\bar{z}(\varphi,x)). \end{aligned} \tag{5.1.69}$$

We look for $z(\varphi,x)$ such that $r_1\equiv 0$. If we look for solutions of the form $1+z(\varphi,x)=\exp(s(\varphi,x))$ we have that $r_1=0$ become

$$2m_2s_x+a_1^{(4)}-m_1=0, \tag{5.1.70}$$

that has solution

$$s(\varphi,x):=\frac{1}{2m}\partial_x^{-1}(a_1^{(4)}-m_1)(\varphi,x) \tag{5.1.71}$$

where $\partial_x^{-1}$ is defined in (5.1.43). Moreover, since $a_1^{(4)} \in i\mathbb{R}$, one has that $s(\varphi, x) \in i\mathbb{R}$. Clearly the operator $\mathcal{T}_5$ is invertible for $\varepsilon$ small, then we obtain $\mathcal{L}_5 := \mathcal{T}_5^{-1}\mathcal{L}_4\mathcal{T}_5$ with

$$\mathcal{L}_5 := \omega \cdot \partial_\varphi \mathbb{1} + i \begin{pmatrix} m_2 & 0 \\ 0 & -m_2 \end{pmatrix} \partial_{xx} + i \begin{pmatrix} m_1 & 0 \\ 0 & -\bar{m}_1 \end{pmatrix} \partial_x + imE + iA_0^{(5)} \tag{5.1.72}$$

that has the form (5.1.20) with $m_2$ and $m_1$ are defined respectively in (5.1.53) and (5.1.63), while the coefficients of $A_5^{(i)}$ are

$$\begin{aligned} a_0^{(5)}(\varphi, x) &:= (1 + z(\varphi, x))^{-1} c(\varphi, x), \\ b_0^{(5)}(\varphi, x) &:= (1 + z(\varphi, x))^{-1} d(\varphi, x). \end{aligned} \tag{5.1.73}$$

It remains to check that the transformation $\exp(s(\varphi, x)) = 1 + z$ is symplectic. One has

$$\Omega(e^s u, e^s v) = \mathrm{Re} \int_{\mathbb{T}} i e^{s(\varphi,x)} u(\varphi, x) e^{-s(\varphi,x)} \bar{v}(\varphi, x) \mathrm{d}x = \Omega(u, v), \tag{5.1.74}$$

where we used that $\bar{s} = -s$, that follows by $s \in i\mathbb{R}$. Hence the operator $\mathcal{L}_5$ is Hamiltonian, with corresponding hamiltonian function

$$\begin{aligned} H_5(u, \bar{u}) = &\int_{\mathbb{T}} m_2 |u_x|^2 - \frac{i}{2} \mathrm{Im}(m_1)(u_x \bar{u} - u \bar{u}_x) - \mathrm{Re}(a_0^{(5)})|u|^2 \mathrm{d}x \\ &+ \int_{\mathbb{T}} -m|u|^2 - \frac{1}{2}(b_0^{(5)} \bar{u}^2 + \bar{b}_0^{(5)} u^2) \mathrm{d}x := \int_{\mathbb{T}} f_5(\varphi, x, u, \bar{u}, u_x, \bar{u}_x) \mathrm{d}x. \end{aligned} \tag{5.1.75}$$

Again using the Hamiltonian structure, see (2.3.41), we can conclude that

$$\mathrm{Im}(a_0^{(5)})(\varphi, x) = \frac{\mathrm{d}}{\mathrm{d}x} \mathrm{Im}(m_1) \equiv 0, \tag{5.1.76}$$

that implies $a_0^{(5)} \in \mathbb{R}$.

**Remark 5.1.93.** *We have* $b_2^{(5)} = b_1^{(5)} \equiv 0$ *and* $\mathcal{R}_5 \equiv 0$.

**Step 6. Descent Method: conjugation by pseudo-differential operator** In this section we want to conjugate $\mathcal{L}_5$ in (5.1.72) to an operator of the form $\omega \cdot \partial_\varphi + iM\partial_{xx} + iM_1\partial_x + \mathcal{R}$ where

$$M = \begin{pmatrix} m_2 & 0 \\ 0 & -m_2 \end{pmatrix}, \quad M_1 = \begin{pmatrix} m_1 & 0 \\ 0 & -\bar{m}_1 \end{pmatrix}, \tag{5.1.77}$$

and $\mathcal{R}$ is a pseudo differential operator of order 0.

We consider an operator of the form

$$\widetilde{\mathcal{S}} := \begin{pmatrix} 1 + w\Upsilon & 0 \\ 0 & 1 + \bar{w}\Upsilon \end{pmatrix}, \tag{5.1.78}$$

where $w : \mathbb{T}^{d+1} \to \mathbb{R}$ and $\Upsilon = (1 - \partial_{xx})\frac{1}{i}\partial_x$ is defined by linearity as

$$\Upsilon e^{ijx} = \frac{1}{1+j^2} j e^{ijx}$$

We have that the difference

$$\mathcal{L}_5 \widetilde{\mathcal{S}} - \widetilde{\mathcal{S}} \left[ \omega \cdot \partial_\varphi \mathbb{1} + i \begin{pmatrix} m_2 & 0 \\ 0 & -m_2 \end{pmatrix} \partial_{xx} + i \begin{pmatrix} m_1 & 0 \\ 0 & -\bar{m}_1 \end{pmatrix} \partial_x + i \begin{pmatrix} m + \hat{a}_0^{(5)} & b_0^{(5)} \\ -\bar{b}_0^{(5)} & -m - \hat{a}_0^{(5)} \end{pmatrix} \right] =$$
$$= i \begin{pmatrix} r_0 & 0 \\ 0 & -\bar{r}_0 \end{pmatrix} + \mathcal{R}$$

where $b_0^{(5)}$ is defined in (5.1.73) and

$$\begin{aligned} r_0(\varphi, x) &:= 2m_2 w_x \Lambda \partial_x + (a_0^{(5)}(\varphi, x) - \hat{a}_0^{(5)}(\varphi)), \quad \mathcal{R} = i \begin{pmatrix} \widetilde{p}_0 & \widetilde{q}_0 \\ -\bar{\widetilde{q}}_0 & -\bar{\widetilde{p}}_0 \end{pmatrix} \\ \widetilde{p}_0(\varphi, x) &:= -i(\omega \cdot \partial_\varphi w)\Upsilon + m_2 w_{xx} \Upsilon + m_1 w_x \Upsilon + (a_0^{(5)} - \hat{a}_0^{(5)}) w \Upsilon, \\ \widetilde{q}_0(\varphi, x) &:= b_0^{(5)} \bar{w} \Upsilon - w \Upsilon b_0^{(5)}. \end{aligned} \tag{5.1.79}$$

We are looking for $w$ such that $r_0 \equiv 0$ or at least $r_0$ is "small" in some sense. The operator $\mathcal{R}$ is a pseudo-differential operator of order $-1$. We can also note that

$$\Upsilon \partial_x u = iu - i(1 - \partial_{xx})^{-1} u$$

Since the second term is of order $-2$, we want to solve the equation

$$2imw_x + (a_0^{(5)} - \hat{a}_0^{(5)})u \equiv 0.$$

This equation has solution if and only if we define

$$\hat{a}_0^{(5)}(\varphi) := \frac{1}{2\pi} \int_{\mathbb{T}} a_0^{(5)}(\varphi, x) \mathrm{d}x, \tag{5.1.80}$$

and it is real thanks to (5.1.76). Now, we define

$$w(\varphi, x) := i \frac{1}{2m} \partial_x^{-1} (a_0^{(5)} - \hat{a}_0^{(5)})(\varphi, x), \tag{5.1.81}$$

that is a purely imaginary function. In this way we can conjugate the operator $\mathcal{L}_5$ to an operator of the form $\omega \cdot \partial_\varphi + iM\partial_{xx} + iM_1 \partial_x + iM_0 + O(\partial_x^{-1})$ with the diagonal part of $M_0$ constant in the space variable. Unfortunately, this transformation in not symplectic. We reason as follow. Let $w = i(w + \bar{w}) := ia$ and consider the Hamiltonian function

$$H(u, \bar{u}) = \frac{1}{2} \int_{\mathbb{T}} -(a\Upsilon + \Upsilon a) u \cdot \bar{u} dx.$$

Since the function $a$ is real, and the operator $\Upsilon : L^2(\mathbb{T};\mathbb{C}) \to L^2(\mathbb{T},\mathbb{C})$ is self-adjoint, then the operator $a\Upsilon + \Upsilon a$ is self-adjont. As consequence the hamiltonian $H$ is real-valued on $L^2$. The corresponding (linear) vector field is

$$\chi_H(u,\bar{u}) = -i\begin{pmatrix}\partial_{\bar{u}}H\\ \partial_u H\end{pmatrix} = \begin{pmatrix}\frac{i}{2}(a\Upsilon+\Upsilon a)u\\ -\frac{i}{2}(a\Upsilon+\Upsilon a)\bar{u}\end{pmatrix}.$$

Then, the 1-flow of $\chi_H$ generates a symplectic transformation of coordinates, given by

$$\begin{aligned}\mathcal{T}_6 &:= \exp(\chi_H(u,\bar{u})) := \begin{pmatrix}e^{i\chi} & 0\\ 0 & e^{-i\chi}\end{pmatrix}\begin{pmatrix}u\\ \bar{u}\end{pmatrix},\\ e^{i\chi}u &:= \left(\sum_{m=0}^{\infty}\frac{1}{m!}\Big(\frac{1}{2}(a\Upsilon+\Upsilon a)\Big)^m\right)u.\end{aligned} \tag{5.1.82}$$

We can easily check that, the operators in (5.1.82) and (5.1.78) differs only for an operator of order $O(\partial_x^{-2})$. Indeed one has

$$\begin{aligned}e^{i\chi}u &= u+\frac{i}{2}(a\Upsilon+\Upsilon a)u+O(\partial_x^{-2}) = u+\frac{i}{2}a\Upsilon u+\frac{i}{2}\Upsilon(au)+O(\partial_x^{-2})\\ &= u+\frac{i}{2}a\Upsilon u+\frac{i}{2}a\Upsilon u+\frac{i}{2}\frac{1}{i}\partial_x\Big((1-\partial_{xx})^{-1}a_{xx}(1-\partial_{xx})^{-1}u\\ &+2(1-\partial_{xx})^{-1}a_x(1-\partial_{xx})^{-1}\partial_x u\Big)+O(\partial_x^{-2})\\ &= (\mathbb{1}+ia\Upsilon)u+O(\partial_x^{-2}).\end{aligned} \tag{5.1.83}$$

In (5.1.83) we essentially studied the commutator of the pseudo-differential operator $(1-\partial_{xx})^{-1}$ with the operator of multiplication by the function $a$. Since the transformation $\mathcal{T}_6$ is symplectic we obtain the hamiltonian operator

$$\begin{aligned}\mathcal{L}_6 &= \mathcal{T}_6^{-1}\mathcal{L}_5\mathcal{T}_6 = \widetilde{\mathcal{L}}_5+\widetilde{\mathcal{R}},\\ \widetilde{\mathcal{L}}_5 &:= \omega\cdot\partial_\varphi\mathbb{1}+im_2E\partial_{xx}+i\begin{pmatrix}m_1 & 0\\ 0 & -\bar{m}_1\end{pmatrix}\partial_x+imE+i\begin{pmatrix}\hat{a}_0^{(5)}(\varphi) & b_0^{(5)}\\ -\bar{b}_0^{(5)} & -\hat{a}_0^{(5)}(\varphi)\end{pmatrix},\\ \widetilde{\mathcal{R}} &:= \mathcal{T}_6^{-1}\Big[\mathcal{L}_5\mathcal{T}_6-\mathcal{T}_6\widetilde{\mathcal{L}}_5\Big],\end{aligned} \tag{5.1.84}$$

where $\mathcal{R}$ is hamiltonian and of order $O(\partial_x^{-1})$.

**Remark 5.1.94.** *Here we have that $\mathcal{L}_6$ has the form (5.1.20) where $b_2^{(6)} = b_1^{(6)} \equiv 0$, $a_0^{(6)} := \hat{a}_0^{(5)}$, $b_0^{(6)} := b_0^{(5)}$ and $\mathcal{R}_6 := \widetilde{\mathcal{R}}$.*

**Step 7. Descent Method: conjugation by multiplication operator II** In this section we want to eliminate the dependance on the time variable of the coefficients $\hat{a}_0^{(5)}(\varphi)$ in (5.1.80).

Consider the operator

$$\mathcal{T}_7 := \begin{pmatrix} 1+k(\varphi) & 0 \\ 0 & 1+\bar{k}(\varphi) \end{pmatrix}, \tag{5.1.85}$$

with $k : \mathbb{T}^d \to \mathbb{C}$. By direct calculation we have that

$$\begin{aligned} &\mathcal{L}_6\mathcal{T}_7 - \mathcal{T}_7\left[\omega\cdot\partial_\varphi \mathbb{1} + im_2E + i\begin{pmatrix} m_1 & 0 \\ 0 & -\bar{m}_1\end{pmatrix}\partial_x + i\begin{pmatrix} m_0 & 0 \\ 0 & -m_0\end{pmatrix}\right] \\ &\qquad = i\begin{pmatrix} r_1 & 0 \\ 0 & -\bar{r}_1\end{pmatrix} + \left[i\begin{pmatrix} 0 & b_0^{(6)} \\ -\bar{b}_0^{(6)} & 0\end{pmatrix} + \mathcal{R}_6\right]\mathcal{T}_7, \end{aligned} \tag{5.1.86}$$

where

$$r_1(\varphi) = \omega\cdot\partial_\varphi k(\varphi) + i(a_0^{(6)}(\varphi) - m_0)(1+k(\varphi)), \tag{5.1.87}$$

We are looking for $\Gamma$ such that $r_1 \equiv 0$. As done in step 5, we write $1+k(\varphi) = \exp(\Gamma(\varphi))$, then equation $r_1 \equiv 0$ reads

$$\omega\cdot\partial_\varphi\Gamma(\varphi) + i(a_0^{(6)}(\varphi) + m - m_0) = 0, \tag{5.1.88}$$

that has a unique solution if and only if we define

$$m_0 := m + \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} a_0^{(6)}(\varphi)\mathrm{d}\varphi. \tag{5.1.89}$$

Hence we can set

$$\Gamma(\varphi) := -i(\omega\cdot\partial_\varphi)^{-1}(a_0^{(6)} + m - m_0)(\varphi). \tag{5.1.90}$$

It turns out that the trasformation $\mathcal{T}_7$ is invertible, then, by conjugation, we obtain $\mathcal{L}_7 := \mathcal{T}_7^{-1}\mathcal{L}_6\mathcal{T}_7$ with

$$\mathcal{L}_7 := \omega\cdot\partial_\varphi\mathbb{1} + i\begin{pmatrix} m & 0 \\ 0 & -m\end{pmatrix}\partial_{xx} + i\begin{pmatrix} m_1 & 0 \\ 0 & -\bar{m}_1\end{pmatrix}\partial_x + i\begin{pmatrix} m_0 & b_0^{(7)} \\ -\bar{b}_0^{(7)} & -m_0\end{pmatrix} + \mathcal{R}_7 \tag{5.1.91}$$

where we have defined

$$\begin{pmatrix} 0 & b_0^{(7)} \\ -\bar{b}_0^{(7)} & 0\end{pmatrix} := \mathcal{T}_7^{-1}\begin{pmatrix} 0 & b_0^{(6)} \\ -\bar{b}_0^{(6)} & 0\end{pmatrix}\mathcal{T}_7, \quad \mathcal{R}_7 := \mathcal{T}_7^{-1}\mathcal{R}_6\mathcal{T}_7. \tag{5.1.92}$$

Moreover, since by (5.1.90) the function $\Gamma$ is purely imaginary, then the transformation is symplectic. Indeed

$$\Omega(e^\Gamma u, e^\Gamma v) := \mathrm{Re}\int_{\mathbb{T}} ie^\Gamma u e^{-\Gamma}\bar{v}\mathrm{d}x = \Omega(u,v), \tag{5.1.93}$$

hence the linearized operator $\mathcal{L}_7$ is Hamiltonian.

### 5.1.1 Non-degeneracy Condition

Here we give the proof of formula (5.1.16) Let us study the properties of the average of the coefficients of the first order differential operator. In particular we are interested in how these quantities depends explicitly on $\omega$, see Remark 5.1.88. Consider $a_1(\varphi,x)=a_1(\varphi,x,u)$ where $u$ satisfies (5.1.12) and $a_i$ is defined in (2.3.38). One has

$$\left|\int_{\mathbb{T}^{d+1}} a_1(\varphi,x)\right| \geq \varepsilon\mathfrak{e} - C\varepsilon\|u\|_{\mathfrak{s}_0+\eta_1} \geq \frac{\mathfrak{e}}{2}\varepsilon, \tag{5.1.94}$$

if $\varepsilon\gamma_0^{-1}$ is small enough. Essentially, by using (5.1.22), (5.1.23), (5.1.24a) and (5.1.24b), one can repeat the reasoning followed in (5.1.94) for the average of $a_1^{(i)}$ for $i=1,2,3,4$ and prove the (5.1.16a) with a constant $c<\mathfrak{e}/16$. Let us check (5.1.16b). At the starting point there is no explicit dependence on the parameters $\omega$ in $a_1$, hence we get also for $\omega_1\neq\omega_2$

$$0=\left|\int_{\mathbb{T}^{d+1}} a_1(\varphi,x,\omega_1,u(\omega)) - a_1(\varphi,x,\omega_2,u(\omega))\right| \leq \varepsilon^2 C|\omega_1-\omega_2|. \tag{5.1.95}$$

Now, by (5.1.27) one has that

$$a_1^{(1)}(\varphi,x,\omega,u(\omega)) := a_1^{(1)}(\varphi,x,u(\omega)) := i(2\mathcal{T}_1^{-1}(E+A_2)\partial_x\mathcal{T}_1 + \mathcal{T}_1^{-1}A_1\mathcal{T}_1)_1^1,$$

and again we do not have explicit dependence on $\omega$ since the matrix $\mathcal{T}_1$ depends on the external parameters only trough the function $u$. Hence bound (5.1.95) holds. Now consider the coefficients $a_1^{(2)}$ in (5.1.37). There is explicit dependence on $\omega$ only in the term

$$\mathcal{T}_2^{-1}(\omega\cdot\partial_\varphi\xi) = \sqrt{1+\hat{\xi}_y(\varphi,y)}\,\omega\cdot\partial_\varphi\xi(\varphi,y+\hat{\xi}(\varphi,y)). \tag{5.1.96}$$

Recall that the functions $\xi$ in (5.1.42) and $\hat{\xi}$ depends on $\omega$ only through $u$. Hence one has

$$\begin{aligned}
&\left|\int_{\mathbb{T}^{d+1}} \sqrt{1+\hat{\xi}_y}(\omega_1-\omega_2)\cdot\partial_\varphi\xi(\varphi,y+\hat{\xi}(\varphi,y))\mathrm{d}\varphi\mathrm{d}y\right| = \\
&= \left|\int_{\mathbb{T}_{d+1}} \frac{(\omega_1-\omega_2)\cdot\partial_\varphi\xi(\varphi,x)}{\sqrt{1+\hat{\xi}_y(\varphi,x+\xi(\varphi,x))}}\mathrm{d}\varphi\mathrm{d}x\right| \\
&\leq |\omega_1-\omega_2|\left|\int_{\mathbb{T}^{d+1}} \partial_\varphi\xi(\varphi,x)\mathrm{d}\varphi\mathrm{d}x\right| \\
&+ |\omega_1-\omega_2|\left|\int_{\mathbb{T}_{d+1}} \left(\frac{1}{\sqrt{1+\hat{\xi}_y}}-1\right)\partial_\varphi\xi(\varphi,x)\mathrm{d}\varphi\mathrm{d}x\right|
\end{aligned} \tag{5.1.97}$$

By defining $|u|_s^\infty := ||u||_{W^{s,\infty}}$ and using the standard estimates of the Sobolev embedding on the function $\xi$ in (5.1.42) we get

$$|\xi|_s^\infty \leq C(s)||\xi||_{s+\mathfrak{s}_0} \leq C(s)||\rho_0||_{s+\mathfrak{s}_0} \leq \varepsilon C(s)(1+||u||_{s+\mathfrak{s}_0+2}),. \tag{5.1.98a}$$

The function $\hat{\xi}$ satisfies the same bounds by Lemma A.167. Hence, since the first integral in (5.1.97) is zero, using the interpolation estimates in Lemma A.164, we get

$$\left|\int_{\mathbb{T}^{d+1}} a_1^{(2)}(\varphi,x)\mathrm{d}\varphi\mathrm{d}x\right|^{lip} \leq C\varepsilon^2. \tag{5.1.99}$$

Let us study the coefficients $a_1^{(3)}$ defined in (5.1.50). In particular one need to control the difference $a_1^{(3)}(\omega_1) - a_1^{(3)}(\omega_2)$. To do this one can uses standard formulæ of propagation of errors for Lipschitz functions. In order to perform the quantitative estimates one can check that the function $\alpha(\varphi)$ defined in (5.1.54) satisfies the tame estimates (see also Lemma 4.2.53):

$$|\alpha|_s^\infty \leq \varepsilon\gamma_0^{-1}C(s)(1+||u||_{s+d+\mathfrak{s}_0+2}), \tag{5.1.100a}$$

$$|d_u\alpha(u)[h]|_s^\infty \leq \varepsilon\gamma_0^{-1}C(s)(||h||_{s+d+\mathfrak{s}_0+2} + ||u||_{s+d+\mathfrak{s}_0+2}||h||_{d+\mathfrak{s}_0+2}), \tag{5.1.100b}$$

$$|\alpha|_{s,\gamma}^\infty \leq \varepsilon\gamma_0^{-1}C(s)(1+||u||_{s+d+\mathfrak{s}_0+2,\gamma}), \tag{5.1.100c}$$

while by the (5.1.49) one has $\rho = 1 + \mathcal{T}_3^{-1}(\omega\cdot\partial_\varphi\alpha)$. By using Lemma A.167 and the bounds (5.1.100) on $\alpha$ and (5.1.12) one can prove

$$|\rho-1|_{s,\gamma}^\infty \leq \varepsilon\gamma_0^{-1}C(s)(1+||u||_{s+d+\mathfrak{s}_0+4,\gamma}) \tag{5.1.101a}$$

$$|d_u\rho(u)[h]|_s^\infty \leq \varepsilon\gamma_0^{-1}C(s)(||h||_{s+d+\mathfrak{s}_0+3} + ||u||_{s+d+\mathfrak{s}_0+4}||h||_{d+\mathfrak{s}_0+3}). \tag{5.1.101b}$$

The bounds above follows by classical tame estimates in Sobolev spaces, anyway the proof can be found in Section 3 of [31]. Now by taking the integral of (5.1.99) and by using (5.1.100a)-(5.1.101b), the tame estimates in Lemma A.167 and the (5.1.24a), (5.1.24b) one obtain the result on the . For the last step one can reason in the same way. Indeed the most important fact is to prove (5.1.99). At the starting point we have no explicit dependence on $\lambda$ in the average of $a_1$, but, once that dependence appear, then we have the estimates (5.1.99) that is quadratic in $\varepsilon$.

One has also the following result.

**Lemma 5.1.95.** *Under the Hypotheses of Lemma 5.1.87 possibly with smaller $\epsilon_0$, if (5.1.12) holds, one has that the $\mathcal{T}_i$, $i \neq 3$ identify operators $\mathcal{T}_i(\varphi)$, of the phase space $\mathbf{H}_x^s := \mathbf{H}^s(\mathbb{T})$. Moreover they are invertible and the following estimates hold for $\mathfrak{s}_0 \leq s \leq q-\eta_1$ and i=1,2,4,5,6,7:*

$$||(\mathcal{T}_i^{\pm1}(\varphi)-\mathbb{1})h||_{\mathbf{H}_x^s} \leq \varepsilon C(s)(||h||_{\mathbf{H}_x^s} + ||u||_{s+d+2\mathfrak{s}_0+4}||h||_{\mathbf{H}_x^1}), \tag{5.1.102a}$$

The Lemma is essentially a consequence if the discussion above. We omit the details because the proof follows basically the same arguments used Lemma 5.1.95.

## 5.2 Reduction to constant coefficients

In this Section we conclude the proof of Proposition 5.0.85 through a reducibility algorithm. First we need to fix some notations. Let $b \in \mathbb{N}$, we consider the exponential basis $\{e_i : i \in \mathbb{Z}^b\}$ of $L^2(\mathbb{T}^b)$. In this way we have that $L^2(\mathbb{T}^2)$ is the space $\{u = \sum u_i e_i : \sum |u_i|^2 < \infty\}$. A linear operator $A : L^2(\mathbb{T}^b) \to L^2(\mathbb{T}^b)$ can be written as an infinite dimensional matrix

$$A = (A_i^j)_{i,j \in \mathbb{Z}^b}, \quad A_i^j = (Ae_j, e_i)_{L^2(\mathbb{T}^b)}, \quad Au = \sum_{i,j} A_i^j u_j e_i.$$

where $(\cdot,\cdot)_{L^2(\mathbb{T}^{d+1})}$ is the usual scalar product on $L^2$. In the following we also use the decay norm

$$|A|_s^2 := \sup_{\sigma,\sigma' \in \boldsymbol{C}} |A_\sigma^{\sigma'}|_s^2 := \sup_{\sigma,\sigma' \in \boldsymbol{C}} \sum_{h \in \mathbb{Z}\times\mathbb{Z}^d} \langle h \rangle^{2s} \sup_{k-k'=h} |A_{\sigma,k}^{\sigma',k'}|^2. \tag{5.2.103}$$

Note that this is the same definition of the norm in (4.3.85) with the only difference that here the indexes are in $\mathbb{Z}\times\mathbb{Z}^d$ while in (4.3.85) one has $h \in \mathbb{N}\times\mathbb{Z}^d$. Clearly all the properties proved in Lemmata 4.3.62, 4.3.63, 4.3.64 and 4.3.66 hold.

**Theorem 5.2.96.** *Let $\mathtt{f} \in C^q$ satisfy the Hypotheses of Proposition 5.0.85 with $q > \eta_1 + \beta + \mathfrak{s}_0$ where $\eta_1$ defined in (5.1.11) and $\beta = 7\tau + 5$ for some $\tau > d$. Let $\gamma \in (0,\gamma_0)$, $\mathfrak{s}_0 \le s \le q - \eta_1 - \beta$ and $u(\lambda) \in \mathbf{H}^0$ be a family of functions depending on a Lipschitz way on a parameter $\omega \in \Lambda_o \subset \Lambda : [1/2, 3/2]$. Assume that*

$$||u||_{\mathfrak{s}_0+\eta_1+\beta,\Lambda_o,\gamma} \le 1. \tag{5.2.104}$$

*Then there exist constants $\epsilon_0$, $C$, depending only on the data of the problem, such that, if $\varepsilon\gamma^{-1} \le \epsilon_0$, then there exists a sequence of purely imaginary numbers as in Proposition 4.1.51, namely $\Omega_{\sigma,j}^{j}, \Omega_{\sigma,j}^{-j} : \Lambda \to \mathbb{C}$ of the form*

$$\begin{aligned} \Omega_{\sigma,j}^{j} &:= -i\sigma m_2 j^2 - i\sigma|m_1|j + i\sigma m_0 + i\sigma r_j^j, \\ \Omega_{\sigma,j}^{-j} &:= i\sigma r_j^{-j}, \end{aligned} \tag{5.2.105}$$

*where*

$$m_2, m_0 \in \mathbb{R}, \;\; m_1 \in i\mathbb{R}, \quad \overline{r_j^k} = r_k^j, \;\; k = \pm j \tag{5.2.106}$$

*for any $\sigma \in \boldsymbol{C}$, $j \in \mathbb{N}$, moreover*

$$|r_{\sigma,j}^{k}|_\gamma \le \frac{\varepsilon C}{\langle j \rangle}, \quad \forall\, \sigma \in \boldsymbol{C}, \; j \in \mathbb{Z}, \;\; k = \pm j, \tag{5.2.107}$$

*and such that, for any* $\omega \in \Lambda_\infty^{2\gamma}(u)$*, defined in (4.1.23), there exists a bounded, invertible linear operator* $\Phi_\infty(\omega) : \mathbf{H}^s \to \mathbf{H}^s$*, with bounded inverse* $\Phi_\infty^{-1}(\omega)$*, such that*

$$\begin{aligned} &\mathcal{L}_\infty(\omega) := \Phi_\infty^{-1}(\omega) \circ \mathcal{L}_7 \circ \Phi_\infty(\omega) = \omega \cdot \partial_\varphi \mathbb{1} + i\mathcal{D}_\infty, \\ \text{where} \qquad &\mathcal{D}_\infty := \text{diag}_{\mathrm{h}=(\sigma,\mathrm{j})\in \mathrm{C}\times\mathbb{N}}\{\Omega_{\sigma,\underline{j}}(\omega)\}, \end{aligned} \tag{5.2.108}$$

*with* $\mathcal{L}_7$ *defined in* (5.1.13) *and where*

$$\Omega_{\sigma,\underline{j}} := \begin{pmatrix} \Omega_{\sigma,j}^{\;j} & \Omega_{\sigma,j}^{-j} \\ \Omega_{\sigma,-j}^{\;j} & \Omega_{\sigma,-j}^{-j} \end{pmatrix} \tag{5.2.109}$$

*Moreover, the transformations* $\Phi_\infty(\lambda)$, $\Phi_\infty^{-1}$ *are symplectic and satisfy*

$$|\Phi_\infty(\lambda) - \mathbb{1}|_{s,\Lambda_\infty^{2\gamma},\gamma} + |\Phi_\infty^{-1}(\lambda) - \mathbb{1}|_{s,\Lambda_\infty^{2\gamma},\gamma} \leq \varepsilon\gamma^{-1}C(s)(1 + ||u||_{s+\eta_1+\beta,\Lambda_o,\gamma}). \tag{5.2.110}$$

*In addition to this, for any* $\varphi \in \mathbb{T}^d$*, for any* $\mathfrak{s}_0 \leq s \leq q - \eta_1 - \beta$ *the operator* $\Phi_\infty(\varphi) : \mathbf{H}_x^s \to \mathbf{H}_x^s$ *is an invertible operator of the phase space* $\mathbf{H}_x^s := \mathbf{H}^s(\mathbb{T})$ *with inverse* $(\Phi_\infty(\varphi))^{-1} := \Phi_\infty^{-1}(\varphi)$ *and*

$$||(\Phi_\infty^{\pm 1}(\varphi) - \mathbb{1})h||_{\mathbf{H}_x^s} \leq \varepsilon\gamma^{-1}C(s)(||h||_{\mathbf{H}_x^s} + ||u||_{s+\eta_1+\beta+\mathfrak{s}_0}||h||_{\mathbf{H}_x^1}). \tag{5.2.111}$$

**Remark 5.2.97.** *Note that since the* $\Phi_\infty$ *is symplectic then the operator* $\mathcal{L}_\infty$ *is hamiltonian.*

The main point of the Theorem 5.2.96 is that the bound on the low norm of $u$ in (5.2.104) guarantees the bound on *higher* norms (5.2.110) for the transformations $\Phi_\infty^{\pm 1}$. This is fundamental in order to get the estimates on the inverse of $\mathcal{L}$ in high norms.

Moreover, the definition (5.0.4) of the set where the second Melnikov conditions hold, depends only on the final eigenvalues. Usually in KAM theorems, the non-resonance conditions have to be checked, inductively, at each step of the algorithm. This formulation, on the contrary, allow us to discuss the measure estimates only once. Indeed, the functions $\mu_h(\omega)$ are well-defined even if $\Lambda_\infty = \emptyset$, so that, we will perform the measure estimates as the last step of the proof of Theorem 1.1.2.

We need some technical lemmata on finite dimensional matrices.

**Lemma 5.2.98.** *Given a matrix* $M \in \mathcal{M}_n(\mathbb{C})$*, where* $\mathcal{M}_n(\mathbb{C})$ *is the space of the* $n \times n$ *matrix with coefficients in* $\mathbb{C}$*, we define the norm* $\|M\|_\infty := \max_{i,j=1,\dots,n}\{A_i^j\}$*. One has*

$$\|M\|_\infty \leq \|M\|_2 \leq n\|M\|_\infty, \tag{5.2.112}$$

*where* $\|\cdot\|_2$ *is the* $L^2-$*operatorial norm.*

*Proof.* It follow straightforward by the definitions. □

**Lemma 5.2.99.** *Take two self adjoint matrices $A, B \in \mathcal{M}_n(\mathbb{C})$. Let us define the operator $M : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$*

$$M : C \mapsto MC := AC - CB. \tag{5.2.113}$$

*Let $\lambda_j$ and $\beta_j$ for $j = 1, \dots, n$ be the eigenvalues respectively of $A$ and $B$. Then, for any $R \in \mathcal{M}_n(\mathbb{C})$ one has that the equation $MC = R$ has a solution with*

$$\|C\|_\infty \le K \left( \min_{i,j=1,\dots,n} \{\lambda_j - \beta_i\} \right)^{-1} \|R\|_\infty, \tag{5.2.114}$$

*where the constant $K$ depends only on $n$.*

*Proof.* Define the operator $\mathcal{T} : \mathcal{M}_n(\mathbb{C}) \to \mathbb{C}^{n^2}$ that associate to a matrix the vector of its components. Then the equation $MC = R$ can be rewritten as

$$(A \otimes \mathbb{1} - \mathbb{1} \otimes B^T)\mathcal{T}(C) = \mathcal{T}(R),$$

where $\mathbb{1}$ is the $n \times n$ identity. Then, by using Lemma 5.2.98, one has

$$\begin{aligned} \|C\|_\infty &= \max_{i=1,\dots,n^2} \|[\mathcal{T}(C)]_i\|_\infty \le n\|(A \otimes \mathbb{1} - \mathbb{1} \otimes B^T)^{-1}\|_\infty \max_{i=1,\dots,n^2} |[\mathcal{T}(R)]_i| \\ &\le n^2 \|(A \otimes \mathbb{1} - \mathbb{1} \otimes B^T)^{-1}\|_2 \max_{i=1,\dots,n^2} |[\mathcal{T}(R)]_i| \le n^2 c \left( \min_{i,j=1,\dots,n} \{\lambda_j - \beta_i\} \right)^{-1} \|R\|_\infty, \end{aligned} \tag{5.2.115}$$

that is the (5.2.114). □

### Hamiltonian operators

Here we give a characterization, in terms of the Fourier coefficients, of hamiltonian linear operators. This is important since we want to show that our algorithm is closed for such class of operators.

**Lemma 5.2.100.** *Consider a linear operator $B := (i\sigma R_\sigma^{\sigma'}) : \mathbf{H}^s \to \mathbf{H}^s$. Then, $B$ is hamiltonian with respect to the symplectic form (1.1.14) if and only if*

$$R_{\sigma,h}^{\sigma',h'} = R_{-\sigma',h'}^{-\sigma,h}, \qquad \overline{R_{\sigma,h}^{\sigma',h'}} = R_{\sigma',-h}^{\sigma,-h'} \tag{5.2.116}$$

*Proof.* In coordinates, an Hamiltonian function for such operator, is a quadratic form real and symmetric,

$$H = \sum_{\substack{\sigma,\sigma' \in \boldsymbol{C} \\ h,h' \in \mathbb{Z}^{d+1}}} Q_{\sigma,h}^{\sigma',h'} z_h^\sigma z_{h'}^{\sigma'},$$

where we denote $\overline{z_h^\sigma} = z_{-h}^{-\sigma}$ and $h = (j,p), h' = (j',p')$. This means that, $Q$ satisfies

$$\overline{Q_{\sigma,h}^{\sigma',h'}} = Q_{-\sigma,-h}^{-\sigma',-h'}, \quad Q_{\sigma,h}^{\sigma',h'} = Q_{\sigma',h'}^{\sigma,h} \tag{5.2.117}$$

Now, since the hamiltonian vector field associated to the Hamiltonian $H$ is given by $B = iJQ$, then writing

$$B = i\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} JQ$$

we set $R_{\sigma}^{\sigma'} = Q_{-\sigma}^{\sigma'}$ follow the (5.2.116). □

Since the operator $\mathcal{L}_\infty$ in Theorem 5.2.96 is hamiltonian, thanks to the characterization in Lemma 5.2.100 we can note that the blocks $\Omega_{\sigma,\underline{j}}$ defined in (5.2.109) as purely imaginary eigenvalues.

### 5.2.1 Reduction algorithm

We prove Theorem 5.2.96 by means of the following Iterative Lemma on the class of linear operators

**Definition 5.2.101.**

$$\omega\cdot\partial_\varphi \mathbb{1} + \mathcal{D} + \mathcal{R} :\ \mathbf{H}^0 \to \mathbf{H}^0, \tag{5.2.118}$$

*where* $\omega$ *is as in* (1.1.7)*, and*

$$\mathcal{D} = diag_{(\sigma,j)\in\boldsymbol{C}\times\mathbb{Z}}\{\Omega_{\sigma,\underline{j}}\} := diag_{(\sigma,j)\in\boldsymbol{C}\times\mathbb{Z}}\left\{\begin{pmatrix}\Omega_{\sigma,j}^{\ j} & \Omega_{\sigma,j}^{-j}\\ \Omega_{\sigma,-j}^{j} & \Omega_{\sigma,-j}^{-j}\end{pmatrix}\right\}, \tag{5.2.119}$$

*where*

$$\begin{aligned}\Omega_{\sigma,j}^{\ j} &:= -i\sigma m_2 j^2 - i\sigma|m_1|j + i\sigma m_0 + i\sigma r_j^j,\\ \Omega_{\sigma,j}^{-j} &:= i\sigma r_j^{-j},\end{aligned} \tag{5.2.120}$$

*and*

$$m_2, m_0 \in \mathbb{R},\ \ m_1 \in i\mathbb{R},\ \ \ \overline{r_j^k} = r_k^j,\ \ \ r_j^k = O(\frac{\varepsilon}{\langle j\rangle})\ \ k=-j,\ \ r_j^k = O(\frac{\varepsilon}{\langle j\rangle}),\ k=j \tag{5.2.121}$$

*for any* $(\sigma,j)\in\boldsymbol{C}\times\mathbb{N}$*, with* $\mathcal{R}$ *is a Töpliz in time Hamiltonian operator such that* $\mathcal{R}_\sigma^\sigma = O(\varepsilon\partial_x^{-1})$ *and* $\mathcal{R}_\sigma^{-\sigma} = O(\varepsilon)$ *for* $\sigma=\pm1$*. Moreover we set* $\mu_{\sigma,j}$ *for* $\sigma\in\boldsymbol{C}$ *the eigenvalues of* $\Omega_{\sigma,\underline{j}}$*.*

Note that the operator $\mathcal{L}_7$ has the form (5.2.118) and satisfies the (5.2.119) and (5.2.120) as well as the estimates (5.1.17) and (5.1.18). Note moreover that for $\mathcal{L}_7$ the matrix $\mathcal{D}$ is completely diagonal. This fact is not necessary for our analysis, and it cannot be preserved during the algorithm.

Define

$$N_{-1} := 1,\quad N_\nu := N_{\nu-1}^{\chi} = N_0^{\chi^\nu},\ \ \forall\,\nu\geq 0,\ \ \chi=\frac{3}{2}. \tag{5.2.122}$$

and

$$\alpha = 7\tau+3,\qquad \eta_3 := \eta_1+\beta, \tag{5.2.123}$$

where $\eta_1$ is defined in (5.1.11) and $\beta = 7\tau + 5$. Consider $\mathcal{L}_7 = \mathcal{L}_0$.Note that $\mathcal{L}_7$ belongs to the class of Definition 5.2.101. Indeed in this case we have that

$$\mathcal{R}_0 := \begin{pmatrix} 0 & q_0(\varphi, x) \\ -\bar{q}_0(\varphi, x) & 0 \end{pmatrix} + \mathcal{R},$$

(see (5.1.13)) and $\mathcal{R}$ is a pseudo differential operator of order $O(\partial_x^{-1})$. We have the following lemma:

**Lemma 5.2.102.** *The operator $\mathcal{R}$ defined in Lemma 4.2.53 satisfies the bounds*

$$|\mathcal{R}(u)|_{s,\gamma} \leq \varepsilon C(s)(1 + ||u||_{s+\eta_1,\gamma}), \tag{5.2.124a}$$

$$|d_u\mathcal{R}(u)[h]|_s \leq \varepsilon C(s) \left(||h||_{\mathfrak{s}_0+\eta_1} + ||h||_{s+\eta_1} + ||u||_{s+\eta_1}||h||_{\mathfrak{s}_0}\right), \tag{5.2.124b}$$

*where $\eta_1$ is defined in Lemma 5.1.87.*

*Proof.* By the proof of Lemma 5.1.87 we have that in the operator $\mathcal{L}_5$ in (5.1.72) the remainder is just a multiplication operator by the functions $a_0^{(5)}, b_0^{(5)}$. Hence by Remark 4.3.67 one has that the decay norm of the operator is finite. We need to check that the transformation $\mathcal{T}_6$ has a finite decay norm. First of all we have that the function $w$ in (5.1.81) satisfies the following estimates:

$$\begin{aligned} ||w||_{s,\gamma} &\leq_s \varepsilon(1 + ||u||_{s+\tau_1,\gamma}), \\ ||\partial_u w(u)[h]||_s &\leq_s \varepsilon(||h||_{s+\tau_1} + ||u||_{s+\tau_1}||h||_{\tau_1}), \end{aligned} \tag{5.2.125}$$

with $\tau_1$ a constant depending only on the data of the problem and much small than $\eta_1$.[1]

The operator $\widetilde{S} = \mathbb{1} + w\Upsilon$ defined in (5.1.78) satisfies the following estimates in norm $|\cdot|_s$ defined in (4.3.85):

$$\begin{aligned} |\widetilde{S} - \mathbb{1}|_{s,\gamma} &\leq_s \varepsilon(1 + \|u\|_{s+\tau_1,\gamma}), \\ |\partial_u\widetilde{S}(u)[h]|_s &\leq_s \varepsilon(\|h\|_{s+\tau_1} + \|u\|_{s+\tau_1}\|h\|_{\tau_1}), \end{aligned} \tag{5.2.126}$$

The (5.2.126) follow by the (5.2.125) and the fact that $|\Upsilon|_s \leq 1$ using Lemma 4.3.67. Clearly also the transformation $\mathcal{T}_6$ defined in (5.1.82) satisfies the same estimates as in (5.2.126). Hence using Lemma **??** one has that the remainder $\widetilde{R}$ of the operator $\mathcal{L}_6$ in (5.1.84) satisfies bounds like (5.2.124) with a different constant $\tau_2$ (possibly greater than $\tau_1$) instead of $\eta_1$. Now the last transformation $\mathcal{T}_7$ is a multiplication operator, then, by using again Lemmata 4.3.62 and 4.3.67 one obtain the (5.2.124) on the remainder of the operator $\mathcal{L}_7$ in (5.1.91). □

[1]to prove Lemma 5.1.87 one prove bounds like (5.1.17) and (5.1.19) on the coefficients of each $\mathcal{L}_i$ with loss of regularity $\tau_i$ at each step. The constant $\eta_1$ of the Lemma is obtained by collecting together the loss of regularity of each step.

**Lemma 5.2.103.** *Let $q > \eta_1 + \mathfrak{s}_0 + \beta$. There exist constant $C_0 > 0$, $N_0 \in \mathbb{N}$ large, such that if*

$$N_0^{C_0}\gamma^{-1}|\mathcal{R}_0|_{\mathfrak{s}_0+\beta} \leq 1, \tag{5.2.127}$$

*then, for any $\nu \geq 0$:*

$(\mathbf{S1})_\nu$ *There exists operators*

$$\mathcal{L}_\nu := \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}_\nu + \mathcal{R}_\nu, \;\; \mathcal{D}_\nu = \mathrm{diag}_{h \in \boldsymbol{C} \times \mathbb{Z}}\{\Omega^\nu_{\sigma,\underline{j}}\}, \tag{5.2.128}$$

*where*

$$\Omega^\nu_{\sigma,\underline{j}}(\omega) = \begin{pmatrix} \Omega^{\nu,j}_{\sigma,j} & \Omega^{\nu,-j}_{\sigma,j} \\ \Omega^{\nu,j}_{\sigma,-j} & \Omega^{;\nu-j}_{\sigma,-j} \end{pmatrix}, \tag{5.2.129}$$

*and*

$$\begin{aligned} \Omega^{\nu,j}_{\sigma,j} &:= -i\sigma m_2 j^2 - i\sigma|m_1|j + i\sigma m_0 + i\sigma r^{\nu,j}_j =: \Omega^{0,j}_{\sigma,j} + i\sigma r^{\nu,j}_j, \\ \Omega^{\nu,-j}_{\sigma,j} &:= i\sigma r^{\nu,-j}_j =: \Omega^{0,-j}_{\sigma,j} + i\sigma r^{\nu,-j}_j, \end{aligned}$$

*with $(\sigma, j) \in \boldsymbol{C} \times \mathbb{Z}$, and defined for $\lambda \in \Lambda^\gamma_\nu := \Lambda^\gamma_\nu$, with $\Lambda^\gamma_0 := \Lambda_o$ and for $\nu \geq 1$,*

$$\begin{aligned} &\Lambda^\gamma_\nu := \mathcal{P}^\gamma_\nu(u) \cap \mathcal{O}^\gamma_\nu, \\ &\mathcal{S}^\gamma_\nu(u) := \left\{ \omega \in \Lambda^\gamma_{\nu-1} : \quad \begin{array}{l} |i\omega\cdot\ell + \mu^{\nu-1}_h(\omega) - \mu^{\nu-1}_{h'}(\omega)| \geq \frac{\gamma|\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^\tau}, \\ \forall |\ell| \leq N_{\nu-1}, h, h' \in \boldsymbol{C} \times \mathbb{Z} \end{array} \right\}, \\ &\mathcal{O}^\gamma_\nu(u) := \left\{ \omega \in \Lambda^\gamma_{\nu-1} : \quad \begin{array}{l} |i\omega\cdot\ell + \mu^{\nu-1}_{\sigma,j} - \mu^{\nu-1}_{\sigma,k}| \geq \frac{\gamma}{\langle \ell \rangle^\tau \langle j \rangle}, \\ \ell \in \mathbb{Z}^d \setminus \{0\}, j \in \mathbb{Z}, k = \pm j, \sigma \in \boldsymbol{C} \end{array} \right\}, \end{aligned} \tag{5.2.130}$$

*where*

$$\begin{aligned} \mu^\nu_{\sigma,j} &:= i\sigma\left(-m_2 j^2 + m_0 + r^{\nu,j}_j + r^{\nu,-j}_{-j} + \frac{1}{2}a_j b_j\right), \\ b^\nu_j &:= \sqrt{\left(-2|m_1| + \frac{r^{\nu,j}_j - r^{\nu,-j}_{-j}}{a_j}\right)^2 + 4\frac{|r^{\nu,-j}_j|^2}{(a_j)^2}}, \\ a_j &= j, \text{ if } j \neq 0, \;\; a_j = 1, \text{ if } j = 0, \end{aligned} \tag{5.2.131}$$

*are the eigenvalues of the matrix $\Omega^\nu_{\sigma,\underline{j}}$. For $\nu \geq 0$ one has $\overline{r^{\nu,k}_j} = r^{\nu,j}_k$, for $k = \pm j$ and*

$$|r^{\nu,k}_j|_\gamma := |r^{\nu,k}_j|_{\Lambda^\gamma_\nu,\gamma} \leq \frac{\varepsilon C}{\langle j \rangle}, \qquad |r^{\nu,-j}_j|_\gamma \leq \frac{\varepsilon C}{\langle j \rangle}, \quad |b^\nu_j|_\gamma \leq \varepsilon C. \tag{5.2.132}$$

*The remainder $\mathcal{R}_\nu$ satisfies $\forall\, s \in [\mathfrak{s}_0, q - \eta_1 - \beta]$ ($\alpha$ is defined in (5.2.123))*

$$\begin{aligned} |\mathcal{R}_\nu|_s &\leq |\mathcal{R}_0|_{s+\beta} N^{-\alpha}_{\nu-1}, \\ |\mathcal{R}_\nu|_{s+\beta} &\leq |\mathcal{R}_0|_{s+\beta} N_{\nu-1}, \end{aligned} \tag{5.2.133}$$

$$|(\mathcal{R}_\nu)^{-\sigma}_\sigma|_s + |D(\mathcal{R}_\nu)^\sigma_\sigma|_s \lessdot |\mathcal{R}_\nu|_s, \ \sigma \in \boldsymbol{C}, \quad \text{where} \quad D := \mathrm{diag}_{j\in\mathbb{Z}}\{j\}. \tag{5.2.134}$$

*Moreover there exists a map $\Phi_{\nu-1}$ of the form $\Phi_{\nu-1} := \exp(\Psi_{\nu-1}) : \mathbf{H}^s \to \mathbf{H}^s$, where $\Psi_{\nu-1}$ is Töplitz in time, $\Psi_{\nu-1} := \Psi_{\nu-1}(\varphi)$ (see (4.3.98)), such that*

$$\mathcal{L}_\nu := \Phi^{-1}_{\nu-1}\mathcal{L}_{\nu-1}\Phi_{\nu-1} \tag{5.2.135}$$

*and for $\nu \geq 1$ one has:*

$$|\Psi_{\nu-1}|_{s,\gamma} \leq |\mathcal{R}_0|^0_{s+\beta} N^{2\tau+1}_{\nu-1} N^{-\alpha}_{\nu-2}. \tag{5.2.136}$$

*One has that the operators $\Phi^{\pm 1}_{\nu-1}$ are symplectic and the operator $\mathcal{R}_\nu$ is hamiltonian. Finally the eigenvalues $\mu^\nu_{\sigma,j}$ are purely imaginary.*

$(\mathbf{S2})_\nu$ *For all $j \in \mathbb{Z}$ there exists Lipschitz extensions $\widetilde{\Omega}^{\nu,k}_{\sigma,j} : \Lambda \to i\mathbb{R}$ of $\Omega^{\nu,k}_{\sigma,j} : \Lambda^\gamma_\nu \to i\mathbb{R}$, for $k = \pm j$, and $\widetilde{\mu}^\nu_h(\cdot) : \Lambda \to i\mathbb{R}$ of $\mu^\nu_h(\cdot) : \Lambda^\gamma_\nu \to i\mathbb{R}$, such that for $\nu \geq 1$,*

$$\begin{aligned} &|\widetilde{\Omega}^{\nu,k}_{\sigma,j} - \widetilde{\Omega}^{\nu-1,k}_{\sigma,j}|_\gamma \leq |(\mathcal{R}_{\nu-1})^\sigma_\sigma|_{\mathfrak{s}_0}, \quad \sigma \in \boldsymbol{C}, j \in \mathbb{Z}, k = \pm j, \\ &|\widetilde{\mu}^\nu_{\sigma,j} - \widetilde{\mu}^{\nu-1}_{\sigma,j}|^{\sup} \leq |(\mathcal{R}_{\nu-1})^\sigma_\sigma|_{\mathfrak{s}_0}, \quad \sigma \in \boldsymbol{C}, j \in \mathbb{Z}. \end{aligned} \tag{5.2.137}$$

$(\mathbf{S3})_\nu$ *Let $u_1(\lambda)$, $u_2(\lambda)$ be Lipschitz families of Sobolev functions, defined for $\lambda \in \Lambda_o$ such that (5.2.104), (5.2.127) hold with $\mathcal{R}_0 = \mathcal{R}_0(u_i)$ with $i = 1, 2$. Then for $\nu \geq 0$, for any $\lambda \in \Lambda^{\gamma_1}_\nu \cap \Lambda^{\gamma_2}_\nu$, with $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$, one has*

$$|\mathcal{R}_\nu(u_1) - \mathcal{R}_\nu(u_2)|_{\mathfrak{s}_0} \leq \varepsilon N^{-\alpha}_{\nu-1}||u_1 - u_2||_{\mathfrak{s}_0+\eta_3}, \tag{5.2.138a}$$

$$|\mathcal{R}_\nu(u_1) - \mathcal{R}_\nu(u_2)|_{\mathfrak{s}_0+\beta} \leq \varepsilon N_{\nu-1}||u_1 - u_2||_{\mathfrak{s}_0+\eta_3}, \tag{5.2.138b}$$

*and moreover, for $\nu \geq 1$, for any $s \in [\mathfrak{s}_0, \mathfrak{s}_0 + \beta]$, for any $(\sigma, j) \in \boldsymbol{C} \times \mathbb{Z}$ and $k = \pm j$,*

$$|(r^{\nu,k}_{\sigma,j}(u_2) - r^{\nu,k}_{\sigma,j}(u_1)) - (r^{\nu-1,k}_{s,j}(u_2) - r^{\nu-1,k}_{\sigma,j}(u_1))| \leq |\mathcal{R}_{\nu-1}(u_1) - \mathcal{R}_{\nu-1}(u_2)|_{\mathfrak{s}_0},$$

$$|(r^{\nu,k}_{\sigma,j}(u_2) - r^{\nu,k}_{\sigma,j}(u_1))| \leq \varepsilon C||u_1 - u_2||_{\mathfrak{s}_0+\eta_3}, \tag{5.2.139}$$

$$|b^\nu_j(u_1) - b^\nu_j(u_2)| \leq \varepsilon C||u_1 - u_2||_{\mathfrak{s}_0+\eta_3}. \tag{5.2.140}$$

$(\mathbf{S4})_\nu$ *Let $u_1, u_2$ be as in $(\mathbf{S3})_\nu$ and $0 < \rho < \gamma/2$. For any $\nu \geq 0$ one has that, if*

$$\begin{aligned} &C N^\tau_{\nu-1}||u_1 - u_2||^{\sup}_{\mathfrak{s}_0+\eta_3} \leq \rho\varepsilon \quad \Rightarrow \\ &P^\gamma_\nu(u_1) \subseteq P^{\gamma-\rho}_\nu(u_2), \quad \mathcal{O}^\gamma_\nu(u_1) \subseteq \mathcal{O}^{\gamma-\rho}_\nu(u_2). \end{aligned} \tag{5.2.141}$$

*Proof.* We start by proving that $(\mathbf{Si})_\mathbf{0}$ hold for $i = 0, \ldots, 4$.

$(\mathbf{S1})_\mathbf{0}$. Clearly the properties (5.2.132)-(5.2.133) hold by (5.2.118), (5.2.119) and the form of $\mu^0_k$ in (5.2.131), recall that $r^0_k = 0$ . Moreover, $m_2, |m_1|$ and $m_0$ real imply that $\mu^0_k$ are imaginary. In addition to this, our hypotheses guarantee that $\mathcal{R}_0$ and $\mathcal{L}_0$ are hamiltonian operators.

**(S2)$_0$**. We have to extend the eigenvalues $\mu_k^0$ from the set $\Lambda_0^\gamma$ to the entire $\Lambda$. Namely we extend the functions $m_2(\lambda), m_1(\lambda)$ and $m_0(\lambda)$ to a $\widetilde{m}_i(\lambda)$ for $i = 0, 1, 2$ which are Lipschitz in $\Lambda$, with the same sup norm and Lipschitz semi-norm, by Kirszbraum theorem.

**(S3)$_0$**. It holds by (5.1.17) and (5.1.18) for $\mathfrak{s}_0$, $\mathfrak{s}_0 + \beta$ using (5.2.104) and (5.2.123).

**(S4)$_0$**. By definition one has $\Lambda_0^\gamma(u_1) = \Lambda_o = \Lambda_0^{\gamma-\rho}(u_2)$, then the (5.2.141) follows trivially.

**Kam step**

In this Section we show in detail one step of the KAM iteration. In other words we will show how to define the transformation $\Phi_\nu$ and $\Psi_\nu$ that trasform the operator $\mathcal{L}_\nu$ in the operator $\mathcal{L}_{\nu+1}$. For simplicity we shall avoid to write the index, but we will only write $+$ instead of $\nu + 1$.

We consider a transformation of the form $\Phi = \exp(\Psi)$, with $\Psi := (\Psi_\sigma^{\sigma'})_{\sigma,\sigma'=\pm 1}$, acting on the operator

$$\mathcal{L} = \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D} + \mathcal{R}$$

with $\mathcal{D}$ and $\mathcal{R}$ as in (4.3.122), We define the operator

$$e^{ad(\Psi)} L := \sum_{m=0}^{\infty} \frac{1}{m!} [\Psi, L]^m, \quad \text{with} \quad [\Psi, L]^m = [\Psi, [\Psi, L]^{m-1}], \quad [\Psi, L] = \Psi L - L\Psi$$

acting on the matrices $L$. One has that

$$e^{ad(\Psi)} L = e^{-\Psi} L e^{\Psi}. \tag{5.2.142}$$

Clearly the (5.2.142) hold since $\Psi$ is a linear operator. Then, $\forall\, h \in \mathbf{H}^s$, by conjugation one has

$$\begin{aligned}
\Phi^{-1} \mathcal{L} \Phi &= e^{ad(\Psi)} (\omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}) + e^{ad(\Psi)} \mathcal{R} \\
&= \omega \cdot \partial_\varphi + \mathcal{D} + [\Psi, \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}] + \Pi_N \mathcal{R} \\
&\quad + \sum_{m \geq 2} \frac{1}{m!} [\Psi, \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}]^m + \Pi_N^\perp \mathcal{R} + \sum_{m \geq 1} \frac{1}{m!} [\Psi, \mathcal{R}]^m
\end{aligned} \tag{5.2.143}$$

where $\Pi_N$ is defined in (4.3.102). The smoothing operator $\Pi_N$ is necessary for technical reasons: it will be used in order to obtain suitable estimates on the high norms of the transformation $\Phi$.

In the following Lemma we will show how to solve the *homological* equation

$$\begin{aligned}
&[\Psi, \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}] + \Pi_N \mathcal{R} = [\mathcal{R}], \quad \text{where} \\
&[\mathcal{R}]_{\sigma,j}^{\sigma',j'}(\ell) := \begin{cases} (\mathcal{R})_{\sigma,j}^{\sigma',k}(0), & \sigma = \sigma',\ k = j, -j,\ \ell = 0 \\ 0 \quad \text{otherwise}, \end{cases}
\end{aligned} \tag{5.2.144}$$

for $k, k' \in \boldsymbol{C} \times \mathbb{N} \times \mathbb{Z}^d$.

**Lemma 5.2.104 (Homological equation).** *For any $\lambda \in \Lambda^{\gamma}_{\nu+1}$ there exists a unique solution $\Psi = \Psi(\varphi)$ of the homological equation (5.2.144), such that*

$$|\Psi|_{s,\gamma} \leq CN^{2\tau+1}\gamma^{-1}|\mathcal{R}|_{s,\gamma} \tag{5.2.145}$$

*Moreover, for $\gamma/2 \leq \gamma_1, \gamma_2 \leq 2\gamma$, and if $u_1(\lambda), u_2(\lambda)$ are Lipschitz functions, then $\forall\ s \in [\mathfrak{s}_0, \mathfrak{s}_0 + \beta]$, $\lambda \in \Lambda^{\gamma_1}_{+}(u_1) \cap \Lambda^{\gamma_2}_{+}(u_2)$, one has*

$$|\Delta_{12}\Psi|_s \leq CN^{2\tau+1}\gamma^{-1}\left(|\mathcal{R}(u_2)|_s||u_1 - u_2||_{\mathfrak{s}_0+\eta_2} + |\Delta_{12}\mathcal{R}|_s\right), \tag{5.2.146}$$

*where we define $\Delta_{12}\Psi = \Psi(u_1) - \Psi(u_2)$.*
*Finally, one has $\Phi : \mathbf{H}^s \to \mathbf{H}^s$ is symplectic.*

*Proof.* We rewrite the equation (5.2.144) on each component $k = (\sigma, j, p), k' = (\sigma', j', p')$ and we get the following matricial equation

$$i\omega \cdot (p - p')\Psi^{\sigma',\underline{j}',p'}_{\sigma,\underline{j},p} + \Omega_{\sigma,\underline{j}}\Psi^{\sigma',\underline{j}',p'}_{\sigma,\underline{j},p} - \Psi^{\sigma',\underline{j}',p'}_{\sigma,\underline{j},p}\Omega_{\sigma',\underline{j}'} = -\mathcal{R}^{\sigma',\underline{j}'}_{\sigma,\underline{j}}(p - p') \tag{5.2.147}$$

where $\Omega_{\sigma,\underline{j}}$ is defined in (5.2.128) and where we have set

$$\Psi^{\sigma',\underline{j}',p'}_{\sigma,\underline{j},p} := \begin{pmatrix} \Psi^{\sigma',j',p'}_{\sigma,j,p} & \Psi^{\sigma',-j',p'}_{\sigma,j,p} \\ \Psi^{\sigma',j',p'}_{\sigma,-j,p} & \Psi^{\sigma',-j',p'}_{\sigma,-j,p} \end{pmatrix} \tag{5.2.148}$$

the matrix block indexed by $(j, j')$. To solve equation (5.2.147) we can use Lemma 5.2.99 with $A := i\omega \cdot p\mathbb{1} + \Omega_{\sigma,\underline{j}}$ and $B = i\omega \cdot p'\mathbb{1} + \Omega_{\sigma',\underline{j}'}$. Hence if we write $\mu_{\sigma,h}$ and $\mu_{\sigma',h'}$ with $h = j, -j$ and $h' = j', -j'$ the eigenvalues respectively of $\Omega_{\sigma,\underline{j}}$ and $\Omega_{\sigma',\underline{j}'}$,

$$\begin{aligned} \|\Psi^{\sigma',\underline{j}',p'}_{\sigma,\underline{j},p}\|_\infty &\overset{(5.2.130)}{\leq} C\frac{\langle\ell\rangle^\tau\gamma^{-1}}{|\sigma j^2 - \sigma' j'^2|}\max_{h=j,-j,h'=j',-j'}|\mathcal{R}^{\sigma',h'}_{\sigma,h}(\ell)|, \\ &\qquad \sigma = \sigma', j \neq j', \quad \text{or } \sigma \neq \sigma',\ \forall j, j' \\ \|\Psi^{\sigma',\underline{j},p'}_{\sigma,\underline{j},p}\|_\infty &\overset{(5.2.130)}{\leq} C\langle\ell\rangle^\tau|j|\gamma^{-1}\max_{h=j,-j}|\mathcal{R}^{\sigma',h}_{\sigma,h}(\ell)|, \quad \sigma = \sigma',\ j = j', \end{aligned} \tag{5.2.149}$$

where we fixed $p - p' = \ell$. Clearly the solution $\Psi$ is Töpliz in time. Unfortunately bounds (5.2.149) are not sufficient in order to estimate the decay norm of the matrix $\Psi^{\sigma'}_{\sigma}$. Roughly speaking one needs to prove, for any $\ell$, that $\Psi^{\sigma',j'}_{\sigma,j}(\ell) \approx o(1/\langle j - j'\rangle^s)$, and $\Psi^{\sigma',-j'}_{\sigma,j} \approx o(1/\langle j + j'\rangle^s)$. Actually we are able to prove the following.

Assume that either $|j| \leq \frac{C}{\mathfrak{e}}$ or $|j'| \leq \frac{C}{\mathfrak{e}}$ for some large $C > 0$ and $\mathfrak{e}$ defined in (1.1.12). Assume also that

$$\max_{h=j,-j,h'=j',-j'}|\mathcal{R}^{\sigma',h'}_{\sigma,h}(\ell)| = |\mathcal{R}^{\sigma',j'}_{\sigma,j}(\ell)|. \tag{5.2.150}$$

By (5.2.149) we have that

$$
\begin{aligned}
(|\Psi_{\sigma,j}^{\sigma',j'}|^2+|\Psi_{\sigma,-j}^{\sigma',-j'}|^2)\langle j-j'\rangle^{2s}&+(|\Psi_{\sigma,j}^{\sigma',-j'}|^2+|\Psi_{\sigma,-j}^{\sigma',j'}|^2)\langle j+j'\rangle^{2s}\\
&\leq C\frac{\langle\ell\rangle^{2\tau}\gamma^{-2}}{|\sigma j^2-\sigma' j'^2|^2}|\mathcal{R}_{\sigma,j}^{\sigma',j'}(\ell)|^2\left(\langle j-j'\rangle^{2s}+\langle j+j'\rangle^{2s}\right)\\
&\leq \widetilde{C}\frac{\langle\ell\rangle^{2\tau}\gamma^{-2}}{|\sigma j^2-\sigma' j'^2|^2}|\mathcal{R}_{\sigma,j}^{\sigma',j'}(\ell)|^2\langle j-j'\rangle^{2s}
\end{aligned}
\tag{5.2.151}
$$

where we used the fact that, for a finite number of $j$ (or finite $j'$), one has

$$
\langle j+j'\rangle\leq K\langle j-j'\rangle,
$$

for some large $K=K(\mathfrak{c})>0$. Note also that the smaller is $\mathfrak{c}$ the larger is the constant $K$. If the (5.2.150) does not hold one can treat the other cases by reasoning as done in (5.2.151). Assume now that

$$
|j|,|j'|\geq\frac{C}{\mathfrak{c}}
\tag{5.2.152}
$$

holds. Here the situation is more delicate. Consider the matrices $\Omega_{\sigma,\underline{j}},\Omega_{\sigma',\underline{j'}}$ in equation (5.2.147) which have, by (5.2.131), eigenvalues $\mu_{\sigma,j}$, $\mu_{\sigma,-j}$ and $\mu_{\sigma',j'}$, $\mu_{\sigma',-j'}$ respectively. First of all one can note that by (5.2.152)

$$
|\mu_{\sigma,j}-\mu_{\sigma,-j}|,\geq|m_1|\langle j\rangle\geq c\varepsilon\mathfrak{c}\langle j\rangle,\quad |\mu_{\sigma',j'}-\mu_{\sigma',-j'}|\geq|m_1|\langle j'\rangle
\tag{5.2.153}
$$

by the (1.1.12). Hence we can define the invertible matrices

$$
U_{\sigma,\underline{j}}:=\begin{pmatrix}\frac{\Omega_{\sigma,-j}^{-j}-\mu_{\sigma,j}}{\mu_{\sigma,j}-\mu_{\sigma,-j}} & \frac{-\Omega_{\sigma,j}^{-j}}{\mu_{\sigma,j}-\mu_{\sigma,-j}}\\ \frac{-\Omega_{\sigma,-j}^{j}}{\mu_{\sigma,j}-\mu_{\sigma,-j}} & \frac{\Omega_{\sigma,j}^{j}-\mu_{\sigma,-j}}{\mu_{\sigma,j}-\mu_{\sigma,-j}}\end{pmatrix},
\tag{5.2.154}
$$

and moreover one can check that

$$
U_{\sigma,\underline{j}}^{-1}\Omega_{\sigma,\underline{j}}U_{\sigma,\underline{j}}=D_{\sigma,\underline{j}}=\begin{pmatrix}\mu_{\sigma,j} & 0\\ 0 & \mu_{\sigma,-j}\end{pmatrix},
\tag{5.2.155}
$$

In order to simplify the notation we set

$$
f_{\sigma,j}^{(1)}:=\frac{\Omega_{\sigma,-j}^{-j}-\mu_{\sigma,j}}{\mu_{\sigma,j}-\mu_{\sigma,-j}},\qquad f_{\sigma,j}^{(2)}:=\frac{\Omega_{\sigma,j}^{j}-\mu_{\sigma,-j}}{\mu_{\sigma,j}-\mu_{\sigma,-j}},\qquad c_{\sigma,j}:=\frac{-\Omega_{\sigma,j}^{-j}}{\mu_{\sigma,j}-\mu_{\sigma,-j}}.
\tag{5.2.156}
$$

First of all, by using (5.2.154), (5.2.153) and (5.2.132) one has

$$
|f_{\sigma,j}^{(1)}|+|f_{\sigma,j}^{(2)}|\leq 4\frac{C}{c\mathfrak{c}},\qquad |c_{\sigma,j}|\leq\frac{1}{c\varepsilon\mathfrak{c}}|r_j^{-j}|.
\tag{5.2.157}
$$

Hence one has

$$
U_\sigma:=\mathrm{diag}_{|j|\geq C/\mathfrak{c},j\in\mathbb{N}}U_{\sigma,\underline{j}},\qquad |U_\sigma|_{s,\gamma}\leq\frac{C}{|m_1|}|\mathcal{R}_\sigma^{\sigma'}|_{s,\gamma},
\tag{5.2.158}
$$

and moreover $U_\sigma$ diagonalizes the matrix $\Omega_\sigma = \mathrm{diag}_{|j|\geq C/\mathfrak{e}}\Omega_{\sigma,\underline{j}}$. Setting $U_\sigma^{-1}\Psi_\sigma^{\sigma'}U_\sigma = Y_\sigma^{\sigma'}$, equation (5.2.147), for $\sigma, \sigma' = \pm 1$, reads

$$\mathrm{i}\omega\cdot\partial_\varphi Y_\sigma^{\sigma'} + D_\sigma Y_\sigma^{\sigma'} - \sigma' Y_\sigma^{\sigma'} D_\sigma = U_\sigma^{-1} R_\sigma^{\sigma'} U_\sigma. \tag{5.2.159}$$

For $|\ell| \leq N$ we set

$$Y_{\sigma,j}^{\sigma',j'}(\ell) = \frac{(U_\sigma^{-1}\mathcal{R}_\sigma^{\sigma'}U_\sigma)_{\sigma,j}^{\sigma',j'}(\ell)}{\mathrm{i}\omega\cdot\ell + \mu_{\sigma,j} - \mu_{\sigma',j'}} \tag{5.2.160}$$

Now as done in (4.3.136), (4.3.137) and (4.3.141), we get the bound

$$|Y_\sigma^{\sigma'}|_s \leq \gamma^{-1}N^\tau |U_\sigma^{-1}\mathcal{R}_\sigma^{\sigma'}U_\sigma|_s, \tag{5.2.161}$$

where we used the estimates (5.2.130) on the small divisors.

By the definition, the estimate (5.2.158) and the interpolation properties in Lemma 4.3.62 we can bound the decay norm of $\Psi$ as

$$|\Psi|_s \leq C(s)\gamma^{-1}N^\tau|\mathcal{R}|_s, \tag{5.2.162}$$

using that $|\mathcal{R}|_s/|m_1| \leq C$ for some constant $C > 0$. Moreover the following hold:

**Lemma 5.2.105.** *Define the operator $A$ as*

$$A_k^{k'} = A_{\sigma,j}^{\sigma',j'}(\ell) := \begin{cases} \Psi_{\sigma,j}^{\sigma,j'}(\ell), & \sigma = \sigma' \in \boldsymbol{C},\ \ j = \pm j' \in \mathbb{Z}\ \ \ell \in \mathbb{Z}^d, \\ 0, & \text{otherwise}, \end{cases} \tag{5.2.163}$$

*then the operator $\Psi - A$ is regularizing and hold*

$$|D(\Psi - A)|_s \leq \gamma^{-1}N^\tau|\mathcal{R}|_s, \tag{5.2.164}$$

*where $D := diag\{j\}_{j\in\mathbb{Z}}$.*

This Lemma will be used in the study of the remainder of the conjugate operator. In particular we will use it to prove that the reminder is still in the class of operators described in (5.2.119).

Now we need a bound on the Lipschitz semi-norm of the transformation. Then, given $\omega_1, \omega_2 \in \Lambda_{\nu+1}^\gamma$, one has, for $k = (\sigma, j, p), k' = (\sigma', j', p') \in \boldsymbol{C}\times\mathbb{Z}\times\mathbb{Z}^d$, and $\ell := p - p'$,

$$\begin{aligned} \omega_1\cdot\ell\Big[\Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_1) - \Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_2)\Big] + \Omega_{\sigma,\underline{j}}(\omega_1)\Big[\Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_1) - \Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_2)\Big] + \\ - \Big[\Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_1) - \Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_2)\Big]\Omega_{\sigma',\underline{j'}}(\omega_1) \\ + (\omega_1 - \omega_2)\cdot\ell\Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_2) + \\ + \Big[\Omega_{\sigma,\underline{j}}(\omega_1) - \Omega_{\sigma,\underline{j}}(\omega_2)\Big]\Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_2) \\ + \Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_2)\Big[\Omega_{\sigma',\underline{j'}}(\omega_1) - \Omega_{\sigma',\underline{j'}}(\omega_2)\Big] = \\ = \mathcal{R}_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_1) - \mathcal{R}_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_1). \end{aligned} \tag{5.2.165}$$

First we can note that

$$
\begin{aligned}
|\Omega_{\sigma,j}^{\,j}(\omega_1)-\Omega_{\sigma',j'}^{\,j'}(\omega_2)| &\le |m_2(\omega_1)-m_2(\omega_2)||\sigma j^2-\sigma' j'^2|+\varepsilon\gamma^{-1}\\
&\quad+|m_1(\omega_1)-m_1(\omega_2)||\sigma j-\sigma' j'|+|m_0(\omega_1)-m_0(\omega_2)|\\
&\le C|\omega_1-\omega_2|(\varepsilon\gamma^{-1}|\sigma j^2-\sigma' j'^2|+\varepsilon\gamma^{-1}+\varepsilon\gamma^{-1})
\end{aligned}
\tag{5.2.166}
$$

where we used the (5.2.128), (5.2.132) and (5.1.15) to estimate the Lipschitz semi-norm of the constants $m_i$. Following the same reasoning, one can estimate the sup norm of the matrix $\Omega_{\sigma,\underline{j}}(\omega_1)-\Omega_{\sigma,\underline{j}}(\omega_2)$. Therefore by triangular inequality one has

$$
\begin{aligned}
&\|\Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_1)-\Psi_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_2)\|_{\infty}\lessdot|\mathcal{R}_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_1)-\mathcal{R}_{\sigma,\underline{j}}^{\sigma',\underline{j'}}(\ell,\omega_1)|_{max}N^{\tau}\gamma^{-1}+\\
&+|\omega_1-\omega_2|\Big(|\ell|+\varepsilon\gamma^{-1}|\sigma j^2-\sigma' j^{j'}|\Big)\\
&+|\omega_1-\omega_2|\left(\varepsilon\gamma^{-1}|\sigma j-\sigma' j'|\varepsilon\gamma^{-1}\right)\|\mathcal{R}_{\sigma,h}^{\sigma',h'}(\ell,\omega_2)\|_{\infty}\frac{N^{2\tau+1}\gamma^{-2}}{|\sigma j^2-\sigma' j'^2|},
\end{aligned}
\tag{5.2.167}
$$

for $|\ell|\le N$, $j\neq j'$ and $\varepsilon\gamma^{-1}\le 1$. As done for the estimate (5.2.162) for a finite number of $j$ of a finite number of $j'$ the bound (5.2.167) is sufficient to get, for $\omega\in\Lambda_{\nu+1}^{\gamma}$ and using also the bound (5.2.130) with $j=j'$, the estimate

$$
|\Psi|_{s,\gamma}:=|\Psi|_s^{\sup}+\gamma\sup_{\omega_1\neq\omega_2}\frac{|\Psi(\omega_1)-\Psi(\omega_2)|}{|\omega_1-\omega_2|}\le C\gamma^{-1}N^{2\tau+1}|\mathcal{R}|_{s,\gamma},
\tag{5.2.168}
$$

that is the (5.2.145).

On the other hand, in the case of (5.2.152), we can reason as follows. Consider the diagonalizing matrix $U_{\sigma,\underline{j}}$ defined in (5.2.155) and recall that by (5.2.158) also the lipschitz semi-norm of $U_\sigma$ is bounded by the lipschitz semi-norm of $\mathcal{R}_{\sigma}^{\sigma'}$. Hence by (5.2.159), (5.2.160), using again the interpolation properties of the decay norm in Lemma (4.3.62) one get the Lipschitz bound in (5.2.168). Note also that the Lemma 5.2.105 holds with $|\cdot|_{s,\gamma}$ and $N^{2\tau+1}$ instead of $|\cdot|_s$ and $N^{\tau}$.

The proof of the bound (5.2.146) is based on the same strategy used to proof (5.2.168). We refer to the proof of the bound (4.3.134) in Section 4.3.2 of Chapter 4.

Finally we show that $\Psi$ is an hamiltonian vector field, and hence the transformation $\Phi$ is symplectic. By hypothesis $\mathcal{R}$ is hamiltonian, hence by Lemma 5.2.100 we have

$$
\left(\overline{\mathcal{R}_{\sigma}^{\sigma}}\right)^T=-\mathcal{R}_{\sigma}^{\sigma},\quad \overline{\mathcal{R}_{\sigma}^{-\sigma}}=\mathcal{R}_{-\sigma}^{\sigma},\quad \overline{\mathcal{R}_{\sigma}^{\sigma'}}=\mathcal{R}_{-\sigma}^{-\sigma'},\quad \forall\,\sigma,\sigma'\in\boldsymbol{C}.
\tag{5.2.169}
$$

Moreover, by inductive hypothesis $(\mathbf{S1})_\nu$ one can note that

$$
\left(\overline{\Omega_\sigma}\right)^T=-\Omega_\sigma=\Omega_{-\sigma}.
\tag{5.2.170}
$$

By (5.2.169), (5.2.170) one can easily note that the solution of the equation

$$\omega \cdot \partial_\varphi \Psi_\sigma^{\sigma'} + \Omega_\sigma \Psi_\sigma^{\sigma'} - \Psi_\sigma^{\sigma'} \Omega_{\sigma'} = \mathcal{R}_\sigma^{\sigma'},$$

satisfies conditions in (5.2.169), hence, again by Lemma 5.2.100, $\Psi$ is hamiltonian. This concludes the proof of Lemma 5.2.104. □

Next Lemma concludes one step of our KAM iteration.

**Lemma 5.2.106 (The new operator $\mathcal{L}_+$).** *Consider the operator $\Phi = \exp(\Psi)$ defined in Lemma 5.2.104. Then the operator $\mathcal{L}_+ := \Phi^{-1}\mathcal{L}\Phi$ has the form*

$$\mathcal{L}_+ := \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}_+ + \mathcal{R}_+, \tag{5.2.171}$$

*where the diagonal part is*

$$\begin{aligned} \mathcal{D}_+ = \mathrm{diag}_{(\sigma,j)\in \boldsymbol{C}\times\mathbb{Z}}\{\Omega^+_{\sigma,\underline{j}}\}, \quad \Omega^+_{\sigma,\underline{j}}(\lambda) &= \begin{pmatrix} \Omega^{+,j}_{\sigma,j} & \Omega^{+,-j}_{\sigma,j} \\ \Omega^{+,j}_{\sigma,-j} & \Omega^{+,-j}_{\sigma,-j} \end{pmatrix}, \\ \Omega^{+,j}_{\sigma,j} &:= -i\sigma m_2 j^2 - i\sigma|m_1|j + i\sigma m_0 + i\sigma r_j^{+,j}, \\ \Omega^{+,-j}_{\sigma,j} &:= i\sigma r_j^{+,-j}, \\ r_j^{+,h} &:= r_j^h + \mathcal{R}^{\sigma,h}_{\sigma,j}(0), \quad h = \pm j. \end{aligned} \tag{5.2.172}$$

*with $(\sigma, j) \in \boldsymbol{C} \times \mathbb{Z}, \lambda \in \Lambda$. The eigenvalues $\mu^+_{\sigma,h}$, with $h = j, -j$, of $\Omega_{\sigma,\underline{j}}$ satisfy*

$$\begin{aligned} |r_j^{+,h} - r_j^h|^{lip} &\le |(\mathcal{R})^\sigma_\sigma|^{lip}_{\mathfrak{s}_0}, \\ |\mu^+_{\sigma,h} - \mu_{\sigma,h}|^{sup} &\le |(\mathcal{R})^\sigma_\sigma|_{\mathfrak{s}_0,\gamma}, \quad h = j, -j. \end{aligned} \tag{5.2.173}$$

*The remainder $\mathcal{R}_+$ is such that*

$$\begin{aligned} |\mathcal{R}_+|_s &\le_s N^{-\beta}|\mathcal{R}|_{s+\beta,\gamma} + N^{2\tau+1}\gamma^{-1}|\mathcal{R}|_{s,\gamma}|\mathcal{R}|_{\mathfrak{s}_0,\gamma}, \\ |\mathcal{R}_+|_{s+\beta} &\le_{s+\beta} |\mathcal{R}|_{s+\beta,\gamma} + N^{2\tau+1}\gamma^{-1}|\mathcal{R}|_{s+\beta,\gamma}|\mathcal{R}|_{\mathfrak{s}_0,\gamma}, \end{aligned} \tag{5.2.174}$$

*and $(\mathcal{R}_+)^\sigma_\sigma = O(\varepsilon\partial_x^{-1})$ while $(\mathcal{R}_+)^{-\sigma}_\sigma = O(\varepsilon)$ for $\sigma = \pm 1$. More precisely,*

$$|(\mathcal{R}_+)^{-\sigma}_\sigma|_s + |D(\mathcal{R}_+)^\sigma_\sigma|_s \lessdot |\mathcal{R}_+|_s, \ \sigma \in \boldsymbol{C}, \quad \text{where} \quad D := \mathrm{diag}_{j\in\mathbb{Z}}\{j\}. \tag{5.2.175}$$

*Finally, for $\gamma/2 \le \gamma_1, \gamma_2 \le 2\gamma$, and for $u_1(\lambda), u_2(\lambda)$ Lipschitz functions, then for any $s \in [\mathfrak{s}_0, \mathfrak{s}_0 + \beta]$ and $\lambda \in \Lambda_+^{\gamma_1}(u_1) \cap \Lambda_+^{\gamma_2}(u_2)$ one has*

$$\begin{aligned} |\Delta_{12}\mathcal{R}_+|_s \le& |\Pi_N^\perp \Delta_{12}\mathcal{R}|_s + +N^{2\tau+1}\gamma^{-1}\Big(|\mathcal{R}(u_1)|_s + |\mathcal{R}(u_2)|_s\Big)|\Delta_{12}\mathcal{R}|_{\mathfrak{s}_0} \\ &+N^{2\tau+1}\gamma^{-1}\Big(|\mathcal{R}(u_1)|_s + |\mathcal{R}(u_2)|_s\Big)\Big(|\mathcal{R}(u_1)|_{\mathfrak{s}_0} + |\mathcal{R}(u_2)|_{\mathfrak{s}_0}\Big)||u_1 - u_2||_{\mathfrak{s}_0+\eta_3} \\ &+N^{2\tau+1}\gamma^{-1}\Big(|\mathcal{R}(u_1)|_{\mathfrak{s}_0} + |\mathcal{R}(u_2)|_{\mathfrak{s}_0}\Big)|\Delta_{12}\mathcal{R}|_s \end{aligned} \tag{5.2.176}$$

*Proof.* The (5.2.172) follow by the (5.2.144). Note that the term $\mathcal{R}^{\sigma,k}_{\sigma,j}(0) = \mathcal{R}^{-\sigma,k}_{-\sigma,j}$ for $k = j, -j$ and hence the new correction $r_j^{+,h}$ does not depend on $\sigma$. Moreover, by (4.3.87) one has

$$|\Omega^{+,k}_{\sigma,j} - \Omega^{\ k}_{\sigma,j}|^{lip} \leq |(\mathcal{R})^{\sigma}_{\sigma}|^{lip}_{\mathfrak{s}_0}, \quad k = j, -j. \tag{5.2.177}$$

Moreover, one has

$$\begin{aligned} |\mu^+_{\sigma,j} - \mu_{\sigma,j}| &\leq 2 \sup_{h=\pm j} |r_h^{+,h} - r_h^h| + |j||b_j^+ - b_j| \\ &\leq 2 \sup_{h=\pm j} |r_h^{+,h} - r_h^h| + \frac{|j|}{|j|} \sup_{h=\pm j} |r_j^{+,h} - r_j^h| \overset{(5.2.177)}{\lessdot} |(\mathcal{R})^{\sigma}_{\sigma}|, \end{aligned} \tag{5.2.178}$$

then the (5.2.173) follows. Now, by (5.2.143) one has that

$$\mathcal{R}_+ := \Pi_N^{\perp}\mathcal{R} + \sum_{n\geq 2} \frac{1}{n!}[\Psi, \omega\cdot\partial_\varphi \mathbb{1} + \mathcal{D}]^n + \sum_{n\geq 1}\frac{1}{n!}[\Psi, \mathcal{R}]^n := \Pi_N^{\perp}\mathcal{R} + \mathcal{B}. \tag{5.2.179}$$

Here we used the simple fact that $[A, B]^n = [A, [A, B]]^{n-1}$ for any $n \geq 1$. Hence we can estimate

$$\begin{aligned} |\mathcal{R}_+|_{s,\gamma} &\leq_s |\Pi_N^{\perp}\mathcal{R}|_{s,\gamma} + \sum_{k\geq 2}\frac{1}{k!}|[\Psi, \Pi_N\mathcal{R}]^{k-1}|_{s,\gamma} + \sum_{n\geq 1}\frac{1}{n!}|[\Psi,\mathcal{R}]^n|_{s,\gamma} \\ &\leq_s |\Pi_N^{\perp}\mathcal{R}|_{s,\gamma} + \sum_{n\geq 1}\frac{1}{n!}|[\Psi,\mathcal{R}]^n|_{s,\gamma} \leq_s |\Pi_n^{\perp}\mathcal{R}|_s \\ &+ \sum_{n\geq 1}\frac{(nC(\mathfrak{s}_0))^{n-1}}{n!}|\Psi|^{n-1}_{\mathfrak{s}_0,\gamma}|\mathcal{R}|^{n-1}_{\mathfrak{s}_0,\gamma}(|\Psi|_{s,\gamma}|\mathcal{R}|_{\mathfrak{s}_0,\gamma} + |\Psi|_{\mathfrak{s}_0,\gamma}|\mathcal{R}|_{s,\gamma}) \\ &\overset{(4.3.103),(5.2.145)}{\leq} N^{-\beta}|\mathcal{R}|_{s+\beta,\gamma} + N^{2\tau+1}\gamma^{-1}|\mathcal{R}|_{s,\gamma}|\mathcal{R}|_{\mathfrak{s}_0,\gamma}, \end{aligned}$$

where we assumed that

$$\sum_{n\geq 1}\frac{n^{n-1}}{n!}C(\mathfrak{s}_0)^{n-1}|\Psi|^{n-1}_{\mathfrak{s}_0,\gamma}|\mathcal{R}|^{n-1}_{\mathfrak{s}_0,\gamma} < 1. \tag{5.2.180}$$

Now we have to estimate $\Delta_{12}\mathcal{R}_+$ defined for $\lambda \in \Lambda^{\gamma_1}(u_1) \cup \Lambda^{\gamma_2}(u_2)$. We write $\mathcal{R}_i := \mathcal{R}(u_i)$ for $i = 1, 2$. We first need a technical Lemma used to study the variation with respect to the function $u$, of the commutator between two operators.

**Lemma 5.2.107.** *Given operators $A(u), B(u)$ one has that the following identities hold for any $n \geq 1$:*

$$[A_1, B_1]^n = [A_1, \Delta_{12}B]^n + [A_1, B_2]^n; \tag{5.2.181}$$

$$[A_1, B_2]^n = \Big[A_1, [A_2, B_2]\Big]^{n-1} + \Big[A_1, [\Delta_{12}A, B_2]\Big]^{n-1}; \tag{5.2.182}$$

$$\begin{aligned} &\Big[A_1, [A_2, B_2]\Big]^{n-1} - [A_2, B_2]^n = (n-2)\Big[A_1, \big[\Delta_{12}A, [A_2, B_2]\big]\Big]^{n-2} \\ &+ \Big[\Delta_{12}A, [A_2, B_2]^{n-1}\Big]. \end{aligned} \tag{5.2.183}$$

*Proof.* We prove the identities by induction. Let us start from the (5.2.181). For $n = 1$ it clearly holds. We prove it for $n + 1$ assuming that (5.2.181) holds for $n$. One has

$$\begin{aligned} [A_1, \Delta_{12} B]^{n+1} + [A_1, B_2]^{n+1} &= \Big[A_1, [A_1, \Delta_{12} B]^n\Big] + \Big[A_1, [A_1, B_2]^n\Big] \\ &\overset{(5.2.181)}{=} \Big[A_1, [A_1, B_1]^n\Big] =: [A_1, B_1]^{n+1}. \end{aligned} \tag{5.2.184}$$

The remaining formulæ can be proved in the same way. ■

By using Lemma 5.2.107, one can rewrite the term $\mathcal{B}$ in (5.2.179). Then setting $A_s := |\mathcal{R}_1|_s + |\mathcal{R}_2|_s$ for any $s \geq 0$, and using (4.3.95) and (5.2.180), one obtains

$$\begin{aligned} |\Delta_{12}\mathcal{B}|_s \overset{(5.2.145),(5.2.146)}{\leq_s} & N^{2\tau+1}\gamma^{-1} A_s |\Delta_{12}\mathcal{R}|_{\mathfrak{s}_0} + N^{2\tau+1}\gamma^{-1} A_{\mathfrak{s}_0} |\Delta_{12}\mathcal{R}|_s \\ & + 2N^{4\tau+2}\gamma^{-1} A_s A_{\mathfrak{s}_0}^2 ||u_1 - u_2||_{\mathfrak{s}_0+\eta_2} \\ & + 2N^{4\tau+2}\gamma^{-2} A_s A_{\mathfrak{s}_0} |\Delta_{12}\mathcal{R}|_{\mathfrak{s}_0} + N^{4\tau+2}\gamma^{-2} A_s A_{\mathfrak{s}_0}^2 ||u_1 - u_2||_{\mathfrak{s}_0+\eta_2} \\ & + N^{4\tau+2}\gamma^{-2} A_{\mathfrak{s}_0}^2 |\Delta_{12}\mathcal{R}|_s, \end{aligned}$$

where we used the (5.2.145) and (5.2.146). If we assume that

$$N^{2\tau+1}\gamma^{-1} A_{\mathfrak{s}_0} \leq 1, \tag{5.2.185}$$

then, using also (4.3.103) we obtain the (5.2.176). Finally by using Lemma 5.2.105 one can note that $[\Psi, \mathcal{R}]_\sigma^\sigma = O(\varepsilon \partial_x^{-1})$ while $[\Psi, \mathcal{R}]_\sigma^{-\sigma} = O(\varepsilon)$ for $\sigma = \pm 1$, this implies that the new remainder $\mathcal{R}_+$ has the same properties. □

Clearly we proved Lemma 5.2.106 by assuming the (5.2.180) and (5.2.185). These hypotheses have to be verified inductively at each step. In the next Section we prove that the procedure described above, can be iterated infinitely many times.

### 5.2.2 Conclusions and Proof of Theorem 5.2.96

To complete the proof of Lemma 5.2.103 we proceed by induction. The proof of the iteration is essentially standard and based on the estimates of the previous Section.

We omit the proof of properties $(\mathbf{S1})_{\nu+1}, (\mathbf{S2})_{\nu+1}$ and $(\mathbf{S3})_{\nu+1}$ since one can repeat almost word by word the proof of Lemma 4.3.71 in Section 4.3. The $(S4)_{\nu+1}$ is fundamental different. The difference depends on the multiplicity of the eigenvalues. Moreover the result is weaker. This is why, in this case, the set of good parameters is smaller. We will see this fact in Section 6.

$(\mathbf{S4})_{\nu+1}$ Let $\omega \in \Lambda^{\gamma}_{\nu+1}$, then by (5.2.130) and the inductive hypothesis $(\mathbf{S4})_{\nu}$ one has that $\Lambda^{\gamma}_{\nu+1}(u_1) \subseteq \Lambda^{\gamma}_{\nu}(u_1) \subseteq \Lambda^{\gamma-\rho}_{\nu}(u_2) \subseteq \Lambda^{\gamma/2}_{\nu}(u_2)$. Hence the eigenvalues $\mu^{\nu}_{h}(\omega, u_2(\omega))$ are well defined by the $(\mathbf{S1})_{\nu}$. Now, since $\lambda \in \Lambda^{\gamma}_{\nu}(u_1) \cap \Lambda^{\gamma/2}_{\nu}(u_2)$, we have for $h = (\sigma, j) \in \boldsymbol{C} \times \mathbb{Z}$ and setting $h' = (\sigma', j') \in \boldsymbol{C} \times \mathbb{Z}$

$$
\begin{aligned}
|(\mu^{\nu}_{h} - \mu^{\nu}_{h'})(\omega, u_2(\omega)) - (\mu^{\nu}_{h} - \mu^{\nu}_{h'})(\omega, u_1(\omega))| &\leq |\sigma j^2 - \sigma' j'^2||m_2(u_1) - m_2(u_2)| \\
&+ |m_0(u_1) - m_0(u_2)||\sigma - \sigma'| + \max_{j} |r^{\nu,j}_{j}(\omega, u_2(\omega)) - r^{\nu,j}_{j}(\omega, u_1(\omega))| \\
&+ |j||b^{\nu}_{j}(u_1) - b^{\nu}_{j}(u_2)| + |j'||b^{\nu}_{j'}(u_1) - b^{\nu}_{j'}(u_2)| \\
&\overset{(5.1.15),(5.2.139),(5.2.140)}{\leq} \varepsilon C \left(|\sigma j^2 - \sigma' j'^2| + ||j| + |j'||\right) ||u_2 - u_1||_{\mathfrak{s}_0+\eta_2},
\end{aligned}
\tag{5.2.186}
$$

The (5.2.186) implies that for any $|\ell| \leq N_{\nu}$ and $j \neq \pm j'$,

$$
\begin{aligned}
|i\omega \cdot \ell + \mu^{\nu}_{\sigma,j}(u_2) - \mu^{\nu}_{\sigma',j'}(u_2)| &\overset{(5.2.130),(5.2.186)}{\geq} \gamma |\sigma j^2 - \sigma' j'^2| \langle \ell \rangle^{-\tau} \\
&- C|\sigma j^2 - \sigma' j'^2| ||u_2 - u_1||_{\mathfrak{s}_0+\eta_2} \\
&\overset{(\mathbf{S4})_{\nu}}{\geq} (\gamma - \rho)|\sigma j^2 - \sigma' j'^2| \langle \ell \rangle^{-\tau},
\end{aligned}
\tag{5.2.187}
$$

where we used that, for any $\lambda \in \Lambda_0$, one has $C\varepsilon N^{\tau}_{\nu} ||u_1 - u_2||_{\mathfrak{s}_0+\eta_2} \leq \rho$ (note that this condition is weaker with respect to the hypothesis in $(\mathbf{S4})_{\nu}$. Now, the (5.2.187), imply that if $\lambda \in \mathcal{P}^{\gamma}_{\nu+1}(u_1)$ then $\lambda \in \mathcal{P}^{\gamma-\rho}_{\nu+1}(u_2)$. Now assume that $\lambda \in \mathcal{O}^{\gamma}_{\nu+1}(u_1)$. We have two cases: if $|j| \geq 4|\omega||\ell|/\varepsilon\mathfrak{e}$, then we have no small divisors. Indeed one has

$$
\begin{aligned}
b^{\nu}_{j}(u)^2 &= \left(-2|m_1| + \frac{r^{\nu,j}_{j} - r^{\nu,-j}_{-j}}{j}\right)^2 + 4\frac{|r^{\nu,-j}_{j}|^2}{|j|^2} \geq \left(2|m_1| - \frac{\varepsilon C}{|j|}\right)^2 \\
&\overset{(5.1.15)}{\geq} |m_1|^2 \left(2 - \frac{\varepsilon C}{|j|\varepsilon\mathfrak{e}}\right)^2 \geq |m_1|^2 \left(2 - \frac{\varepsilon\mathfrak{e}}{4|\omega||\ell|}\right)^2 \geq \frac{|m_1|^2}{4} \geq \frac{(\varepsilon\mathfrak{e})^2}{4},
\end{aligned}
$$

for any $u$. Hence it is obvious that

$$
\begin{aligned}
|i\omega \cdot \ell + \mu^{\nu}_{\sigma,j}(u_2) - \mu^{\nu}_{\sigma,-j}(u_2)| &\geq \frac{4|\omega||\ell|}{\varepsilon\mathfrak{e}} |b^{\nu}_{j}(u_2)| - |\omega \cdot \ell| \\
&\geq |\omega||\ell| \geq \frac{\gamma - \rho}{\langle \ell \rangle^{\tau} \langle j \rangle}.
\end{aligned}
\tag{5.2.188}
$$

Let us consider the case $|j| \leq 4|\omega||\ell|/\varepsilon\mathfrak{e}$: one has

$$
\begin{aligned}
|i\omega \cdot \ell + \mu^{\nu}_{\sigma,j}(u_2) - \mu^{\nu}_{\sigma,-j}(u_2)| &\overset{(5.2.130),(5.2.186)}{\geq} \gamma \langle \ell \rangle^{-\tau} \langle j \rangle^{-1} - \varepsilon C|j| ||u_2 - u_1||_{\mathfrak{s}_0+\eta_2} \\
&\geq \frac{1}{\langle \ell \rangle^{\tau} \langle j \rangle} \left(\gamma - \varepsilon |j|^2 C N^{-\alpha+\tau+2}_{\nu}\right) \geq \frac{\gamma - \rho}{\langle \ell \rangle^{\tau} \langle j \rangle}
\end{aligned}
$$

that is the $(\mathbf{S4})_{\nu+1}$.

*Proof of Theorem 5.2.96*

We want apply Lemma 5.2.103 to the linear operator $\mathcal{L}_0 = \mathcal{L}_7$ defined in (5.1.13) where

$$\mathcal{R}_0 := \begin{pmatrix} 0 & q_0(\varphi, x) \\ -\bar{q}_0(\varphi, x) & 0 \end{pmatrix} + \mathcal{R}_7,$$

with $\mathcal{R}_7$ defined in (5.1.92). One has that $\mathcal{R}_0$ satisfies the $(iii)$ of Lemma 5.1.87. Then

$$\begin{aligned} |\mathcal{R}_0|_{\mathfrak{s}_0+\beta} &\overset{(5.1.17)}{\leq} \varepsilon C(\mathfrak{s}_0+\beta)(1+||\mathbf{u}||_{\beta+\mathfrak{s}_0+\eta_1,\gamma}) \overset{(4.3.88)}{\leq} 2\varepsilon C(\mathfrak{s}_0+\beta), \qquad \Rightarrow \\ & N_0^{C_0}|\mathcal{R}_0|^0_{\mathfrak{s}_0+\beta}\gamma^{-1} \leq 1, \end{aligned} \tag{5.2.189}$$

if $\varepsilon\gamma^{-1} \leq \epsilon_0$ is small enough, that is the (5.2.127). Then we have to prove that in the set $\cap_{\nu\geq 0}\Lambda_\nu^\gamma$ there exists a final transformation

$$\Phi_\infty = \lim_{\nu\to\infty} \widetilde{\Phi}_\nu = \lim_{\nu\to\infty} \Phi_0 \circ \Phi_1 \circ \ldots \circ \Phi_\nu. \tag{5.2.190}$$

and the normal form

$$\Omega^\infty_{\sigma,\underline{j}} := \Omega^\infty_{\sigma,\underline{j}}(\lambda) = \lim_{\nu\to+\infty} \widetilde{\Omega}^\nu_{\sigma,\underline{j}}(\lambda) = \widetilde{\Omega}^0_{\sigma,\underline{j}}(\lambda) + \lim_{\nu\to+\infty} \begin{pmatrix} i\sigma\widetilde{r}_j^{\nu,j} & i\sigma\widetilde{r}_j^{\nu,-j} \\ i\sigma\widetilde{r}_{-j}^{\nu,j} & i\sigma\widetilde{r}_{-j}^{\nu,-j} \end{pmatrix}. \tag{5.2.191}$$

The proof that limits in (5.2.190) and (5.2.191) exist uses the bounds of Lemma 5.2.103. We refer the reader Section 4.3.2 for more details.

The following Lemma gives us a connection between the Cantor sets defined in Lemma 5.2.103 and Theorem 5.2.96. Again the proof is omitted since it is essentially the same of Lemma 4.3.75 in Section 4.3.

**Lemma 5.2.108.** *One has that*

$$\Lambda_\infty^{2\gamma} \subset \cap_{\nu\geq 0}\Lambda_\nu^\gamma. \tag{5.2.192}$$

Since one prove that in $\Lambda_\infty^{2\gamma}$ the limit in (5.2.190) exists in norm $|\cdot|_{s,\gamma}$ one has

$$\begin{aligned} \mathcal{L}_\nu &\overset{(5.2.128)}{=} \omega\cdot\partial_\varphi\mathbb{1} + \mathcal{D}_\nu + \mathcal{R}_\nu \overset{|\cdot|_{s,\gamma}}{\to} \omega\cdot\partial_\varphi\mathbb{1} + \mathcal{D}_\infty =: \mathcal{L}_\infty, \\ \mathcal{D}_\infty &:= diag_{(\sigma,j)\in C\times\mathbb{Z}}\Omega^\infty_{\sigma,\underline{j}}. \end{aligned} \tag{5.2.193}$$

and moreover

$$\mathcal{L}_\infty = \Phi_\infty^{-1} \circ \mathcal{L}_0 \circ \Phi_\infty, \tag{5.2.194}$$

that is the (5.2.108), while the (5.2.107) follows by the smallness in (5.2.132) and the convergence Finally, Lemma 4.3.62, Lemma 4.3.64 and (5.2.110) implies the bounds (5.2.111). This concludes the proof. ∎

## 5.3 Inversion of the linearized operator

In this Section we prove the invertibility of $\mathcal{L}(u)$, and consequently of $d_uF(u)$ (see 2.3.15), by showing the appropriate tame estimates on the inverse. The following Lemma resume the results obtained in the previous Sections.

We have the following result

**Lemma 5.3.109.** *Let* $\mathcal{L} = W_1\mathcal{L}_\infty W_2^{-1}$ *where*

$$W_i = \mathcal{V}_i\Phi_\infty, \quad \mathcal{V}_1 := \mathcal{T}_1\mathcal{T}_2\mathcal{T}_3\rho\mathcal{T}_4\mathcal{T}_5\mathcal{T}_6\mathcal{T}_7, \quad \mathcal{V}_2 = \mathcal{T}_1\mathcal{T}_2\mathcal{T}_3\mathcal{T}_4\mathcal{T}_5\mathcal{T}_6\mathcal{T}_7. \tag{5.3.195}$$

*where* $\mathcal{V}_i$ *and* $\Phi_\infty$ *are defined in Lemmata 5.1.87 and 5.2.96. Let* $\mathfrak{s}_0 \leq s \leq q-\beta-\eta_1-2$*, with* $\eta_1$ *define in (5.1.11) and* $\beta$ *in Theorem (5.2.96). Then, for* $\varepsilon\gamma^{-1}$ *small enough, and*

$$||u||_{\mathfrak{s}_0+\beta+\eta_1+2,\gamma} \leq 1, \tag{5.3.196}$$

*one has for any* $\lambda \in \Lambda_\infty^{2\gamma}$*,*

$$||W_ih||_{s,\gamma} + ||W_i^{-1}h||_{s,\gamma} \leq C(s)\left(||h||_{s+2,\gamma} + ||u||_{s+\beta+\eta_1+4,\gamma}||h||_{\mathfrak{s}_0,\gamma}\right), \tag{5.3.197}$$

*for* $i = 0, 1$*. Moreover,* $W_i$ *and* $W_i^{-1}$ *symplectic.*

*Proof.* Each $W_i$ is composition of two operators, the $\mathcal{V}_i$ satisfy the (5.1.14) while $\Phi_\infty$ satisfies (5.2.110). We use Lemma 4.3.62 in order to pass to the operatorial norm. Then Lemma A.168 and (A.2) with $p = s - \mathfrak{s}_0$, $q = 2$ implies the bounds (5.3.197). Moreover the transformations $W_i$ and $W_i^{-1}$ symplectic because they are composition of symplectic transformations $\mathcal{V}_i$,$\mathcal{V}_i^{-1}$ and $\Phi_\infty$, $\Phi_\infty^{-1}$ . □

Thanks to Lemma 5.3.109 the proof of Proposition 5.0.85 is almost concluded. We fix the constants $\eta = \eta_1+\beta+2$ (the constant $\eta$ has to be chosen) and $q > \mathfrak{s}_0+\eta$. Let $\Omega_{\sigma,j}^{j}$ and $\Omega_{\sigma,j}^{-j}$ be the functions defined in (5.2.191), and consequently $\mu_{\sigma,j}$ the eigenvalues of the matrices $\Omega_{\sigma,\underline{j}}$. Therefore by Lemmata 5.2.96 and 5.3.109 item $(i)$ in Proposition 5.0.85 hold. The item $(ii)$ follows simply by applying a dynamical point of view to our operator. Indeed by Lemma 5.1.95 and (5.2.111) in Lemma 5.2.96 one has that each transformation in (5.3.195) is a quasi-periodic in time transformation of the phase space $\mathbf{H}_x^s$ plus a reparametrization of the time with $\mathcal{T}_3$. Under a transformation of the form $\mathbf{u} = A(\omega t)\mathbf{v}$, one has that

$$\begin{aligned} &\partial_t\mathbf{u} = L(\omega t)\mathbf{u} \quad \leftrightarrow \quad \partial_t\mathbf{v} = L_+(\omega t)\mathbf{v}, \\ &L_+(\omega t) = A(\omega t)^{-1}L(\omega t)A(\omega t) - A(\omega t)^{-1}\partial_tA(\omega t) \end{aligned} \tag{5.3.198}$$

by conjugation. Moreover the transformation $A(\omega t)$ acts on the functions $\mathbf{u}(\varphi, x)$ as

$$\begin{aligned}(A\mathbf{u})(\varphi, x) := (A(\varphi)\mathbf{u}(\varphi,\cdot))(x) := A(\varphi)\mathbf{u}(\varphi, x),\\ (A^{-1}\mathbf{u})(\varphi, x) = A^{-1}(\varphi)\mathbf{u}(\varphi, x).\end{aligned} \tag{5.3.199}$$

Then on the space of quasi-periodic functions one has that the operator

$$\mathcal{L} := \omega \cdot \partial_\varphi - L(\varphi), \tag{5.3.200}$$

associated to the system (5.3.198), is transformed by $A$ into

$$A^{-1}\mathcal{L}A = \omega \cdot \partial_\varphi - L_+(\varphi), \tag{5.3.201}$$

that represent the system in (5.3.198) acting on quasi-periodic functions. The same consideration hold for a transformation of the type $\mathcal{T}_3$ as explained in Section 4.4.

Now we prove the following Lemma that is the equivalent result of Lemma 4.4.77 in the Hamiltonian case.

**Lemma 5.3.110 (Right inverse of $\mathcal{L}$).** *Under the hypotheses of Proposition 5.0.85, let us set*

$$\zeta := 4\tau + \eta + 8 \tag{5.3.202}$$

*where $\eta$ is fixed in Proposition 5.0.85. Consider a Lipschitz family $\mathbf{u}(\omega)$ with $\omega \in \Lambda_o \subseteq \Lambda \subseteq \mathbb{R}^d$ such that*

$$||\mathbf{u}||_{\mathfrak{s}_0+\zeta,\gamma} \leq 1. \tag{5.3.203}$$

*Define the set*

$$\mathcal{P}^{2\gamma}_\infty(u) := \left\{ \begin{array}{ll} \omega \in \Lambda_o : & |\omega\cdot\ell+\mu_{\sigma,j}(\omega)| \geq \frac{2\gamma\langle j\rangle^2}{\langle\ell\rangle^\tau}, \\ & \ell \in \mathbb{Z}^d, \sigma, \in \boldsymbol{C}, j \in \mathbb{Z} \end{array} \right\}. \tag{5.3.204}$$

*There exists $\epsilon_0$, depending only on the data of the problem, such that if $\varepsilon\gamma^{-1} < \epsilon_0$ then, for any $\omega \in \Lambda^{2\gamma}_\infty(u) \cap \mathcal{P}^{2\gamma}_\infty(u)$ (see (5.0.4)), and for any Lipschitz family $g(\omega) \in \mathbf{H}^s$, the equation $\mathcal{L}h := \mathcal{L}(\omega, u(\omega))h = g$, where $\mathcal{L}$ is the linearized operator $\mathcal{L}$ in* (2.3.36), *admits a solution*

$$h := \mathcal{L}^{-1}g := W_2\mathcal{L}_\infty^{-1}W_1^{-1}g \tag{5.3.205}$$

*such that*

$$||h||_{s,\gamma} \leq C(s)\gamma^{-1}\left(||g||_{s+2\tau+5,\gamma} + ||u||_{s+\zeta,\gamma}||g||_{\mathfrak{s}_0,\gamma}\right), \quad \mathfrak{s}_0 \leq s \leq q - \zeta. \tag{5.3.206}$$

*Proof.* A direct consequence of Lemma 5.3.109 is that, once one has conjugated the operator $\mathcal{L}$ in (2.3.36) to a block-diagonal operator $\mathcal{L}_\infty$ in (5.2.108) is essentially trivial to invert it:

**Lemma 5.3.111.** *For $g \in \mathbf{H}^s$, consider the equation*

$$\mathcal{L}_\infty(u)h = g. \tag{5.3.207}$$

*If $\omega \in \Lambda^{2\gamma}_\infty(\mathbf{u}) \cap \mathcal{P}^{2\gamma}_\infty(u)$ (defined in* (5.0.4) *and* (5.3.204)*), then there exists a unique solution $\mathcal{L}^{-1}_\infty g := h \in \mathbf{H}^s$. Moreover, for all Lipschitz family $g := g(\omega) \in \mathbf{H}^s$ one has*

$$||\mathcal{L}^{-1}_\infty g||_{s,\gamma} \le C\gamma^{-1}||g||_{s+2\tau+1,\gamma}. \tag{5.3.208}$$

*Proof.* One can follows the same strategy used for Lemma refinverselinfty in Section 4.4 and conclude using Lemma 5.2.98. □

In order to conclude the proof of Lemma 4.4.77 it is sufficient to collect the results of Lemmata 5.3.109 and 5.3.111. In particular one uses (5.3.197) and (5.3.208) to obtain the estimate

$$\begin{aligned} ||h||_{s,\gamma} &= \|W_2\mathcal{L}^{-1}_\infty W_1^{-1} g\|_{s,\gamma} \\ &\le C(s)\gamma^{-1}\left(||g||_{s+2\tau+5,\gamma} + ||u||_{s+4\tau+\beta+10+\eta_1,\gamma}||g||_{\mathfrak{s}_0,\gamma}\right), \end{aligned} \tag{5.3.209}$$

□

Note that by Lemma 2.3.15 the estimates (5.3.209) holds also for the linearized operator $d_u\mathcal{F}(u)$.

## 5.4 Measure estimates

In Section 5.1, 5.2 and 5.3 we prove that in the set $\Lambda^{2\gamma}_\infty(\mathbf{u}_n) \cap \mathcal{P}^{2\gamma}_\infty(u)$ we have good bounds (see (5.3.206)) on the inverse of $\mathcal{L}(u_n)$ . We also give a precise characterization of this set in terms of the eigenvalues of $\mathcal{L}$. Now in the Nash-Moser proposition 4.0.46 we defined in an implicit way the sets $\mathcal{G}_n$ in order to ensure bounds on the inverse of $\mathcal{L}(u_n)$. In this section we prove Proposition 5.4.112 that is the analogous of Proposition 4.5.79 of Section 4.5.

**Proposition 5.4.112** (**Measure estimates**)**.** *Set $\gamma_n := (1 + 2^{-n})\gamma$ and consider the set $\mathcal{G}_\infty$ of Proposition 4.0.46 with $\mu = \zeta$ defined in Lemma 5.3.110 and fix $\gamma := \varepsilon^a$ for some $a \in (0, 1)$. We have*

$$\cap_{n\ge 0}\,\mathcal{P}^{2\gamma_n}_\infty(u_n) \cap \Lambda^{2\gamma_n}_\infty(u_n) \subseteq \mathcal{G}_\infty, \tag{5.4.210a}$$

$$|\Lambda\backslash\mathcal{G}_\infty| \to 0, \ \text{ as } \ \varepsilon \to 0. \tag{5.4.210b}$$

**Proof of Proposition 5.4.112**. Let $(u_n)_{\ge 0}$ be the sequence of approximate solutions introduced in Proposition 4.0.46 which is well defined in $\mathcal{G}_n$ and satisfies the hypothesis of Proposition 5.0.85. $\mathcal{G}_n$

in turn is defined in Definition 3.1.17. We now define inductively a sequence of nested sets $G_n \cap H_n$ for $n \geq 0$. Set $G_0 \cap H_0 = \Lambda$ and

$$G_{n+1} := \left\{ \begin{array}{c} \omega \in G_n \ : |i\omega \cdot \ell + \mu_{\sigma,j}(u_n) - \mu_{\sigma',j'}(u_n)| \geq \dfrac{2\gamma_n |\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^\tau}, \\ \forall \ell \in \mathbb{Z}^n, \ \ \sigma, \sigma' \in \boldsymbol{C}, \ \ j, j' \in \mathbb{Z} \end{array} \right\},$$
$$H_{n+1} := \left\{ \begin{array}{c} \omega \in H_n \ : |i\omega \cdot \ell + \mu_{\sigma,j}(u_n) - \mu_{\sigma,-j}(u_n)| \geq \dfrac{2\gamma_n}{\langle \ell \rangle^\tau \langle j \rangle}, \\ \forall \ell \in \mathbb{Z}^n \backslash \{0\}, \ \ \sigma \in \boldsymbol{C}, \ \ j \in \mathbb{Z} \end{array} \right\}, \tag{5.4.211}$$
$$P_{n+1} := \left\{ \begin{array}{c} \omega \in P_n \ : |i\omega \cdot \ell + \mu_{\sigma,j}(u_n)| \geq \dfrac{2\gamma_n \langle j \rangle^2}{\langle \ell \rangle^\tau}, \\ \forall \ell \in \mathbb{Z}^n, \ \ \sigma \in \boldsymbol{C}, \ \ j \in \mathbb{Z} \end{array} \right\},$$

Recall that $\mu_{\sigma,j}(u_n)$ and $\mu_{\sigma,-j}(u_n)$ are the eigenvalues of the matrices $\Omega_{\sigma,\underline{j}}$ defined in Proposition 5.0.85 in (5.0.1) The following Lemma implies (5.4.210a).

**Lemma 5.4.113.** *Under the Hypotheses of Proposition 5.4.112, for any $n \geq 0$, one has*

$$P_{n+1} \cap G_{n+1} \cap H_{n+1} \subseteq \mathcal{G}_{n+1}. \tag{5.4.212}$$

*Proof.* For any $n \geq 0$ and if $\lambda \in G_{n+1}$, one has by Lemmata 5.3.111 and 4.4.77, (recalling that $\gamma \leq \gamma_n \leq 2\gamma$ and $2\tau + 5 < \zeta$)

$$\begin{aligned} &||\mathcal{L}^{-1}(u_n)h||_{s,\gamma} \leq C(s)\gamma^{-1}\left(||h||_{s+\zeta,\gamma} + ||u_n||_{s+\zeta,\gamma}||h||_{\mathfrak{s}_0,\gamma}\right), \\ &||\mathcal{L}^{-1}(u_n)||_{\mathfrak{s}_0,\gamma} \leq C(\mathfrak{s}_0)\gamma^{-1}N_n^\zeta ||h||_{\mathfrak{s}_0,\gamma}, \end{aligned} \tag{5.4.213}$$

for $\mathfrak{s}_0 \leq s \leq q - \mu$, for any $h(\lambda)$ Lipschitz family. The (5.4.213) are nothing but the (3.1.4) in Definition 3.1.17 with $\mu = \zeta$ . It represents the loss of regularity that you have when you perform the regularization procedure in Section 5.1 and during the diagonalization algorithm in Section 5.2. This justifies our choice of $\mu$ in Proposition 5.4.112. □

Now we prove formula (5.4.210b) that is the most delicate point. It turns out, by an explicit computation, that we can write for $j \neq 0$,

$$\mu_{\sigma,j} - \mu_{\sigma,-j} := i\sigma\sqrt{(-2|m_1|j + r_j^j - r_{-j}^{-j})^2 + 4|r_j^{-j}|} := jb_j = jb_j(u_n), \tag{5.4.214}$$

where $r_j^k$, for $j, k \in \mathbb{N}$ are the coefficients of the matrix $R_{\sigma,\underline{j}}$ in (5.2.105), and we define

$$\psi(\omega, u_n) := \omega \cdot \ell + jb_j(u_n). \tag{5.4.215}$$

Now we write for any $\ell \in \mathbb{Z}^d \backslash \{0\}$ and $j \in \mathbb{Z}$,

$$H_n := \bigcap_{\substack{\sigma \in \boldsymbol{C}, \\ (\ell,j) \in \mathbb{Z}^{d+1}}} A^{\sigma}_{\ell,j}(u_n) := \bigcap_{\substack{\sigma \in \boldsymbol{C}, \\ (\ell,j) \in \mathbb{Z}^{d+1}}} \left\{ \omega \in H_{n-1} : |i\omega \cdot \ell + jb_j(u_n)| \geq \frac{\gamma_n}{\langle j \rangle \langle \ell \rangle^{\tau}} \right\}. \tag{5.4.216}$$

Clearly one need to estimate the measure of $\bigcap_{n \geq 0} H_n$. The strategy to get such estimate is quite standard and it is the following:

**a**. First one give an estimate of the resonant set for fixed $(\sigma, j, \ell) \in \boldsymbol{C} \times \mathbb{Z} \times \mathbb{Z}^d$ (namely $(A^{\sigma}_{\ell,j})^c$). This point require a lower bound on the Lipschitz sub-norm of the function $\psi$ in (5.4.215). In this way we can give an estimate of the measure of the bad set using the standard arguments to estimate the measure of sub-levels of Lipschitz functions. This is in general non trivial but in the case of the sets $G_n$ and $P_n$ there is a well established strategy to follow that uses that $\mu_{\sigma,j} \sim O(j^2)$. In the case of the sets $H_n$ the problem is more difficult since $\mu_{\sigma,j} \sim O(\varepsilon j)$, hence, even if $j$ is large, it could happen that $\mu_{\sigma,j} \sim \omega \cdot \ell$. However we prove such lower bound (see (5.4.230)) using result of Lemma 5.4.117 and non-degeneracy condition on $m_1$ (see (5.1.16)). Moreveor we use deeply the fact that we have $d$ parameters $\omega_i$ for $i = 1, \ldots, d$ to move. On the contrary in Section 4.5 (see also [31]) we performed the estimates by choosing a diophantine direction $\bar{\omega}$ and using as frequency the vector $\omega = \lambda \bar{\omega}$ with $\lambda \in [1/2, 3/2]$, hence using just one parameter. In this case this is not possible.

**b**. Item **a**. provides and estimate like $|(A^{\sigma}_{j,\ell})^c| \sim \gamma/(j|\ell|^{\tau})$. The second point is to have some summability of the series in $j$ since one need to control $\bigcup_{j,\ell}(A^{\sigma}_{\ell,j})^c$. One can sum over $\ell$ by choosing $\tau$ large enough. In principle on can think to weaker the Melnikov conditions and ask for a lower bound of the type

$$|\psi| \geq \gamma/|j|^2|\ell|^{\tau}. \tag{5.4.217}$$

This can cause two problems. If one ask (5.4.217) it may be very difficult to prove the lower bound on the Lipschitz norm. Secondly in the reduction algorithm one must have a remainder $\mathcal{R}$ that support the loss of 2 derivatives in the space. Our strategy is different: we use results in Lemmata 5.4.115 and 5.4.116 to prove that the number of $j$ for which $(A^{\sigma}_{\ell,j})^c \neq \emptyset$ is controlled by $|\ell|$.

**c**. Finally one has to prove some "relation" between the sets $H_n$ and $H_k$ for $k \neq n$. Indeed the first two points imply only that the set $H_n$ has large measure as $\varepsilon \to 0$. But in principle as $n$ varies this sets can be unrelated, so that the intersection can be empty. Roughly speaking in Lemma 5.4.118 we prove that lots of resonances at the step $n$ have been already removed at the step $n-1$. In other words we prove that, if $|\ell|$ is sufficiently small, if $\psi(u_{n-1})$ satisfies the Melnikov conditions, then also $\psi(u_n)$ automatically has the good bounds. Again this point is different from the case studied in Section 4.5. Indeed with double eigenvalues one is able to prove the previous claim only for $n$ large enough and not for any $n$. This is the reason in this case the set of good parameters is small, but in any case of full measure.

In the following Lemma we resume the key result one need to prove Proposition 5.4.112.

**Lemma 5.4.114.** *For any $n \geq 0$ one has*

$$|P_n \backslash P_{n+1}|, |G_n \backslash G_{n+1}|, |H_n \backslash H_{n+1}| \leq C\sqrt{\gamma}. \tag{5.4.218}$$

*Moreover, if $n \geq \bar{n}(\varepsilon)$ (where $\bar{n}(\varepsilon)$ is defined in Lemma 5.4.118), then one has*

$$|P_n \backslash P_{n+1}|, |G_n \backslash G_{n+1}|, |H_n \backslash H_{n+1}| \leq C\sqrt{\gamma} N_n^{-1}. \tag{5.4.219}$$

*In particular $\bar{n}(\varepsilon)$ has the form*

$$\bar{n}(\varepsilon) := a\mathrm{loglog}\left[b\frac{1}{c\gamma\varepsilon}\right], \tag{5.4.220}$$

*with $a, b, c > 0$ independent on $\varepsilon$.*

By Lemma 5.4.114 follows the (4.5.220b). Indeed on one hand we have

$$|\Lambda \backslash \cap_{n\geq 0} H_n| \leq \sum_{n=0}^{\bar{n}(\varepsilon)} |H_n \backslash H_{n+1}| + \sum_{n > \bar{n}(\varepsilon)} |H_n \backslash H_{n+1}| \leq C\gamma \bar{n}(\varepsilon). \tag{5.4.221}$$

The same bounds holds for $|\Lambda \backslash \cap_{n\geq 0} G_n|, |\Lambda \backslash \cap_{n\geq 0} P_n|$. Now, fixing $\gamma := \gamma(\varepsilon) = \varepsilon^a$ with $a \in (0,1)$, one has that

$$|\Lambda \backslash \mathcal{G}_\infty| \leq C\sqrt{\gamma(\varepsilon)}(1 + \bar{n}(\varepsilon)) \to 0, \quad \text{as} \quad \varepsilon \to 0.$$

This concludes the proof of Proposition 5.4.112. It remains to check Lemma 5.4.114 following the strategy in three point explained above. We will give the complete proof only for the sets $H_n$ that is more difficult. The inductive estimates on $G_n$ and $P_n$ is very similar, anyway one can follows essentially word by word from the proof of Proposition 4.5.79 in Section 4.5. Similar measure estimates can be also found in [5].

**Lemma 5.4.115.** *If $|b_j||j| \geq 2|\omega \cdot \ell|$ or $|b_j||j| \leq |\omega \cdot \ell|/2$ then $(A^\sigma_{\ell,j}(u_n))^c = \emptyset$.*

*Proof.* Lemma follows by the fact that $\omega$ is diophantine with constant $\tau_0$ and $\tau > \tau_0$ and from the smallness of $|m_1|$. □

Thanks Lemma 5.4.115 in the following we will consider only the $j \in \mathcal{S}_{\ell,n} \subseteq \mathbb{Z}$ where

$$\mathcal{S}_{\ell,n} := \left\{ j \in \mathbb{Z} \ \frac{|\omega \cdot \ell|}{2} \leq |j| b_j(u_n) \leq 2|\omega \cdot \ell| \right\} \tag{5.4.222}$$

for some constant $C > 0$. In order to estimate the measure of $(A^\sigma_{\ell,j}(u_n))^c$ we need the following technical Lemma.

**Lemma 5.4.116.** *If $j \in \mathcal{S}_{\ell,n} \cap (A_{\ell,n})^c$, where*

$$A_\ell := \{j \in \mathbb{Z} : |j| \le 4|\ell| C/\mathfrak{e}\},$$

*then one has that $|b_j(u_n)| \ge |m_1(u_n)|/2$.*

*Proof.* It follow by

$$\begin{aligned} b_j^2 &= \left(-2|m_1| + \frac{r_j^j - r_{-j}^{-j}}{j}\right)^2 + 4\frac{|r_j^{-j}|^2}{|j|^2} \ge \left(2|m_1| - \frac{\varepsilon C}{|j|}\right)^2 \\ &\overset{(5.1.15)}{\ge} |m_1|^2 \left(2 - \frac{\varepsilon C}{|j|\varepsilon\mathfrak{e}}\right)^2 \ge |m_1|^2 \left(2 - \frac{1}{4|\ell|}\right)^2 \ge \frac{|m_1|^2}{4}. \end{aligned} \tag{5.4.223}$$

■

An consequence of Lemmata 5.4.115 and 5.4.116 is that we need to study the sets $A_{\ell,j}^{\sigma}$ only for

$$|j| \le \frac{C|\ell|}{\varepsilon\mathfrak{e}}. \tag{5.4.224}$$

It is essentially what explained in item **b**. Note the here we used the non-degeracy of the constant $m_1$.

**Lemma 5.4.117.** *For any $n \ge 0$ and $j \in \mathcal{S}_{\ell,n}$ one has*

$$|b_j(u_n)|^{lip} \le K\frac{1}{|j|}\left[|m_1|^{lip}|j| + \varepsilon C\right], \tag{5.4.225}$$

*for some $K > 0$.*

*Proof.* One can note that,

$$\begin{aligned} |b_j(\omega_1) - b_j(\omega_2)| &= \left|\frac{b_j^2(\omega_1) - b_j^2(\omega_2)}{b_j(\omega_1) + b_j(\omega_2)}\right| \le \\ &\le |\omega_1 - \omega_1| \left[|m_1|^{lip} + \frac{1}{|j|}(|r_j^j|^{lip} + |r_{-j}^{-j}|^{lip} + |r_j^{-j}|^{lip})\right], \end{aligned} \tag{5.4.226}$$

using that

$$\frac{|(-2|m_1(\omega_1)| + (r_j^j - r_{-j}^{-j})(\omega_1)/j)| + |(-2|m_1(\omega_1)| + (r_j^j - r_{-j}^{-j})(\omega_1)/j)|}{b_j(\omega_1) + b_j(\omega_2)} \le 2, \tag{5.4.227}$$

and that the same bound holds also for $|(r_j^{-j})(\omega_1)|/|j|(b_j(\omega_1) + b_j(\omega_2))$. ■

An immediate consequence of (5.4.225) is that

$$|j||b_j|^{lip} \overset{(5.1.15)}{\le} 4|\ell|\frac{C}{\mathfrak{e}}2K\varepsilon C, \qquad j \in \mathcal{S}_{\ell,n} \cap A_\ell \tag{5.4.228}$$

$$|j||b_j|^{lip} \overset{(5.1.16)}{\leq} K|j|\frac{1}{|j|}\left[\varepsilon|m_1(0)|C\frac{|\ell|}{\varepsilon\mathfrak{e}} + \varepsilon C\right] \leq \widetilde{K}\varepsilon|\ell|, \qquad j \in \mathcal{S}_{\ell,n} \cap (A_\ell)^c \tag{5.4.229}$$

By Lemmata 5.4.116 and 5.4.117 we deduce the following fundamental estimates on the function $\psi$ defined in (5.4.215). First we note that, since there exists $i \in \{1, \dots, d\}$ such that $|\ell_i| \geq |\ell|/2d$, one has

$$|\partial_{\omega_i}\omega \cdot \ell| \geq \frac{|\ell|}{2d}.$$

Hence one has

$$|\psi|^{lip} \geq \left(\frac{|\ell|}{2d} - |j||b_j|^{lip}\right) \overset{(5.4.228)}{\geq} \frac{|\ell|}{4d}, \tag{5.4.230}$$

for $\varepsilon$ small enough for any $j \in \mathcal{S}_{\ell,n}$. The (5.4.230) is fundamental in order to estimate the measure of a single resonant set and this is what we claimed in item **a**. The following Lemma is the part **c**. of the strategy,

**Lemma 5.4.118.** *For $|\ell| \leq N_n$ one has that for any $\varepsilon > 0$ there exists $\bar{n} := \bar{n}(\varepsilon)$ such that if $n \geq \bar{n}(\varepsilon)$ then*

$$(A^\sigma_{\ell,j}(u_n))^c \subseteq (A^\sigma_{\ell,j}(u_{n-1}))^c. \tag{5.4.231}$$

*Proof.* We first have to estimate

$$\begin{aligned} |j||b_j(u_n) - b_j(u_{n-1})| \leq 4\max_{h=\pm j}\{|r_j^{-h}(u_n) - r_j^{-h}(u_{n-1})|\} \\ + 2|m_1(u_n) - m_1(u_{n-1})||j|. \end{aligned} \tag{5.4.232}$$

By Lemma 5.2.103, using the $(\mathbf{S4})_{n+1}$ with $\gamma = \gamma_{n-1}$ and $\gamma - \rho = \gamma_n$, and with $u_1 = u_{n-1}$, $u_2 = u_n$, we have

$$\Lambda^{\gamma_{n-1}}_{n+1}(u_{n-1}) \subseteq \Lambda^{\gamma_n}_{n+1}(u_n), \tag{5.4.233}$$

since, for $\varepsilon\gamma^{-1}$ small enough, and $n \geq \bar{n}(\varepsilon)$ defined as

$$\bar{n}(\varepsilon) := \frac{1}{\log(3/2)}\log\left[\frac{1}{(\kappa - \tau - 3)\log N_0}\log\left(\frac{1}{C\gamma\varepsilon}\right)\right] \tag{5.4.234}$$

$$CN_n^\tau \sup_{\lambda \in G_n} ||u_n - u_{n-1}||_{\mathfrak{s}_0+\mu} \leq \varepsilon(\gamma_{n-1} - \gamma_n) =: \varepsilon\rho = \varepsilon\gamma 2^{-n}. \tag{5.4.235}$$

where $\kappa$ is defined in (4.0.1) with $\nu = 2$, $\mu = \zeta$ defined in (4.4.206) with $\eta = \eta_1 + \beta$, $\mu > \tau$ (see Lemmata 5.4.112, 5.4.113 and (5.2.123), (5.1.11)). We also note that,

$$G_n \cap H_n \overset{(5.4.211),(4.1.23)}{\subseteq} \Lambda^{2\gamma_{n-1}}_\infty(u_{n-1}) \overset{(5.2.192)}{\subseteq} \Lambda^{\gamma_{n-1}}_{n+1}(u_{n-1}) \overset{(C.10)}{\subseteq} \Lambda^{\gamma_n}_{n+1}(u_n). \tag{5.4.236}$$

This means that $\lambda \in H_n \cap G_n \subset \Lambda^{\gamma_{n-1}}_{n+1}(u_{n-1}) \cap \Lambda^{\gamma_n}_{n+1}(u_n)$, and hence, we can apply the $(\mathbf{S3})_\nu$, with $\nu = n+1$, in Lemma 5.2.103 to get for any $h, k = \pm j$,

$$\begin{aligned} |r_h^k(u_n) - r_h^k(u_{n-1})| &\leq |r_h^{n+1,k}(u_n) - r_h^{n+1,k}(u_{n-1})| \\ &\quad + |r_h^k(u_n) - r_h^{n+1,k}(u_n)| + |r_h^k(u_{n-1}) - r_h^{n+1,k}(u_{n-1})| \\ &\overset{(5.1.19a),(5.2.139),(4.0.1)}{\leq} C\varepsilon^2\gamma^{-1}N_n^{-\kappa} + \varepsilon\left(1 + ||u_{n-1}||_{\mathfrak{s}_0+\eta_1+\beta} + ||u_n||_{\mathfrak{s}_0+\eta_1+\beta}\right) N_n^{-\alpha}. \end{aligned}$$

Now, first of all $\kappa > \alpha$ by (4.0.1), (5.2.123), moreover $\eta_1 + \beta < \eta_5$ then by $(\mathbf{S1})_n$, $(\mathbf{S1})_{n-1}$, one has $||u_{n-1}||_{\mathfrak{s}_0+\eta_5} + ||u_n||_{\mathfrak{s}_0+\eta_5} \leq 2$, we obtain

$$|r_h^k(u_n) - r_h^k(u_{n-1})| \overset{(5.4)}{\leq} \varepsilon N_n^{-\alpha}. \tag{5.4.237}$$

Then, by (B.16), (5.1.15) and (C.12) one has that

$$|(\mu_{\sigma,j} - \mu_{\sigma,-j})(u_n) - (\mu_{\sigma,j} - \mu_{\sigma,-j})(u_{n-1})| \leq C\varepsilon|j|N_n^{-\alpha}, \tag{5.4.238}$$

hence for $|\ell| \leq N_n$, and $\lambda \in G_n \cap H_n$, we have

$$|i\omega\cdot\ell + \mu_{\sigma,j}(u_n) - \mu_{\sigma,j}(u_n)| \overset{(C.8)}{\geq} \frac{2\gamma_{n-1}}{\langle\ell\rangle^\tau\langle j\rangle} - C\varepsilon|j|N_n^{-\alpha} \geq \frac{2\gamma_n}{\langle\ell\rangle^\tau\langle j\rangle}, \tag{5.4.239}$$

since $j \in \mathcal{S}_{\ell,n}$, hence $|j| \leq 4|\omega||\ell|/\varepsilon\mathfrak{c}$, and $n$ is such that $N_n^{\tau-\alpha+2} \lessdot \gamma 2^{-n}\varepsilon$. The (C.9) implies the (5.4.231). □

An immediate consequence of Lemma 5.4.118 is the following.

*Proof. Proof of Lemma 5.4.114.* First of all, write

$$H_n \backslash H_{n+1} := \bigcup_{\substack{\sigma\in\boldsymbol{C}, j\in\mathbb{Z} \\ \ell\in\mathbb{Z}^d}} (A^\sigma_{\ell,j}(u_n))^c. \tag{5.4.240}$$

By using Lemma 5.4.118 and equation (5.4.222), we obtain

$$\begin{aligned} &H_n\backslash H_{n+1} \subseteq H_n^{(1)} \cup H_n^{(2)} \cup H_n^{(3)} \cup H_n^{(4)} \\ &H_n^{(1)} := \Big(\bigcup_{\substack{\sigma\in\boldsymbol{C}, \\ j\in\mathcal{S}_\ell\cap A_\ell \\ |\ell|\leq N_n}} (A^\sigma_{\ell,j}(u_n))^c\Big), \quad H_n^{(2)} := \Big(\bigcup_{\substack{\sigma\in\boldsymbol{C}, \\ j\in\mathcal{S}_\ell\cap A_\ell \\ |\ell|> N_n}} (A^\sigma_{\ell,j}(u_n))^c\Big), \\ &H_n^{(3)} := \Big(\bigcup_{\substack{\sigma\in\boldsymbol{C}, \\ j\in\mathcal{S}_\ell\cap (A_\ell)^c \\ |\ell|\leq N_n}} (A^\sigma_{\ell,j}(u_n))^c\Big), \quad H_n^{(4)} := \Big(\bigcup_{\substack{\sigma\in\boldsymbol{C}, \\ j\in\mathcal{S}_\ell\cap (A_\ell)^c \\ |\ell|> N_n}} (A^\sigma_{\ell,j}(u_n))^c\Big). \end{aligned} \tag{5.4.241}$$

One has that the cardinality if the set $\mathcal{S}_{\ell,n} \cap A_\ell$ is less than $4|\ell|C/\mathfrak{e}$. This implies that

$$|H^{(2)}| \leq \sum_{|\ell|>N_n} \frac{4|\ell|C\gamma_n}{\mathfrak{e}\langle j\rangle\langle\ell\rangle^\tau}\frac{4d}{|\ell|} \lessdot C\gamma N_n^{-1}. \tag{5.4.242}$$

Let us estimate the measure of the sets $H^{(i)}$ for $i = 3, 4$. The cardinality of $\mathcal{S}_{\ell,n} \cap (A_\ell)^c$ is less than $K|\ell|/\varepsilon\mathfrak{e}$, hence we have to study the case $j \in \mathcal{S}_{\ell,n} \cap (A_\ell)^c$ more carefully. We introduce the sets

$$B^\sigma_{\ell,j} := \left\{ \omega \in H_{n-1} : |i\omega \cdot \ell + jb_j(u_n)| \geq \frac{\gamma'_n \alpha_n}{\langle\ell\rangle^{\tau_1}} \right\}, \tag{5.4.243}$$

for $\ell \in \mathbb{Z}^d \backslash \{0\}$, $j \in \mathcal{S}_{\ell,n} \cap (A_\ell)^c$, where $\alpha_n := \inf_j |b_j(u_n)|$, $\gamma'_n = (1 + 2^{-n})\gamma'$, $\gamma' \leq \gamma_0$ and $\tau_1 > 0$. We have the following result.

**Lemma 5.4.119.** *Given $\gamma'$ and $\tau_1$, there exist $\gamma$ and $\tau$ such that if $\lambda \in B^c_{\ell,j}$ then $\lambda \in A^\sigma_{\ell,j}$ for $j \in \mathcal{S}_{\ell,n} \cap (A_\ell)^c$.*

*Proof.* First of all

$$j \in \mathcal{S}_\ell, \ \Rightarrow b_j \geq \frac{|\omega \cdot \ell|}{2|j|}, \ \Rightarrow \alpha_n \geq \frac{\gamma_0}{2\langle\ell\rangle^{\tau_0}\langle j\rangle},$$

hence

$$|\omega \cdot \ell + jb_j| \geq \frac{\gamma'_n\alpha_n}{\langle\ell\rangle^{\tau_1}} \geq \frac{\gamma'_n\gamma_0}{\langle j\rangle\langle\ell\rangle^{\tau_1+\tau_0}2} \geq \frac{\gamma_n}{\langle j\rangle\langle\ell\rangle^\tau},$$

if $\gamma'\gamma_0 \geq 2\gamma$ and $\tau \geq \tau_1 + \tau_0$. ☐

By Lemma 5.4.119 follows that

$$|H^{(4)}_n| \leq \sum_{|\ell|>N_n} \sum_{j\in\mathcal{S}_{\ell,n}\cap(A_\ell)^c} |B^\sigma_{\ell j}| \leq \sum_{|\ell|>N_n} \frac{4|\ell|K\gamma'_n\alpha_n}{\varepsilon\mathfrak{e}\langle\ell\rangle^{\tau_1}}\frac{4d}{|\ell|} \lessdot C\gamma' N_n^{-1} \tag{5.4.244}$$

Unfortunately, for the sets $H^{(1)}_n$ and $H^{(3)}_n$ we cannot provide an estimate like (5.4.244); by the summability of the series in $\ell$ we can only conclude

$$|H^{(1)}_n|, |H^{(3)}_n| \leq C\gamma'. \tag{5.4.245}$$

This implies the (5.4.218) for any $n \geq 0$. Moreover by Lemma 5.4.118 we have that if $n \geq \bar{n}(\varepsilon)$ then $H^{(1)}_n = H^{(3)}_n = \emptyset$, hence the (5.4.219) follows by (5.4.242) and (5.4.244). Lemma 5.4.114 implies (5.4.210b) by choosing, for instance, $\gamma := (\gamma')^2 \leq \gamma_0 \leq 1$. ☐

### 5.4.1 Proof of Theorem 1.1.2

Theorem 1.1.2 essentially follows by Propositions 4.0.46 and 5.4.112. The measure estimates performed in the last section guarantee that the "good" sets defined in Prop. 5.0.85 are not empty, but on the contrary have "full" measure. In particular one uses the result of Proposition 5.0.85 in order to prove Lemma refinverseoflham. Indeed one one has diagonalized the linearized operator it is trivial to get estimate (5.3.206). From formula (5.3.206) essentially follows (5.4.210a). Concerning the proof of Proposition 4.0.46, we have omitted since it is the same of Section 4.1.1. The only differences is that in Section 4.1.1 one deals with a functional that is diagonal plus a non linear perturbation. In this case the situation is slightly different. However the next Lemma guarantees that the subspaces $H_n$ in (2.1.8) are preserved by the linear part of our functional $\mathcal{F}$ in (2.1.11),

**Lemma 5.4.120.** *One has that*

$$D_\omega : H_n \to H_n. \tag{5.4.246}$$

*Proof.* Let us consider $u = (u^{(1)}, u^2) \in H_n$, then

$$\begin{aligned} D_\omega u &= D_\omega \sum_{|(\ell,j)|\leq N_n} u_j^{(i)}(\ell) e^{i\ell\cdot\varphi + ijx} \\ &= \begin{pmatrix} \sum_{|(\ell,j)|\leq N_n} (i\omega\cdot\ell) u^1(\ell)_j - [(ij)^2 + m] u_j^{(2)} e^{i\ell\cdot\varphi+ijx} \\ \sum_{|(\ell,j)|\leq N_n} (i\omega\cdot\ell) u^2(\ell)_j + [(ij)^2 + m] u_j^{(1)} e^{i\ell\cdot\varphi+ijx} \end{pmatrix} \in H_n. \end{aligned}$$

□

We fix $\gamma := \varepsilon^a$, $a \in (0,1)$. Then the smallness condition $\varepsilon\gamma^{-1} = \varepsilon^{1-a} < \epsilon_0$ of Proposition 4.0.46 is satisfied. Then we can apply it with $\mu = \zeta$ in (5.3.202) (see Proposition 5.4.112). Hence by (4.0.2) we have that the function $u_\infty$ in $\mathbf{H}^{\mathfrak{s}_0+\zeta}$ is a solution of the perturbed NLS with frequency $\omega$. Moreover, one has

$$|\Lambda \backslash \mathcal{G}_\infty| \overset{(5.4.210b)}{\to} 0, \tag{5.4.247}$$

as $\varepsilon$ tends to zero. To complete the proof of the theorem, it remains to prove the linear stability of the solution. Since the eigenvalues $\mu_{\sigma,j}$ are purely imaginary, we know that the Sobolev norm of the solution $v(t)$ of (4.1.27) is constant in time. We just need to show that the Sobolev norm of $h(t) = W_2 v(t)$, solution of $\mathcal{L}h = 0$ does not grow on time. Again to do this one can follow the same strategy used in Section 4.5.1 (or in [31]). In particular one uses the results of Lemma 5.1.95 in Section 5.1 and estimates (5.2.111) in Proposition 5.2.96 in order to get the estimates

$$||h(t)||_{H^s_x} \leq K||h(0)||_{H^s_x}, \tag{5.4.248a}$$

$$||h(0)||_{H^s_x} - \varepsilon^b K||h(0)||_{H^{s+1}_x} \leq ||h(t)||_{H^s_x} \leq ||h(0)||_{H^s_x} + \varepsilon^b K||h(0)||_{H^{s+1}_x}, \tag{5.4.248b}$$

for $b \in (0,1)$. Clearly the (5.4.248) imply the linear stability of the solution, so we concluded the proof of Theorem 1.1.2.

# 6. Autonomous NLS

In this last Chapter we give the proof of Theorem 1.2.5 on the analytic autonomous NLS. We see that one can obtain Theorem 1.2.6 just by following words by words the proof given for the analytic case. This is due to the fact that we never exploit analytic properties of functions defined on the complex strip $\mathbb{T}_a$. As we explained in the Introduction we just use the property of the functions to be Sobolev on the boundary. In other words we never use Cauchy estimates and we never control the (polynomial) loss of regularity, due to the small divisors, by using the exponential decay of the Fourier coefficients of analytic functions. We exploit the "tame" properties as essentially done for forced cases.

We start by studying the preliminary steps of Birkhoff normal form, in order to introduce a frequency-amplitude modulation of the frequency. In such a way we transform the system (1.2.16) into the form (15) in order to apply Theorem 3.2.39. In Section 6.5 we give a more explicit formulation of the sets of "good" parameters given by Theorem (3.2.39), and hence, in Section 6.6 we concludes the proof of Theorem 1.2.5 by giving measure estimates of such sets.

## 6.1 Weak Birkhoff Normal Form

Let fix some notation. Given a finite set of distinct numbers $\{j_1,\dots,j_N\} = E^+ \subset \mathbb{N}$ we define $E := \{\pm j_1,\dots,\pm j_N\} \subset \mathbb{Z}$. This decomposes naturally $h^{a,p}$ into two orthogonal subspaces $u = (\{u_j\}_{j\in E}, \{u_j\}_{j\notin E})$. We write

$$u(x) = \Pi_E u + \Pi_E^{\perp} u., \quad h^{a,p} = \Pi_E h^{a,p} \oplus \Pi_E^{\perp} h^{a,p}. \tag{6.1.1}$$

We choose $S^+ = \{\mathtt{v}_1,\dots,\mathtt{v}_d\} \subset \mathbb{N}$ as above and denote $v = \Pi_S u$ the tangential variables and $z = \Pi_S^{\perp} u$ the normal ones. As notation we will also indicate with $R(v^q z^r)$ a homogeneous polynomial

$$R(v^q z^r) := M[\overbrace{v^+,\dots,v^+,v^-,\dots v^-}^{q-times},\overbrace{z^+,\dots,z^+,z^-,\dots,z^-}^{r-times}],$$

with $M$ a $q,r$–multilinear operator in the variables $v^{\pm}, z^{\pm}$.

**Definition 6.1.121.** *For any natural $k$ consider a $2k$-uple $\vec{\jmath} = (j_1, \dots, j_{2k}) \in \mathbb{Z}^{2k}$. We say that $\vec{\jmath}$ is a* $k$-**resonance** *if*

$$\sum_{i=1}^{2k} (-1)^i j_i = 0\,, \quad \sum_{i=1}^{2k} (-1)^i j_i^2 = 0.$$

*We say that a $k$-resonance is* **trivial** *if $j_i = j_{i+1}$ up to a permutation of the $\{j_{2l}\}_{l=1}^k$.*
*We say that a $2k$-uple is* **non-resonant**, *$\vec{\jmath} \in \mathtt{N}$ if*

$$\sum_{i=1}^{2k} (-1)^i j_i = 0\,, \quad \sum_{i=1}^{2k} (-1)^i j_i^2 \neq 0$$

**Remark 6.1.122.** *Note that all 2-resonances are trivial. Indeed if $j_1 - j_2 + j_3 - j = 0$ then $j_1^2 - j_2^2 + j_3^2 - j^2 = 2(j_1 - j_2)(j_2 - j_3) = 0$.*

**Lemma 6.1.123.** *For $S$ generic one has that that there are no non-trivial 3-resonances with at least five points in $S$*

*Proof.* We just need to exhibit the polynomial which gives such genericity condition. □

Given $\vec{\jmath} = (j_1, \dots, j_n) \in \mathbb{Z}^n$ for some $n$ we say that

$$\vec{\jmath} \in \mathcal{A}_l\,, \quad \text{if at most } l \text{ of the } j_i \text{ are not in } S\,.$$

Note that for each fixed $n$ the set $(j_1, \dots, j_n) \in \mathcal{A}_1$ is finite dimensional. For a finite dimensional subspace $E := \mathrm{span}\{e^{ijx} : |j| < C\}, C > 0$ we denote $\Pi_E$ its $L^2$ projector.

**Proposition 6.1.124** (Weak Birkhoff Normal Form). *There exists an analytic change of variables of the form*

$$\mathbf{u}_+ = \Phi(\mathbf{u}) = \mathbf{u} + \Psi(\mathbf{u}), \tag{6.1.2}$$

*where $\Psi$ is a* finite *rank map. The map $\Phi(\mathbf{u})$ is defined for all* [1] *$\mathbf{u} \in \mathbf{h}^{a,p}$ such that $\|\mathbf{u}\|_{a,\mathfrak{p}_1} \le \epsilon_0$, and satisfies the bounds:*

$$\|\Psi(\mathbf{u})\|_{a,p} \le C_p \epsilon_0^2 \|\Pi_E \mathbf{u}\|_{a,p}\,, \quad \|d_{\mathbf{u}}\Psi(\mathbf{u})[h]\|_{a,p} \le C_p \left(\epsilon_0^2 \|\Pi_E h\|_{a,p} + \epsilon_0 \|\Pi_E u\|_{a,p} \|\Pi_E h\|_{a,\mathfrak{p}_1}\right) \tag{6.1.3}$$

*for all $\mathbf{u} : \|\mathbf{u}\|_{a,\mathfrak{p}_1} \le \epsilon_0$. Actually $\Psi$ is* tame modulus *in the sense of [6], namely it respects interpolation bounds also for the higher order derivatives.*

[1] note that $\epsilon_0$ is fixed in terms of $\mathtt{r}_0$ and $a, \mathfrak{p}_1$.

*Finally $\Phi$ preserves $\mathbf{h}^{a,p}_{\rm odd}\cap\mathcal{U}$ and the new vector field $\Phi_*\chi := \Upsilon$ restricted to $\mathcal{U}$ is reversible and has the form*

$$\begin{aligned}\overline{\Upsilon^-} &= \Upsilon^+,\\ \Pi_S\Upsilon^+ &:= -\mathrm{i}(v^+_{xx} + A(\mathbf{u}) + B^+_1(\mathbf{u})),\\ \Pi_S^\perp\Upsilon^+ &:= -\mathrm{i}(z^+_{xx} + Q(\mathbf{u}) + B^+_2(\mathbf{u})),\end{aligned} \tag{6.1.4}$$

*where*

$$\begin{aligned}B^\sigma_1(\mathbf{u}) &:= \sum_{i,j,k\in S} C_{ikj}u^+_i u^-_i u^+_k u^-_k u^\sigma_j \partial_{u^\sigma_j} + \sum_{j\in S}\sum_{k\in S^c,} \chi_{kkjj}u^+_k u^-_k u^\sigma_j\partial_{u^\sigma_j}, + \sum_{q=0}^3 R(v^q z^{5-q}) + \Pi_S h^{(>5)}(u)),\\ B^\sigma_2(\mathbf{u}) &:= \sum_{q=0}^4 R(v^q z^{5-q}) + \Pi^\perp_S h^{(>5)}(u)),\end{aligned} \tag{6.1.5}$$

*$h^{(>5)}$ collects all terms with degree greater than 5, and for*[2] $\mathbf{u}\in\mathcal{U}$

$$\begin{aligned}A(u) &:= \sum_{j\in S} C^j_j |u_j|^2 u_j\partial_{u_j} + \sum_{j\in S}(\sum_{\substack{i\in S\\ k\neq j}} C^k_j |u_k|^2)u_j\partial_{u_j}, \quad Q(u) = \sum_{j\in S^c}\sum_{\substack{j_1-j_2+j_3-j=0,\\ (j_1,j_2,j_3,j)\notin\mathcal{A}_1}} \chi_{j_1j_2j_3j}u_{j_1}\bar u_{j_2}u_{j_3}\partial_{u_j},\\ C^j_j &:= \chi_{jjjj}, \qquad C^k_j := \chi_{kkjj} + \chi_{jkkj},\\ \chi_{j_1j_2j_3j} &:= \mathtt{a}_1 - \mathtt{a}_2 j_3^2 + \mathtt{a}_3 j_1 j_2 - \mathtt{b}_2 j_2^2 - \mathtt{b}_3 j_1 j_2 - \mathtt{a}_4 j_1 j_2 j_3^2 + \mathtt{b}_4 j_1 j_2^2 j_3 - \mathtt{a}_6 j_1^2 j_2^2 j_3^2\end{aligned} \tag{6.1.6}$$

*Proof.* Now consider the equation (1.2.16). As notation for a vector field $F$ we denote by $F_{j_1\dots j_{2k+1}j}$ the coefficient of the monomial $u^+_{j_1}u^-_{j_2}\dots u^+_{j_{2k+1}}\partial_{u^+_j}$. We divide

$$\chi(\mathbf{u}) = \mathcal{N} + \chi_3(\mathbf{u}) + \chi_5(\mathbf{u}) + \chi_{>5}(\mathbf{u})$$

a direct computation gives

$$\mathcal{N} = \mathrm{i}\sum_j j^2 u^+_j\partial_{u^+_j} - \mathrm{i}\sum_j j^2 u^-_j\partial_{u^-_j}$$

while

$$\chi_3(\mathbf{u}) = -\mathrm{i}\sum_{j\in\mathbb{Z}}\sum_{\substack{j_1-j_2+j_3=j\\ j_i\in\mathbb{Z}, i=1,2,3}} \chi_{j_1j_2j_3j}u^+_{j_1}u^-_{j_2}u^+_{j_3}\partial_{u^+_j} + \mathrm{i}\sum_{j\in\mathbb{Z}}\sum_{\substack{j_1-j_2+j_3=j\\ j_i\in\mathbb{Z}, i=1,2,3}} \bar\chi_{j_1j_2j_3j}u^-_{j_1}u^+_{j_2}u^-_{j_3}\partial_{u^-_j}$$

with $\chi_{j_1j_2j_3j}\in\mathbb{R}$ defined in (6.1.6), comes from $f$. The other terms collect respectively the part of degree 5 and $>5$ of $g$.

We want to eliminate from $\chi$ all the monomials $u^+_{j_1}u^-_{j_2}u^+_{j_3}\partial_{u^+_j}$ or $u^-_{j_1}u^+_{j_2}u^-_{j_3}\partial_{u^-_j}$ such that the list $(j_1,j_2,j_3,j)\in\mathcal{A}_1\cap\mathtt{N}$. Note that this is a finite set of monomials.

[2] extending polynomials outside $\mathcal{U}$ is trivial just apply $u\to u^+$ and $\bar u\to u^-$.

We define the transformation $\Phi^{(1)}$ as the time$-1$ flow map generated by the vector field

$$
\begin{aligned}
F_3(\mathbf{u}) = &- \sum_{\sigma=+1,-1} \sigma \sum_{j\in\mathbb{Z}} \sum_{\substack{j_1-j_2+j_3=j \\ (j_1,j_2,j_3,j)\in\mathcal{A}_1\cap\mathtt{N}}} \frac{\chi_{j_1j_2j_3j}}{(j_1^2-j_2^2+j_3^2-j^2)} u^\sigma_{j_1}u^{-\sigma}_{j_2}u^\sigma_{j_3}\partial_{u^\sigma_j} \\
&- \sum_{\sigma=+1,-1} \sigma \sum_{j\in S} \sum_{\substack{j_1-j_2+j_3=j \\ (j_1,j_2,j_3,j)\in(\mathcal{A}_1)^c\cap\mathtt{N}}} \frac{\chi_{j_1j_2j_3j}}{(j_1^2-j_2^2+j_3^2-j^2)} u^\sigma_{j_1}u^{-\sigma}_{j_2}u^\sigma_{j_3}\partial_{u^\sigma_j}
\end{aligned}
\tag{6.1.7}
$$

By construction the transformation $\Phi^{(1)}$ has finite rank. Moreover one has

$$
\chi_{j_1j_2j_3j} \in \mathbb{R}, \quad \chi_{(-j_1)(-j_2)(-j_3)(-j)} = \chi_{j_1j_2j_3j},
$$

hence the vector field $F_3$ in (6.1.7) is reversibility preserving.

By construction the push-forward of the vector field $\Phi^{(1)}_*\chi := \mathcal{Y}$ has the form $\mathcal{Y} = \mathcal{N}+\mathcal{Y}_3+\mathcal{Y}_5+\mathcal{Y}_{>5}$, where $\mathcal{Y}_3$ contains only monomials $u^\sigma_{j_1}u^{-\sigma}_{j_2}u^\sigma_{j_3}\partial_{u^\sigma_j}$ such that $(j_1,j_2,j_3,j)$ is either a trivial resonance or is not in $\mathcal{A}_1$ or $j\in S$ and at least two among $j_1,j_2,j_3$ are in $S^c$ (see the second summand in $B^\sigma_1$ in (6.1.5)) . The trivial resonances in $\mathcal{A}_1$ give $A(u)$, all the other terms either contribute to $B_1$ or to $Q$. More explicitly In this way the system $\dot{\mathbf{u}} = \mathcal{Y}(\mathbf{u})$ possesses an invariant subspace $H_S$ and its dynamics is integrable and, as we will see, non-isochronous.

In order to enter a perturbative regime we need to cancel further term from the vector field. In particular we look for a transformation $\Phi^{(2)}$ such that the field $\Upsilon := \Phi^{(2)}_*\mathcal{Y}$ does not contain monomials $u^\sigma_{j_1}u^{-\sigma}_{j_2}u^\sigma_{j_3}u^{-\sigma}_{j_4}u^\sigma_{j_5}\partial_{u^\sigma_j}$ such that the list $(j_1,j_2,j_3,j_4,j_5,j)\in\mathcal{A}_1\cap\mathtt{N}$, as in degree three this is a finite set of monomials. $\Phi^{(2)}$ is the time 1 flow of the vector field $F_5$ of the form

$$
\begin{aligned}
F_5(\mathbf{u}) = &-\sum_{j\in\mathbb{Z}} \sum_{\substack{j_1-j_2+j_3-j_4+j_5=j \\ (j_1,j_2,j_3,j_4,j_5,j)\in\mathcal{A}_1\cap\mathtt{N}}} \frac{\mathcal{Y}_{j_1j_2j_3j_4j_5j}}{(j_1^2-j_2^2+j_3^2-j_4^2+j_5^2-j^2)} u^+_{j_1}u^-_{j_2}u^+_{j_3}u^-_{j_4}u^+_{j_5}\partial_{u^+_j}+ \\
&+\sum_{j\in\mathbb{Z}} \sum_{\substack{j_1-j_2+j_3-j_4+j_5=j \\ (j_1,j_2,j_3,j_4,j_5,j)\in\mathcal{A}_1\cap\mathtt{N}}} \frac{\overline{\mathcal{Y}}_{j_1j_2j_3j_4j_5j}}{(j_1^2-j_2^2+j_3^2-j_4^2+j_5^2-j^2)} u^-_{j_1}u^+_{j_2}u^-_{j_3}u^+_{j_4}u^-_{j_5}\partial_{u^-_j}.
\end{aligned}
\tag{6.1.8}
$$

Again by construction $\Phi^{(2)}$ has finite rank. Moreover since $\mathcal{Y}$ is reversible then $F_5$ is reversibility preserving. Finally $\Upsilon := \Phi^{(2)}_*\mathcal{Y} = \mathcal{N}+\mathcal{Y}_3+\Upsilon_5+\Upsilon_{>5}$ contains only monomials such that $(j_1,j_2,j_3,j_4,j_5,j)$ is either a resonance or is not in $\mathcal{A}_1$. By Lemma 6.1.123 all the resonances in $\mathcal{A}_1$ are trivial and hence contribute to the first summand in $B_1$. Now we perform the last step in order to cancel out from $\mathcal{Y}_3$

For the tame estimates (6.1.3) we refer to [6]. $\square$

**Remark 6.1.125.** *Note that, by construction the change of variables written on functions $H^p(\mathbb{T}_a)$ is*

$$
\mathbf{u}(x) \rightsquigarrow \mathbf{q}(x) = \mathbf{u}(x) + \sum_{j\in E} \Psi_j(\Pi_E\mathbf{u})e^{ijx}
$$

*hence setting* $\mathbf{u} = \Phi^{-1}(\mathbf{q}) = \mathbf{q}(x) + \sum_{j\in E} \widetilde{\Psi}_j(\Pi_E \mathbf{q})e^{ijx}$ *one has*

$$\Upsilon(\mathbf{q}) := d\Phi(\Phi^{-1}(\mathbf{q}))\chi(\Phi^{-1}(\mathbf{q})) = -\mathrm{i}E\left[\mathbf{u}_{xx} + \begin{pmatrix}\mathtt{f}^+(\mathbf{u},\mathbf{u}_x,\mathbf{u}_{xx})\\ \mathtt{f}^-(\mathbf{u},\mathbf{u}_x,\mathbf{u}_{xx})\end{pmatrix}\right] + d\Psi(\mathbf{u})\chi(\mathbf{u})\,.$$

*Then ( see Lemma 7.1 of [4] for a detailed proof)*

$$d_{\mathbf{q}}\Upsilon(\mathbf{q}) = -\mathrm{i}E\left[\partial_{xx} + d_{\mathbf{u}}\mathcal{F}(\mathbf{u})\right] + \mathcal{R}(\mathbf{q})$$

*where the first term is described in* (2.1.19)*, while for some fixed* $N$

$$\mathcal{R}(\mathbf{q})[h] = \sum_{m=1}^{N}\Big(h(x), a^{(m)}(\mathbf{q};x)\Big)_{L^2} b^{(m)}(\mathbf{q};x)\,.$$

*Here* $a^{(m)}, b^{(m)}$ *are functions in* $h^{a,p}$ *depending on* $\mathbf{q}$ *and such that, for any* $m$*, one of the* $a^{(m)}, b^{(m)}$ *is equal to* $e^{\mathrm{i}jx}$ *for some* $j \in E$*. In other words* $\mathcal{R}$ *is a linear operator sum of two terms, one maps* $\Pi_E h^{a,p}$ *into* $h^{a,p}$ *and the other maps* $h^{a,p}$ *into* $\Pi_E h^{a,p}$*.*

## 6.2 Action-angles variables

In the previous paragraph we have worked in the Fourier basis and we have shown that the change of variables preserves $\mathbf{h}^{a,p}_{\rm odd}$. Now we restrict our vector field to $\mathbf{h}^{a,p}_{\rm odd}$ where it is natural to use the sin basis indexed by $\mathbb{N}$.

We want to switch the tangential variables to polar coordinates. We set

$$\begin{aligned} 2u^\sigma_{\pm \mathtt{v}_i} &:= \pm\sigma\frac{1}{2\mathrm{i}}\sqrt{\xi_i + y_i}e^{\sigma \mathrm{i}\mathtt{e}_i}, \qquad i = 1,\dots,d \\ u^\sigma_j &:= \sigma\frac{\mathrm{sign}(j)}{2\mathrm{i}}z^\sigma_{|j|}, \qquad\qquad j\in S^c, \end{aligned} \tag{6.2.9}$$

this is a well defined , analytic change of variables for $\xi_i > 0$, $|y_i| < \xi_i$.

For $\varepsilon$ small we consider $\xi \in \varepsilon^2\Lambda = \varepsilon^2[1/2,3/2]^d$ and the domain

$$D_{a_0,p+\nu}(s_0,r_0) := \Big\{\theta\in\mathbb{T}^d_{s_0},\ |y|\le r_0^2, \|w\|_{a_0,\mathtt{p}_1}\le r_0\Big\} \subseteq \mathbb{T}^d_{s_0}\times\mathbb{C}^d\times\Pi_S^\perp \mathbf{h}^{a_0,p+\nu}_{\rm odd}. \tag{6.2.10}$$

One can check that there exist constant $c_1$ and $c_2$ such that, if

$$r_0 < c_1\varepsilon < \varepsilon/2, \qquad \sqrt{2d}c_2\kappa^{\mathtt{p}_1}e^{s+a\kappa}\varepsilon < \epsilon_0, \quad \kappa := \max(|j_i|), \tag{6.2.11}$$

then one has $\Phi^{(\xi)} : D_{a,p+\nu}(s_0,r_0) \to B_{\epsilon_0}$, where the vector field $\Upsilon$ is well defined. We assume that our parameters $\varepsilon, r_0, s_0$ satisfy (6.2.11) sto that we can apply $\Phi_\xi$ to our vector field. In the new variables one has $\mathbf{u}(x) = \mathbf{v}(\theta, y; x) + w(x)$ with

$$\mathbf{v} = (v^+, v^-)\,, \quad v^{\pm} := \sum_{i=1}^{d} \sqrt{\xi_i + y_i} e^{\pm \mathrm{i}\theta_i} \sin(\mathtt{v}_i x), \quad w = (z^+, z^-)\,, z^\sigma = \sum_{j \in S^c} z_j^\sigma \sin(jx), \tag{6.2.12}$$

and

$$F := \Phi_*^{(\xi)} \Upsilon = F^{(\theta)}(\theta, y, w)\partial_\theta + F^{(y)}(\theta, y, w)\partial_y + F^{(w)}(\theta, y, w)\partial_w, \tag{6.2.13}$$

reads

$$\begin{aligned} F^{(\theta_k)}(\theta, y, w) &= \frac{\Upsilon^+_{\mathtt{v}_k}}{2\mathrm{i}v^+_{\mathtt{v}_k}} - \frac{\Upsilon^-_{\mathtt{v}_k}}{2\mathrm{i}v^-_{\mathtt{v}_k}} = {\mathtt{v}_k}^2 - (\mathcal{M}(\xi + y))_k - \frac{e^{-\mathrm{i}\theta_k}(B_1^+)_{\mathtt{v}_k} + e^{\mathrm{i}\theta_k}(B_1^-)_{\mathtt{v}_k}}{\sqrt{\xi_k + y_k}} \qquad k = 1, \dots, d, \\ F^{(y_k)}(\theta, y, w) &= 4v^-_{\mathtt{v}_k}\Upsilon^+_{\mathtt{v}_k} + 4v^+_{\mathtt{v}_k}\Upsilon^-_{\mathtt{v}_k} = 2\mathrm{i}\sqrt{\xi_k + y_k}(e^{-\mathrm{i}\theta_k}(B_1^+)_{\mathtt{v}_k} - e^{\mathrm{i}\theta_k}(B_1^-)_{\mathtt{v}_k}), \qquad k = 1, \dots, d \\ F^{(w)}(\theta, y, w)) &= 2\Pi_S^\perp \Upsilon \end{aligned} \tag{6.2.14}$$

where $\mathcal{M}$ is the twist matrix

$$\mathcal{M}_{kh} = \frac{1}{4}\big(C^{\mathtt{v}_h}_{\mathtt{v}_k} + C^{-\mathtt{v}_h}_{\mathtt{v}_k}\big) \quad \text{for} \quad k, h = 1, \dots, d\,, \quad \mathtt{v}_k, \mathtt{v}_h \in S^+. \tag{6.2.15}$$

We define $\widetilde{\omega} \in \mathbb{R}^d$ the vector of *unperturbed* frequencies as

$$\widetilde{\omega}_j = \widetilde{\omega}_j(\xi) := \omega_j^{-1} + \omega_j^{(0)}(\xi), \qquad \omega_j^{(-1)} := j^2, \quad \omega_j^{(0)}(\xi) := -(\mathcal{M}\xi)_j, \; j \in S^+. \tag{6.2.16}$$

In the new variables $\mathcal{U}$ becomes

$$\mathcal{U} := \{(\theta, y, w) \in \mathbb{C}^{2d} \times \Pi_S^\perp \mathbf{h}^{0,0}_{\mathrm{odd}} : \quad \theta \in \mathbb{T}^d\,, \quad y \in \mathbb{R}^d\,, \quad \bar{z}^- = z^+\}$$

which is clearly preserved by the dynamics of $F$. We now are in the setting of (3.2.43) with $p' = p + \nu$, $\nu = 2$ and $\ell_{a,q} \rightsquigarrow \Pi_S^\perp h^{a,q}$.

**Remark 6.2.126.** *Our identification of $\ell_{a,p}$ with Sobolev functions on $\mathbb{T}_a$ implies that a map $\theta \to f(\theta) \in \ell_{a,p}$ with $\|f\|_{s,a,p} < \infty$ is identified with a function $f(\theta, x) \in H^p(\mathbb{T}^d_s \times \mathbb{T}_a)$ and $\|f(\theta, x)\|_{H^p(\mathbb{T}^d_s \times \mathbb{T}_a)} \sim \|f\|_{s,a,p}$. See Lemma A.170 in the Appendix A.1 for the equivalence.*

**Properties of $F$** We describe here some fundamental properties of the vector field $F$.

- *Tameness* $F$ is tame according to Definition (3.2.26). This properties follows by reasoning as in Lemma 4.1.50 and by the properties of the map $\Psi$ in Proposition 6.1.124.

- *Reversibility*: $F$ restricted to $\mathcal{U}$ is reversible w.r.t. the involution

$$S : (\theta_j, y_j, z_j, \bar{z}_j) \mapsto (-\theta_j, y_j, \bar{z}_j, z_j), \quad j \in \mathbb{N}, \quad S^2 = \mathbb{1}. \tag{6.2.17}$$

hence we have that

$$\begin{aligned} &F^{(\theta)}(-\theta, y, -\bar{w}) = F^{(\theta)}(\theta, y, w), \qquad F^{(y)}(-\theta, y, -\bar{w}) = -F^{(y)}(\theta, y, w), \\ &F^{(w)}(-\theta, y, -\bar{w}) = \overline{F^{(w)}(\theta, y, w)}. \end{aligned} \tag{6.2.18}$$

In particular the component $F^{(\theta)}$ is even in the variables $\theta$ while $F^{(y)}$ is odd in $\theta$.

- *Constants of motion*: The vector field $F$ preserves the mass $M = \sum_i y_i + \sum_j |z_j|^2$, this means that

$$F = \sum_{l,h,\alpha,\beta,x} F^{(x)}_{l,h,\alpha,\beta} e^{\mathrm{i}\theta\cdot l} y^h z^\alpha \bar{z}^\beta \partial_x \,, \quad x = \theta, y, z_j^\sigma$$

is such that

$$\sum \ell_i + \sum (\alpha_j - \beta_j) = \begin{cases} 0, & x = \theta, y \\ \sigma, & x = z_j^\sigma \end{cases}$$

- *Pseudo-differential structure*: Setting $u = (\theta, y, w)$ one has that

$$d_w F(u) = \mathscr{P}(u) + \mathscr{R}(u) \tag{6.2.19}$$

where

$$\begin{aligned} \mathscr{P}(u) = -\mathrm{i}E\, \Pi_S^\perp &\left[ \begin{pmatrix} 1 + a_2(u;x) & b_2(u;x) \\ \bar{b}_2(x) & 1 + \bar{a}_2(u;x) \end{pmatrix} \partial_{xx} + \begin{pmatrix} a_1(u;x) & b_1(u;x) \\ \bar{b}_1(x) & \bar{a}_1(u;x) \end{pmatrix} \partial_x \right] \Pi_S^\perp \\ - \mathrm{i}E\, \Pi_S^\perp &\left[ \begin{pmatrix} a_0(u;x) & b_0(u;x) \\ \bar{b}_0(u;x) & \bar{a}_0(u;x) \end{pmatrix} \right] \Pi_S^\perp \end{aligned}$$

while

$$\begin{aligned} \mathscr{R}(u)[h] = A(u) \begin{pmatrix} h^{(\theta)} \\ h^{(y)} \end{pmatrix} &+ \sum_{l=1}^{d} f_l(u) h^{(\theta_l)} + f_{l+d}(u) h^{(y_l)} + \mathsf{e}^{(\theta_l)} g_l \cdot h^{(w)} + \mathsf{e}^{(y_l)} g_{l+d} \cdot h^{(w)} + \\ + \sum_{m=1}^{N} (c_m \cdot h^{(w)}) d_m & \end{aligned} \tag{6.2.20}$$

where $A(u) \in \mathrm{Mat}(\mathbb{C}^{2d})$ while $c_l, d_l, f_l, g_l \in \ell_{a,p}$ for each $l$. In other words we have

$$\mathscr{R}(u)[h] = \left(\begin{array}{c|c} \text{A} & \begin{matrix} —f_1^T— \\ \vdots \\ —f_{2d}^T— \end{matrix} \\ \hline \begin{matrix} | & & | \\ g_1 & \cdots & g_{2d} \\ | & & | \end{matrix} & 0 \end{array}\right) + \mathscr{K}(u)[h] := (\mathscr{A}(u) + \mathscr{K}(u))[h] \tag{6.2.21}$$

where

$$\mathscr{K}(u)[h] = \sum_{m=1}^{N} (c_m \cdot h^{(w)}) d_m \tag{6.2.22}$$

is the linear operator on $\ell_{a,p} \to \ell_{a,p}$

**Definition 6.2.127 (Pseudo-differential vector fields).** *We call finite rank operators the linear operators of the form $\mathscr{R}$ and Schrödinger pseudo-differential operators the ones of the form $\mathscr{P}$.*

*We say that a tame vector field $F = N_0 + G$ is of Schrödinger pseudo-differential type if its differential in a neighborhood of zero has the form (6.2.19) for some $\mathscr{R}$,$\mathscr{P}$ for a fixed $N \geq 0$ and where $a_i(u), b_i(u)$ for $i = 0, 1, 2$ and $c_l, d_l, f_l, g_l$ for any $l$ are tame functions form $D_{a,p}(s, r)$ to $\ell_{a,p}$ with tame constant controlled by the tameness constant $C_{\vec{v},p}(G)$.*

**Remark 6.2.128.** *Finite rank operators are bounded linear operators on $\mathbb{C}^{2d} \times \ell_{a,p}$ and form a bilateral ideal in the algebra of bounded linear operators. More precisely given $R$ of finite rank and a linear operator $L : \mathbb{C}^{2d} \times \ell_{a,p} \to \mathbb{C}^{2d} \times \ell_{a,q}$ then $RL$ is of finite rank and maps $\mathbb{C}^{2d} \times \ell_{a,p} \to \mathbb{C}^{2d} \times \ell_{a,p}$ while $LR$ is of finite rank and maps $\mathbb{C}^{2d} \times \ell_{a,p} \to \mathbb{C}^{2d} \times \ell_{a,q}$.*

## 6.3 Inizialization

In this Section we consider the field $F$ defined in (6.2.14) and we prove that satisfies hypotheses of the abstract theorem 3.2.39. In particular we need to fix the subspace of vector field $\mathcal{E}$ on which we will work and the constants given in (3.2.69). We also perform directly a first step in which we find the first map of the sequence $\mathcal{T}_n$. We follows the notations of the abstract theorem 3.2.39 and we set $F_0 := F$ defined on the domain in (6.2.10) $D_{a_0,p+\nu}(s_0, r_0)$ where the parameters are given by formula (6.2.11). Moreover by 6.2.16 we set

$$N_0 := (\omega^{-1} + \omega^{(0)}(\xi)) \cdot \partial_\theta + \Omega^{-1} w \partial_\theta = \widetilde{\omega}(\xi) \cdot \partial_\theta + \Omega^{-1} w \partial_\theta, \tag{6.3.23}$$

and $(\Omega^{-1})^{\sigma}_{\sigma} = i\sigma \text{diag} \mathrm{j}^2$, $(\Omega^{-1})^{-\sigma}_{\sigma} = 0$. Hence by equation (6.2.14) we have $F_0 = N_0 + G_0$ and we define

$$
\begin{aligned}
\Pi_{\mathcal{N}} F &= \omega(\xi,\theta)\partial_\theta + \Omega(\theta,\xi) w \partial_w, \\
\omega(\xi,\theta) &= \omega^{(-1)} + \omega^{(0)}(\xi) + G_0^{(\theta,0)}(\xi,\theta), \\
\Omega(\theta,\xi) &= \Omega^{(-1)} + \Omega^{(0)}(\theta,\xi) = d_w F_0^{(w)}(\theta,0,0)[\cdot] = \Omega^{-1} + d_w G_0^{(w)}(\theta,0,0)[\cdot]
\end{aligned}
\tag{6.3.24}
$$

Let us study in particular the linear operator $\Omega$ on $\ell_S^{a_0,p}$. In order to do this it is more simple to identify $\ell_S^{a_0,p}$ with $\Pi_S^{\perp} \mathbf{h}_{\text{odd}}^{a_0,p} := \Pi_S^{\perp} h_{\text{odd}}^{a_0,p} \times \Pi_S^{\perp} h_{\text{odd}}^{a_0,p}$ (recall definition in (6.1.1)). We have that $\Omega^{-1} := -iE\partial_{xx} : \Pi_S^{\perp} \mathbf{h}_{\text{odd}}^{a_0,p} \to \Pi_S^{\perp} \mathbf{h}_{\text{odd}}^{a_0,p}$ where $E := \text{diag}\{1,-1\}$. Moreover one has that $\Omega^{(0)} = ((\Omega^{(0)})^{\sigma'}_{\sigma})_{\sigma,\sigma'=\pm1}$ can be seen as a 2 times 2 matrix whose components are operator on $\Pi_S^{\perp} h_{\text{odd}}^{a_0,p}$. In particular

$$
\begin{aligned}
&\Omega^{(0)}(\theta,\xi) = \Pi_S^{\perp}\left(-iE\left[T_1(\theta,\xi) + T_2(\theta,\xi)\right]\right)\Pi_S^{\perp} + \mathscr{K}_0, \\
&\qquad T_1 = \left[\begin{pmatrix} a_2^{(0)}(\theta,\xi) & b_2^{(0)}(\theta,\xi) \\ \bar{b}_2^{(0)}(\theta,\xi) & \bar{a}_2^{(0)}(\theta,\xi)\end{pmatrix}\partial_{xx} + \begin{pmatrix} a_1^{(0)}(\theta,\xi) & b_1^{(0)}(\theta,\xi) \\ \bar{b}_1^{(0)}(\theta,\xi) & \bar{a}_1^{(0)}(\theta,\xi)\end{pmatrix}\partial_x + \begin{pmatrix} a_0^{(0)}(\theta,\xi) & b_0^{(0)}(\theta,\xi) \\ \bar{b}_0^{(0)}(\theta,\xi) & \bar{a}_0^{(0)}(\theta,\xi)\end{pmatrix}\right] \\
&a_2^{(0)}(\theta,\xi) := \mathtt{a}_2|v|^2 + \mathtt{a}_4|v_x|^2 + 2\mathtt{a}_6|v_{xx}|^2, \quad b_2^{(0)}(\theta,\xi) := \mathtt{b}_2|v|^2 + \mathtt{b}_4(v_x)^2 + \mathtt{a}_6 v_{xx}^2 \\
&a_1^{(0)}(\theta,\xi) := \mathtt{a}_3\bar{v}_x v + \mathtt{a}_4\bar{v}_x v_{xx} + 2\mathtt{b}_3\bar{v}v_x + 2\mathtt{b}_4 v_x\bar{v}_{xx}, \quad b_1^{(0)}(\theta,\xi) := \mathtt{a}_3 v v_x + \mathtt{a}_4 v_x v_{xx} \\
&a_0^{(0)}(\theta,\xi) := 2\mathtt{a}_1|v|^2 + \mathtt{a}_2\bar{v}v_{xx} + \mathtt{a}_3|v_x|^2 + \mathtt{b}_2\bar{v}\bar{v}_{xx}, \quad b_0^{(0)}(\theta,x) := \mathtt{a}_1 v^2 + \mathtt{a}_2 v v_{xx} + \mathtt{b}_2 v\bar{v}_{xx} + \mathtt{b}_3(v_x)^2,
\end{aligned}
\tag{6.3.25}
$$

where $T_2(\theta,\xi)$ has the same form of $T_1$ but collects all terms that are at least of degree 5 in $v$ and $\mathscr{K}_0$ has the form (6.2.22). Clearly in formula (6.3.25) the function $v(\theta,y)$ is evaluated at $y=0$. We can see that $T_1$ comes from the cubic term $Q(u)$ defined in (6.1.6) while $T_2$ comes from the linearization of the term $B_2$. For convenience we split $\Pi_{\mathcal{N}} F_0$ in several terms using equation (6.3.25). In particular we split the term $T_1$ in (6.3.25). This is done because each term will have different sizes and will be treated in a different way. More explicitly we define

$$
F_0 := N_0 + N^{(1)} + N^{(2)} + H_0, \tag{6.3.26}
$$

where $N^{(1)}, N^{(2)} \in \mathcal{N}$, are

$$
\begin{aligned}
N^{(1)} &:= \Pi_S^{\perp}\left(-iE\begin{pmatrix} a_2^{(0)}(\theta,\xi) & b_2^{(0)}(\theta,\xi) \\ \bar{b}_2^{(0)}(\theta,\xi) & \bar{a}_2^{(0)}(\theta,\xi)\end{pmatrix}\partial_{xx}\right)\Pi_S^{\perp} w\cdot\partial_w, \\
N^{(2)} &:= \Omega^{(0)}(\theta)w - N^{(1)}, \qquad H_0 := G_0 - N^{(1)} - N^{(2)}
\end{aligned}
\tag{6.3.27}
$$

where the coefficients $a_2^{(0)}, b_2^{(0)}$ and the field $\Omega^{(0)}$ are defined in (6.3.25) . Since the vector $\widetilde{\omega}(\xi) \in \{\widetilde{\omega}(\xi) : \xi \in \varepsilon^2\Lambda\}$ (see (6.2.10)) we have that $\widetilde{\omega}$ is $\xi-$close to the integer vector $\omega^{-1}$ in (6.3.24), hence we fix the size of $\gamma_0$ by requiring that $\widetilde{\omega}(\xi)$ satisfies

$$
|\widetilde{\omega}\cdot l| \geq \frac{\gamma_0}{\langle l\rangle^{\tau}}, \;\; \forall\, l \in \mathbb{Z}^d,\, \tau \geq d+1,\, \gamma_0 = \boldsymbol{c}\xi, \tag{6.3.28}
$$

with the constant $\boldsymbol{c}$ small enough.

We define

$$\mathcal{E} := \big\{F \in \mathcal{V}_{\vec{v}_0,p} \ : \ F \ \text{satisfies} \ (6.2.18)\big\} \tag{6.3.29}$$

where $\vec{v}_0 := (\gamma_0, \mathcal{O}_0, s_0, a_0)$ with $\mathcal{O}_0 := \varepsilon^2\Lambda = \varepsilon^2[1/2, 3/2]^d$.

**Lemma 6.3.129.** *The vector field* $F_0$ *given by* (6.2.14) *satisfies the hypotheses of Theorem* (3.2.39).

*Proof.* By definition each term in (6.3.26) is tame, now we want to estimate their tameness constant in order determine $\Theta, \mathfrak{A}_0$ in (3.2.69). By usign the Definition 3.2.26 and the explicit form in (6.3.25) of the coefficients one has that

$$C_{\vec{v}_0,\mathfrak{p}_2}(N^{(1)}) \leq \mathfrak{M}_0\xi, \quad C_{\vec{v}_0,\mathfrak{p}_2}(N^{(2)}) \leq \mathfrak{M}_0\xi, \quad C_{\vec{v}_0,\mathfrak{p}_2}(H_0) \leq \mathfrak{M}_0\xi^{\frac{3}{2}}, \tag{6.3.30}$$

where the constant $\mathfrak{M}_0$ depends on the constants $\mathtt{a}_i$, $\mathtt{b}_j$ for $i = 1, 2, 3, 4, 6$, $j = 2, 3, 4$ in (6.3.25) and on $\mathfrak{p}_2$. This scaling justifies the splitting (6.3.26). Indeed we have separated the terms $N^{(1)}, N^{(2)}$ that are not "perturbative" with respect to the size of the small divisors $\gamma_0 \approx \xi$.

We check explicitly the (6.3.30). Recall that $F_0$ in (6.2.14) is defined in terms of equations (6.1.4), (6.1.5) and (6.1.6). We start to study the terms $N^{(1)}$ and $N^{(2)}$ that are terms that contribute to $F^{(w)}$ that comes form $Q$ in (6.1.6). For instance we can bound using the interpolation properties of the norm $\|\cdot\|_{s,a,p}$

$$\gamma_0^{-1}\|a_2|v|^2\partial_{xx}z\|_{\vec{v}_0,\mathfrak{p}_2} \leq \frac{1}{\boldsymbol{c}\xi}\left(C(\mathfrak{p}_2)\xi + C(\mathfrak{p}_0)\xi\|z\|_{\vec{v}_0,p+\nu}\right), \tag{6.3.31}$$

where we used that $\|z\|_{a,\mathfrak{p}_1} \leq r_0$. Hence one can check Definition (3.2.26) with a constant $C_{\vec{v}_0,p}(a_2|v|^2 z_{xx}) \leq \mathfrak{A}_0 = \mathfrak{A}_0(\mathfrak{p}_2, \boldsymbol{c})$. All the other terms in (6.3.25) can be estimated in the same way. Indeed all those terms are quadratic on $v$ and linear in $z$. Recall that the norm $\|\cdot\|_{\vec{v}_0,p}$ is a weighted norm and on the $w-$ component the weight is $r_0$ (see (3.2.34)). This determines the constant $\mathfrak{A}_0$ by setting

$$\gamma_0^{-1}C_{\vec{v}_0,\mathfrak{p}_2}(N^{(1)}), \gamma_0^{-1}C_{\vec{v}_0,\mathfrak{p}_2}(N^{(2)}) \leq \frac{\mathfrak{M}_0}{\boldsymbol{c}} = \mathfrak{A}_0(\mathfrak{p}_2, \boldsymbol{c}), \tag{6.3.32}$$

so we can fix also a large number $K_0 > \mathfrak{A}_0$. Now let us estimate the several components of the field $H_0$ starting from $\Pi_{\mathcal{N}}H_0$. Consider $\Pi_{\mathcal{N}}H_0^{(w)}$. All these terms comes form $B_2^+$ in (6.1.4) so that the linear term in $z$ is at least of degree 4 in $v$. Hence one has

$$\gamma_0^{-1}C_{\vec{v}_0,\mathfrak{p}_2}(\Pi_{\mathcal{N}}(H_0^{(w)})) \leq \mathfrak{A}_0\xi. \tag{6.3.33}$$

We now note that $\Pi_{\mathcal{N}}H_0^{(\theta)} := H_0^{(\theta,0)}(\theta)$ is of the form (6.2.14) where $B_1$ is defined in (6.1.5) where the term independent on $z$ of degree minimum has degree 5 in $v$. Hence

$$\gamma_0^{-1}C_{\vec{v}_0,\mathfrak{p}_2}(\Pi_{\mathcal{N}}(H_0^{(\theta)})) \leq \mathfrak{A}_0\xi. \tag{6.3.34}$$

Let us study $\Pi_{\mathcal{A}} H_0$. Recall that by Lemma B.177 we can directly to estimate $|\Pi_{\mathcal{A}} H_0|_{\vec{v}_0,p}$ defined in (B.7) in order to estimate the tameness constant of the field. We have by equation (6.1.5) that $\Pi_{\mathcal{A}} H_0^{(w)} = \Pi_{\mathcal{A}} \Pi_S^{\perp} h^{(>5)}(u)$, hence the term $H_0^{(w,0)}(\theta)$ is of degree at least 6 in $v$. Hence

$$\gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(\Pi_{\mathcal{A}} H_0^{(w)}) \le \mathfrak{A}_0 \xi^2 r_0^{-1}. \tag{6.3.35}$$

Now by (6.2.14) and (6.1.5) one has

$$\gamma_0^{-1} |\Pi_{\mathcal{A}} H_0^{(y)}|_{\vec{v}_0,\mathfrak{p}_2} \le \mathfrak{A}_0 \xi^2 \sqrt{\xi} r_0^{-2} + \mathfrak{A}_0 \xi + \mathfrak{A}_0 \xi^2 r_0^{-1} \tag{6.3.36}$$

To get bound (6.3.36) we used an important property coming from reversibility condition and from the fact that the only term of degree 5 in $v$ in $B_1$ are integrable as one can see in (6.1.5). This two fact implies that such term of degree 5 cancel out and hence do not contribute to the $y$ component of $F_0$. Collecting the bounds (6.3.35) and (6.3.36) we have

$$\gamma_0^{-1} |\Pi_{\mathcal{A}} H_0|_{\vec{v}_0,\mathfrak{p}_2} \le \mathfrak{A}_0 \xi^b, \quad b < \frac{3}{2}, \tag{6.3.37}$$

provided

$$\xi^{\frac{5-2b}{4}} \le r_0 \le c_1 \sqrt{\xi}. \tag{6.3.38}$$

The second inequality comes form (6.2.11). Now we study $\Pi_{\mathcal{R}} H_0$. The component$\Pi_{\mathcal{R}} H_0^{(\theta)}$ is at least linear in the variables $y, w$ while the terms in $\Pi_{\mathcal{R}} H_0^{(y)}$ comes from the last two summand in (6.1.5) (this follows by the definition of $R^{(1)}$ that collects the terms coming form the second summand and the fact that the integrable terms of order 5 are zero). Following the same reasoning of the previous bounds we get

$$\gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(\Pi_{\mathcal{R}} H_0^{(\theta)}) \le \mathfrak{A}_0 \xi^{\frac{3}{2}}, \qquad \gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(\Pi_{\mathcal{R}} H_0^{(y)}) \le \mathfrak{A}_0 \sqrt{\xi}, \qquad \gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(\Pi_{\mathcal{R}} H_0^{(w)}) \le \mathfrak{A}_0 \xi \tag{6.3.39}$$

without requiring any additional hypotheses on $r_0$. Again in the second bound we use that the integrable terms in $B_1$ in (6.1.5) cancel out. Indeed on terms of the form $R(vz^2)$ in $B_1$ one cannot prove a bound like (6.3.39). Now we fix for convenience $b = 1/2$, and for $\xi$ small enough we set

$$\delta := \gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(\Pi_{\mathcal{A}} H_0) := \xi^{\frac{1}{4}}, \quad \Theta := \gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(H_0) \le \mathfrak{A}_0 \delta, \tag{6.3.40}$$

Hence we have fixed the constants in (3.2.69). Moreover by requiring that $\xi$, (or that $\varepsilon$ in (6.3.29)) is small enough, then equation (3.2.70) is satisfied. □

All the terms that are not "small" with $\delta$ are in $N^{(1)}$ and $N^{(2)}$. In order to use Theorem 3.2.39 we need to identify the sequence of maps $\mathcal{T}_n$ with $n \ge 1$. We first present some abstract results on reversible pseudo-differential vector field of Schrödinger type (see Definition 6.2.127). Then in the last Section we will prove the measure estimates that concludes the proof of Theorem 1.2.5

## 6.4 Pseudo-differential Vector Fields

In this Section we study how pseudo-differential operator changes under special changes of variables. First we prove some lemmata of conjugation of non linear vector fields. Then, in Section 6.4.2, we analyze properties of the linearized operator of Pseudo-differential vector field. In particular we study how to invert it. In the forced cases of Chapters 4 and 5 we shown as to study the asymptotics of the eigenvalues of the linearized operator. In the autonomous case we perform a similar analysis. the important difference is the following. Here we follow the iterative scheme of Theorem 3.2.39 while for forced cases we proved the Nash-Moser theorem 3.1.18. Theorem 3.2.39 is in some sense very close to a classical KAM scheme. Indeed at each step on choose the convenient coordinates in which the approximate solution is trivial. In the Nash-Moser scheme in Theorem 3.1.18 one never changes coordinates. On the contrary here we must study how the entire non linear system changes under the transformations of coordinates we use. In particular in this Section we study the transformation we use to define the $\mathcal{T}_n$ introduced in Theorem 3.2.39. We refer to Lemma 6.5.153 for more details.

### 6.4.1 Regularization

In the following we consider the decay norm $|\cdot|_{s,a,p}$ in we have introduced in (4.3.85) in order to deal with linear operators on $\Pi_S^{\perp}\mathbf{h}_{\rm odd}^{a_0,p+\nu}$. We have the following important property.

**Lemma 6.4.130 (Decay along lines).** *Let $M=(M_i^{i'})_{i,i'\in S^c\times\mathbb{Z}^d}$ be a linear operator on $\Pi_S^{\perp}\mathbf{h}_{\rm odd}^{a_0,p+\nu}$. Then one has*

$$|M|_{s,a,p}\le C\max_{i\in S^c\times\mathbb{Z}^d}\|M_{\{i\}}\|_{s,a,p+d+1}. \tag{6.4.41}$$

The decay norm satisfies the following classical interpolation estimates proved in Lemmata 4.3.62, 4.3.63. Moreover to obtain the stability result on the solutions we will strongly use this property.

**Lemma 6.4.131.** *If $A$ is a Töpliz in time matrix as in (4.3.98), and $\mathfrak{p}_0:=(d+2)/2$, then one has*

$$|A(\theta)|_{s,p}\le C(\mathfrak{p}_0)|A|_{s,a,p+\mathfrak{p}_0}:=C_{\mathfrak{p}_0}\left(\sup_{\sigma,\sigma'}\sum_{(h,l\in)\mathbb{Z}_+\times\mathbb{Z}^d}\langle h,l\rangle^{2p}e^{2|h|a}e^{2|\ell|s}\sup_{k-k'=h}|A_{\sigma,k}^{\sigma',k'}(l)|^2\right)^{\frac{1}{2}},\quad\forall\,\varphi\in\mathbb{T}^d. \tag{6.4.42}$$

*Proof.* It is sufficient to follow word by word the proof of Lemma 6.7 of [31] by substituting the matrix $A$ with $\widetilde{A}$ defined as $|\widetilde{A}_{\sigma,k}^{\sigma',k'}(l)|=|A_{\sigma,k}^{\sigma',k'}(l)|e^{2|l|s}e^{2|k-k'|a}$. □

**Remark 6.4.132.** *The class of linear operators with* decay *is strictly stronger than just being bounded. In particular contains an important class operator, the so called multiplication operators. See Remark 4.34 in Section 4 of [31].*

**Remark 6.4.133.** *The* $(s,p)-$*decay norm is defined on linear operators* $A_k^{k'}$ *with indeces* $k, k' \in \mathbb{Z}_+ \times \mathbb{Z}^d$*. The linearized operator in the normal directions, in our case, is supported on* $k, k' \in S^c \times \mathbb{Z}^d$*. We say that a linear operator* $A$ *on* $\Pi_S^{\perp} \mathbf{h}_{\mathrm{odd}}^{a_0,p+\nu}$ *has the* decay *is it is the restriction to* $S^c$ *of some operator* $B$ *on* $\mathbf{h}_{\mathrm{odd}}^{a_0,p+\nu}$ *with* $(s,p)-$ *decay norm finite.*

We now want to introduce a "decay" norm on generic linear operator on $\mathbb{C}^{2d} \times \Pi_S^{\perp} \mathbf{h}_{\mathrm{odd}}^{a_0,p+\nu}$.

**Definition 6.4.134.** *Consider a linear operator of the form*

$$M := M_1 + M_2 : \mathbb{C}^{2d} \times \Pi_S^{\perp} \mathbf{h}_{\mathrm{odd}}^{a_0,p+\nu} \to \mathbb{C}^{2d} \times \Pi_S^{\perp} \mathbf{h}_{\mathrm{odd}}^{a_0,p+\nu}, \tag{6.4.43}$$

*where* $M_1$ *is a finite rank operator of the form* (6.2.20) *and is a linear operator* $M_2 : \Pi_S^{\perp} \mathbf{h}_{\mathrm{odd}}^{a_0,p+\nu} \to \Pi_S^{\perp} \mathbf{h}_{\mathrm{odd}}^{a_0,p+\nu}$*. We define the decay norm of* $M$ *as*

$$|M|_{s,a,p}^{\mathrm{dec}} := \max_{i,j=1,\dots,2d} (\|A_j^k\|_{s,a,p}) + \max_{i=1,\dots,2d} (\|c_i\|_{s,a,p}, \|f_i\|_{s,a,p}) + |K|_{s,a,p} + |M_2|_{s,a,p} \tag{6.4.44}$$

*where the coefficients* $A_j^k, c_i, f_i, a_i, b_i$ *has the same meaning as in* (6.2.20) *and* $K$ *is the term in* (6.2.22)*. The Lipschitz norm* $|\cdot|_{\gamma,\mathcal{O},s,a,p}$ *is defined as done in* (4.3.86)*.*

Definition 6.4.134 guarantees that a finite rank operator of the form (6.2.20) has a finite "decay" norm if the coefficients $f_l, g_l$ are in $\ell_{a,p}$ and $c_l, d_l \in \ell_{a,p+d+1}$ for any $l$. In other words a finite rank operator has lines in $\ell_{a,p+d+1}$ then has decay. Hence for any operator $F = N_0 + G$ of the Schrödinger type one has that

$$|\mathscr{R}|_{\vec{v},p} \le C_{\vec{v},p+d+1}(G) \tag{6.4.45}$$

In this Section we study the properties of Schrödinger type pseudo-differential equation. In particular we study how this structure changes under the following class of linear changes of variables:

- *Torus embedding.* the transformation $\Phi := \mathbb{1} + f$ with $f \in \mathcal{B}$. Note that this is a finite rank operator (see Definition 6.2.127),
- *Diffeomorphism of the torus 1.* the transformation $\Phi := \Pi_S^{\perp} \mathcal{T}_\alpha \Pi_S^{\perp}$ where $\mathcal{T}_\alpha u := u(\theta, x + \alpha(\theta, x))$,
- *Diffeomorphism of the torus 2.* The transformation $\mathcal{T}_\beta : \theta \to \theta + \beta(\theta)$
- *Multiplication operator.* the transformation $\Phi := \Pi_S^{\perp} M(\theta, x) \Pi_S^{\perp}$ where $M(\theta, x)$ is a multiplication operator on the space $H^p(\mathbb{T}_a^{d+1})$.

**Lemma 6.4.135 (Torus embedding).** *Fix $\rho > 0$. Consider $F : \mathcal{O} \times D_{a,p+\nu}(s+\rho s_0, r+\rho r_0) \to V_{a,p}$ a vector field of pseudo-differential type, then for all* torus embedding $\Phi = \mathbb{1} + f$ *as in Definition 3.2.24 and with $f \in \hat{\mathcal{B}}$ one has that $\Phi_* F$ is tame and pseudo-differential. In particular $d(\Phi_* F)(u)[h] = \mathscr{P}_+ + \mathscr{R}_+$ has the form* (6.2.19) *where $\mathscr{P}_+$ has coefficients $\widetilde{a}_i, \widetilde{b}_i$ tame with constant*

$$C_{\vec{v},p}(\widetilde{a}_i) \leq C_{\vec{v},p}(a_i) + C_{\vec{v},p}(f) C_{\vec{v},\mathfrak{p}_1}(a_i) \qquad \vec{v} := (\gamma, \mathcal{O}, s+\rho s_0, a). \tag{6.4.46}$$

*The coefficients $\widetilde{c}_i, \widetilde{d}_i$ for $i = 1, \dots, N$ of $\mathscr{K}_+$ satisfies the same bound* (6.4.46).

*Proof.* Since the transformation $\Phi$ is tame (see Lemma B.177) by Lemma 3.2.31 one has that the push-forward is tame with tameness constants given by (3.2.49). We study in particular the structure of the linearized operator in $u = (\theta, y, w)$. One has

$$d(\Phi_* F)(u)[h] = dF(\Phi^{-1}(u))[h] + R[h], \tag{6.4.47}$$

where $R$ is a finite rank operator as in (6.2.19). Moreover the pseudo-differential part as the same form of $dF(u)$ but with coefficients $\widetilde{a}_i(x) = a_i(\Phi^{-1}(u); x)$ (same for $b_i$). The bounds (6.4.46) follow by the tameness of the coefficients $a_i$. The range of the finite rank operator does not increase. Indeed $(\Phi_* F)(u) = F \circ \Phi^{-1} + d_u f(\Phi^{-1})[F \circ \Phi^{-1}]$. The second term does not contribute to the $w-$component. If we linearize the first term we have

$$d_u(F \circ \Phi^{-1})(\theta, 0, 0)[h] = d_u F(\widetilde{u})[h].$$

This happens because the map $\Phi^{-1}$ contains translations, and in other words, to linearize in zero the field $F \circ \Phi^{-1}$ is nothing but to linearize in a point near zero the field $F$. The property 6.2.127 requires that the linearized operatorhas the form (6.2.19) in a whole neighbourhood of the origin, hence the rank is again $N$. The estimates of the new coefficients $\widetilde{c}_i, \widetilde{d}_i$ follows. ☐

**Lemma 6.4.136 (Diffemorphism of the torus 1).** *Fix $\rho > 0$. Consider $F : \mathcal{O} \times D_{a,p+\nu}(s+\rho s_0, r+\rho r_0) \to V_{a,p}$ a vector field of pseudo-differential type, and a function $\alpha : \mathbb{T}^d_s \times \mathbb{T}_a \to \mathbb{C}$ such that $\|\alpha\|_{\vec{v},\mathfrak{p}_1} \leq \delta$, for some $\delta > 0$ small. Then setting $\mathcal{T}_\alpha u(\theta, x + \alpha(\theta, x)) := u(\theta, x_+)$ and defining the map*

$$\Phi : \theta_+ = \theta, \;\; y_+ = y, \;\; w_+ = \Pi_S^\perp \mathcal{T}_\alpha w, \tag{6.4.48}$$

*one has that for some $\rho$ small $\Phi_* F : D_{a-2\rho a_0, p+\nu}(s+\rho s_0, r+\rho r_0) \to V_{a-2\rho a_0, p}$ is tame of Schrödinger type and $\delta \ll \rho$. Moreover $d(\Phi_* F)(u)[h] = \mathscr{P}_+ + \mathscr{R}_+$ has the form* (6.2.19) *where $\mathscr{P}_+$ has coefficients $\widetilde{a}_i, \widetilde{b}_i$ given by*

$$\begin{aligned}
&1 + \widetilde{a}_2 = \mathcal{T}_\alpha^{-1}[(1+a_2)(1+\alpha_x)^2], \qquad \widetilde{b}_2 = \mathcal{T}_\alpha^{-1}[(b_2)(1+\alpha_x)^2], \\
&\widetilde{a}_1 = \mathcal{T}_\alpha^{-1}[(1+a_2)\alpha_{xx}] - i\mathcal{T}_\alpha^{-1}[\omega \cdot \partial_\theta \alpha] + \mathcal{T}_\alpha^{-1}[a_1(1+\alpha_x)], \qquad \widetilde{b}_1 = \mathcal{T}_\alpha^{-1}[b_1(1+\alpha_x)], \\
&\widetilde{a}_0 = \mathcal{T}_\alpha^{-1}[a_0], \qquad \widetilde{b}_0 = \mathcal{T}_\alpha^{-1}[b_0],
\end{aligned} \tag{6.4.49}$$

*where* $\mathcal{T}_\alpha^{-1}u = u(\theta, y+\widetilde{\alpha}(\theta,y)$ *is the inverse of the diffeomorphism* $x \to x+\alpha(\theta,x)$*, and* $\mathscr{R}_+ = \mathcal{K}_+ + \mathscr{K}_+$ *(see* (6.4.62)*) where* $\mathscr{K}_+$ *with rank* $N$ *with coefficients* $\widetilde{c}_i, \widetilde{d}_i$ *such that*

$$C_{\vec{v},p}(\widetilde{c}_i) \leq C_{\vec{v},p}(c_i) + \|\alpha\|_{\vec{v},p+\mathfrak{p}_0}(\alpha) C_{\vec{v},\mathfrak{p}_1}(c_i). \tag{6.4.50}$$

*The coefficients* $\widetilde{a}_i = \widetilde{a}_i(\theta, y, w_+; y)$*, with* $w_+ = \mathcal{T}_\alpha w$ *and* $y = x + \alpha(\theta, x)$ *are tame with constant*

$$\begin{aligned} C_{\vec{v}_2,p}(\widetilde{a}_1) &\leq \|\alpha\|_{\vec{v},p+2} + C_{\vec{v},p}(a_2) + \|\alpha\|_{\vec{v}_1,p} C_{\vec{v},\mathfrak{p}_1}(a_2) + C_{\vec{v},p}(a_1) + \|\alpha\|_{\vec{v}_1,p} C_{\vec{v},\mathfrak{p}_1}(a_1) \\ C_{\vec{v}_2,p}(\widetilde{b}_1) &\leq C_{\vec{v},p}(b_1) + \|\alpha\|_{\vec{v}_1,p+2} C_{\vec{v},\mathfrak{p}_1}(b_1) \end{aligned} \tag{6.4.51}$$

*where* $\vec{v} := (\gamma, \mathcal{O}, s+\rho s_0, a)$*,* $\vec{v}_1 := (\gamma, \mathcal{O}, s+\rho s_0, a-\rho a_0)$ *and* $\vec{v}_2 := (\gamma, \mathcal{O}, s+\rho s_0, a-2\rho a_0)$*. Same for* $\widetilde{a}_0, \widetilde{b}_0$*.*

*Proof.* The vector field $\Phi_* F$ is clearly tame, indeed in the new variables the system reads

$$\begin{cases} \dot{\theta}_+ = F^{(\theta)}(\theta_+, y_+, (\Pi_S^\perp \mathcal{T}_\alpha)^{-1} w_+) \\ \dot{y}_+ = F^{(y)}(\theta_+, y_+, (\Pi_S^\perp \mathcal{T}_\alpha)^{-1} w_+) \\ \dot{w}_+ = [F^{(\theta)}(\theta_+, y_+, (\Pi_S^\perp \mathcal{T}_\alpha)^{-1} w_+) \cdot \partial_\theta \alpha] \partial_{x_+} w_+ + \Pi_S^\perp \mathcal{T}_\alpha F^{(w)}(\theta_+, y_+, (\Pi_S^\perp \mathcal{T}_\alpha)^{-1} w_+) \end{cases} \tag{6.4.52}$$

while equation (6.4.49) follows by a direct computation. Bounds (6.4.51) follows by Lemma A.171. Indeed we already note that $\|\cdot\|_{\vec{v},p}$ is equivalent to $\|\cdot\|_{H^p(\mathbb{T}_s^d \times \mathbb{T}_a)}$ on the functions $u(\theta, x)$. Hence Lemma A.171 applies on the diffeomorphism of the toroidal domain $\mathbb{T}_s^d \times \mathbb{T}_a$ $\theta \to \theta$ and $x \to x + \alpha(\theta, x)$. Let us check (6.4.50). We define $\mathscr{R}_+ = \mathscr{A}_+ + \mathscr{K}_+$ with same splitting used in (6.2.21). Again by using Lemma A.171 one gets the bound for the term in $\mathscr{A}_+$. Since $F$ is of Schrödinger type, then the coefficients $c_l, d_l$ in $\mathscr{K}$ are in $\ell_{a,p+d+1}$ and their norms are controlled by $C_{\vec{v},p}(F)$. Now one can see that $\mathscr{K}_+$ is of the form (6.2.22) with the same $N$ of $\mathscr{K}$. Indeed

$$\mathscr{K}_+(w_+) := (\mathcal{T}_\alpha \mathscr{K} \mathcal{T}_\alpha^{-1}) w_+ = \mathcal{T}_\alpha \sum_{m=1}^N (c_m(x), \mathcal{T}_\alpha^{-1} w_+)_{L^2} d_m(x). \tag{6.4.53}$$

Hence formula (6.4.50) holds.

□

**Lemma 6.4.137** (**Diffeomorphism of the torus 2**)**.** *Consider* $F$ *as in Lemma 6.4.136 and the transformation* $\mathcal{T}_\beta : \theta \to \theta + \beta(\theta)$ *with* $\|\beta\|_{\vec{v},\mathfrak{p}_1} \leq \delta$*, for some* $\delta > 0$ *small. Then setting* $\mathcal{T}_\beta u(\theta + \beta(\theta), x) := u(\theta_+, x)$ *and defining the map*

$$\Phi : \theta_+ = \theta + \beta(\theta), \ \ y_+ = y, \ \ w_+ = w, \tag{6.4.54}$$

*one has that for some $\rho$ small $\Phi_* F : D_{a,p+\nu}(s-2\rho s_0, r+\rho r_0) \to V_{a,p}$ is tame of Schrödinger type and $\delta \ll \rho$. Moreover $d(\Phi_* F)(u)[h] := \mathscr{P}_+ + \mathscr{R}_+$ has the form* (6.2.19) *where $\mathscr{P}$ has coefficients $\widetilde{a}_i, \widetilde{b}_i$ such that*

$$C_{\vec{v}_2,p}(\widetilde{a}_i) \le C_{\vec{v},p}(a_i) + \|\beta\|_{\vec{v},p+\mathfrak{p}_0} C_{\vec{v},\mathfrak{p}_1}(a_i), \tag{6.4.55}$$

*where $\vec{v} := (\gamma, \mathcal{O}, s+\rho s_0, a)$, $\vec{v}_1 := (\gamma, \mathcal{O}, s, a)$ and $\vec{v}_2 := (\gamma, \mathcal{O}, s-2\rho s_0, a)$. Same for $\widetilde{b}_i$. Moreover $\mathscr{R}_+ = \mathcal{K}_+ + \mathscr{K}_+$ (see* (6.4.62)*) where $\mathscr{K}_+$ with rank $N$ and coefficients that satisfy the bound* (6.4.50) *with $\alpha \rightsquigarrow \beta$.*

*Proof.* One can reason as in Lemma 6.4.136 and use Lemma A.171. □

**Lemma 6.4.138 (Multiplication operator).** *Consider $F$ as in Lemma 6.4.136 and the transformation $\Phi := \Pi_S^\perp M(\theta,x) \Pi_S^\perp$ where $M(\theta,x) = \mathbb{1} + A(\theta,x) : H^p(\mathbb{T}^d_s \times \mathbb{T}_a) \times H^p(\mathbb{T}^d_s \times \mathbb{T}_a) \to H^p(\mathbb{T}^d_s \times \mathbb{T}_a) \times H^p(\mathbb{T}^d_s \times \mathbb{T}_a)$ with $\|A\|_{\vec{v},\mathfrak{p}_1}$ small. One has that $\Phi_* F$ is tame with tameness constant given by* (3.2.49) *where $C_{\vec{v},p} \rightsquigarrow \|A\|_{\vec{v},p}$. Moreover, writing*

$$d_w F(u) = \Pi_S^\perp \left[(-iE + L_2)\partial_{xx} + L_1 \partial_x + L_0\right] \Pi_S^\perp + \mathscr{R}(u)$$

*where*

$$L_i(\theta, y, w; x) = -iE \begin{pmatrix} a_i(u;x) & b_i(u;x) \\ \bar{b}_i(x) & \bar{a}_i(u;x) \end{pmatrix}, \quad i = 0,1,2,$$

*the linearized operator $d_w(\Phi_* F)(u)$ has the form*

$$\begin{aligned} &\Pi_S^\perp \left( M^{-1}(-iE + L_2) M \partial_{xx} + \left[2M^{-1}(-iE+L_2)\partial_x M + M^{-1} L_1 M\right] \partial_x \right. \\ &\left. + \Big[M^{-1}(F^{(\theta)} \cdot \partial_\theta M) + M^{-1}(-iE+L_2)\partial_{xx} M + M^{-1} L_1 \partial_x M + M^{-1} L_0 M\Big]\right) \Pi_S^\perp + \mathscr{R}_+ \end{aligned} \tag{6.4.56}$$

*hence each coefficient $\widetilde{L}_i$ for $i = 0,1,2$ defined by equation* (6.4.56) *is tame with the same tameness constant of $\Phi_* F$. Moreover one has that $\mathscr{R}_+ = \mathcal{K}_+ + \mathscr{K}_+$ (see* (6.4.62)*) where $\mathscr{K}_+$ with rank $N$ of $\mathscr{R}$ and coefficients such that*

$$C_{\vec{v}_2,p}(\widetilde{c}_i) \le C_{\vec{v},p}(c_i) + \|A\|_{\vec{v},p+\mathfrak{p}_0} C_{\vec{v},\mathfrak{p}_1}(c_i), \quad i = 1,\dots,N, \tag{6.4.57}$$

*same for $\widetilde{d}_i$.*

*Proof.* First of all we note that $\Phi$ is a tame map. It is sufficient to apply the definition and use the fact that $\|\cdot\|_{\vec{v},p}$ is equivalent to $\|\cdot\|_{H^p(\mathbb{T}^d_b \times \mathbb{T}_a)}$ on the functions $u(\theta,x)$. Hence also the push-forward is tame. Equation (6.4.56) follows again by an explicit calculation, and the tame bounds on the coefficients follow by the tameness of the transformation and of the coefficients $L_i$. The bounds (6.4.57) follows by interpolation properties of the decay norm. □

**Remark 6.4.139 (Loss of regularity).** *Note that transformations in Lemma 6.4.135 do not loose analyticity. On the contrary the diffeormorphisms of the torus in Lemmata 6.4.136 and 6.4.137 loose regularity in the analytic case. In the differentiable case such transformations are diffeomorphism of the real torus. The loss of regularity is controlled by the the low norm of the functions $\alpha$ and $\beta$.*

Lemmata 6.4.135, 6.4.136, 6.4.137 and 6.4.138 guarantees that the structure of pseudo-differential operator of the linearized in a neighbourhood of $u = (\theta, 0, 0)$ persists under the change of variables we need to apply. We also have decomposed the linearized operator in homogeneous decreasing symbols of order two, one and zero. By Lemmata above we note that also such decomposition is stable, and we are able to control the tameness constant of each symbol. This not *a priori* obvious.

The following Lemma is the key result of this Section. We give the result on a special class of vector field. Consider tame, pseudo-differential and reversible vector fields $F : \mathcal{O} \times D_{a,p+\nu}(s, r) \to V_{a,p}$ with $F = N_0 + G$ (see Section 6.6) We assume that $F$ has the form

$$F := (1 + h)\Big(\hat{N}_0 + N^{(1)} + N^{(2)} + H\Big) \tag{6.4.58}$$

for some $h : \mathcal{O} \times \mathbb{T}^d_s \to \mathbb{C}$ with

$$\gamma^{-1}\|h\|_{\vec{v},p} := \delta^{(3)}_p, \tag{6.4.59}$$

and where $\hat{N}_0 = \omega \cdot \partial_\theta + \widetilde{\Omega}^{-1} w \cdot \partial_w$ with $\widetilde{\Omega}^{-1} = c\Omega^{-1}$, $\omega \in \mathbb{R}^d$ is diophantine and

$$|\omega - \widetilde{\omega}|_\gamma \le o(\xi), \quad |c - 1|_\gamma \le O(\xi), \tag{6.4.60}$$

$$N^{(1)} = \left(\Pi_S^\perp \left(-\mathrm{i}E \begin{pmatrix} a_2(\theta, x) & b_2(\theta, x) \\ \bar{b}_2(\theta, x) & a_2(\theta, x) \end{pmatrix}\right) \partial_{xx} \Pi_S^\perp\right) w \cdot \partial_w \tag{6.4.61}$$

$$N^{(2)} = -\mathrm{i}E\Pi_S^\perp \left[\begin{pmatrix} a_1(\theta, x) & b_1(\theta, x) \\ \bar{b}_1(\theta, x) & \bar{a}_1(\theta, x) \end{pmatrix} \partial_x + \begin{pmatrix} a_0(\theta, x) & b_0(\theta, x) \\ \bar{b}_0(\theta, x) & \bar{a}_0(\theta, x) \end{pmatrix}\right] \Pi_S^\perp + \mathscr{K}, \tag{6.4.62}$$

and $\mathscr{K}$ of the form (6.2.22). Moreover we assume that

$$d_w H^{(w)}(u)[\cdot] : \Pi_S^\perp \mathbf{h}^{a,p+1}_{\mathrm{odd}} \to \Pi_S^\perp \mathbf{h}^{a,p}_{\mathrm{odd}}. \tag{6.4.63}$$

and

$$C_{\vec{v},p}(H) \le C_{\vec{v},p}(\Pi_{\mathcal{N}^\perp} G), \quad C_{\vec{v},p}(N^{(2)}) \le C_{\vec{v},p}(G). \tag{6.4.64}$$

Note that by the reversibility condition one has that on $\mathcal{U}$ the function $h(\theta)$ is real and *even* in $\theta$. Write

$$\gamma^{-1}\|H^{(\theta,0)}\|_{\vec{v},p} := \delta^{(1)}_p, \qquad \max\{\gamma^{-1}\|a_2\|_{\vec{v},p}, \gamma^{-1}\|b_2\|_{\vec{v},p}\} := \delta^{(2)}_p, \tag{6.4.65}$$

Finally we fix parameter $\mathfrak{p}_0 > (d + 1)/2$ and $\mathfrak{p}_2 > \mathfrak{p}_0$. In this way the norm $\|\cdot\|_{s,a,p}$ for $p \ge \mathfrak{p}_0$ defines a Banach algebra on the space $H^p(\mathbb{T}^d_s \times \mathbb{T}_a)$.

**Lemma 6.4.140 (Regularization).** *Fix $K>0$ large, $K_+ = K^{\frac{3}{2}}$ and $\rho_+>0$ small and consider $F$ as in* (6.4.58). *There exists $\eta=\eta(d,\mathfrak{p}_0)>0$ such that, if*

$$\rho_+^{-1}K^{\mu_1}\max\{\delta^{(1)}_{\mathfrak{p}_1},\delta^{(2)}_{\mathfrak{p}_1}\}\le\epsilon,\quad \delta^{(3)}_{\mathfrak{p}_1}\le\gamma_0\mathfrak{A}_0,\qquad \mathfrak{p}_0+\eta\le\mathfrak{p}_1<\mathfrak{p}_2, \tag{6.4.66}$$

*where $\mathfrak{A}_0>0$, $\mu_1\ge 2\mathfrak{p}_0+2\tau+4$ and $\epsilon=\epsilon(d,\mathfrak{p}_0)$ is small enough then there exists a tame, reversibility preserving map*

$$\mathcal{T}_+=\mathbb{1}+f:\mathcal{O}_0\times D_{a+\rho_+a_0,p+\nu}(s-\rho_+s_0,r-\rho_+r_0)\to D_{a,p+\nu}(s,r), \tag{6.4.67}$$

*with*

$$C_{\vec{v},p}(f)\le C\gamma_0^{-1}K^{\eta}\max\{\|a_2\|_{\vec{v},p},\|b_2\|_{\vec{v},p},\|H^{(\theta,0)}\|_{\vec{v},p},\},\ \ p\le\mathfrak{p}_2 \tag{6.4.68}$$

*that satisfies* (3.2.65), (3.2.66) *and* (3.2.67) *with $K_n\rightsquigarrow K$, $K_{n+1}\rightsquigarrow K_+$, $\rho_{n+1}\rightsquigarrow\rho_+$ and with $\mathfrak{p}_1:=\mathfrak{p}_0+\mu$, $\mathfrak{p}_2=\mathfrak{p}_0+\kappa_2$ defined in* (3.2.61) *and* (3.2.63), *provided $\mu\ge\mu_1$. Moreover for any $\xi\in\mathcal{O}_0$ such that*

$$|\omega\cdot l|\ge\frac{\gamma}{\langle l\rangle^{\tau}},\quad |l|\le K_+ \tag{6.4.69}$$

$$\hat F:=(\mathcal{T}_+)_*F:D_{a-2\rho_+a_0,p+\nu}(s-2\rho_+s_0,r-2\rho_+r_0)\to V_{a-2\rho a_0,p} \tag{6.4.70}$$

*has the form $N_0+\hat G$ and*

$$\hat F:=(1+h_+)\Big(\hat N_0^+ + N_+^{(1)}+N_+^{(2)}+H_+\Big) \tag{6.4.71}$$

*with*

$$\|h_+\|_{\vec{v},\mathfrak{p}_1}\le\delta^{(3)}_{\mathfrak{p}_1}(1+K_+^{(\mathfrak{p}_0+2\tau+2)}\delta^{(2)}_{\mathfrak{p}_1}),\qquad \|h_+\|_{\vec{v},\mathfrak{p}_2}\le\delta^{(3)}_{\mathfrak{p}_1}(1+K_+^{(\mathfrak{p}_0+2\tau+2)}\delta^{(2)}_{\mathfrak{p}_1})+K_+^{\mathfrak{p}_0+2\tau+2}\delta^{(2)}_{\mathfrak{p}_2}, \tag{6.4.72}$$

*$\hat N_0^+=\omega_+\cdot\partial_\theta+\widetilde\Omega_+^{-1}w\cdot\partial_w$ with $\widetilde\Omega^{-1}=c_+\Omega^{-1}$, and*

$$|\omega_+-\omega|\le\delta^{(1)}_{\mathfrak{p}_1},\quad |c_+-c|\le\delta^{(2)}_{\mathfrak{p}_1}. \tag{6.4.73}$$

*One has that $N_+^{(1)}$ as the form* (6.4.61) *with coefficients $a_2^+,b_2^+$, $N_+^{(2)}$ has the form* (6.4.62) *with coefficients $a_i^+,b_i^+$ for $i=0,1$, $\mathscr{K}_+$ of the form* (6.2.22) *with the same $N$ of $\mathscr{K}$ and coefficients $c_i^+,d_i^+$, and one has*

$$\begin{aligned}&C_{\vec{v},\mathfrak{p}_1}(N_+^{(2)})\le(1+K_+^{\mu_1}\delta^{(2)}_{\mathfrak{p}_1})C_{\vec{v},\mathfrak{p}_1}(N^{(2)})\\&C_{\vec{v},\mathfrak{p}_2}(N_+^{(2)})\le(1+K_+^{\mu_1}\delta^{(2)}_{\mathfrak{p}_1})(C_{\vec{v},\mathfrak{p}_2}(N^{(2)})+K_+^{2\mathfrak{p}_0+2\tau+2}(\delta^{(1)}_{\mathfrak{p}_2}+\delta^{(2)}_{\mathfrak{p}_2})\gamma^{-1}C_{\vec{v},\mathfrak{p}_1}(G)),\end{aligned} \tag{6.4.74}$$

$$\begin{aligned}&C_{\vec{v},\mathfrak{p}_1}(N_+^{(1)})\le(1+K_+^{\mu_1}\delta^{(2)}_{\mathfrak{p}_1})\Big[K_+^{-(\mathfrak{p}_2-\mathfrak{p}_1-2\mathfrak{p}_0-2\tau-2)}\delta^{(2)}_{\mathfrak{p}_2}+K_+^{2\mathfrak{p}_0+2\tau+2}(\delta^{(2)}_{\mathfrak{p}_1})^2\Big],\\&C_{\vec{v},\mathfrak{p}_2}(N_+^{(1)})\le C_{\vec{v},\mathfrak{p}_2}(N^{(1)})+K_+^{\mu_1}C_{\vec{v},\mathfrak{p}_1}(N^{(1)}).\end{aligned} \tag{6.4.75}$$

*where* $\vec{v} = (\gamma, \mathcal{O}, s, a)$, $\vec{v}_1 := (\gamma, \mathcal{O}, s - \rho_+ s_0, a - \rho_+ a_0)$ *and* $\vec{v}_2 := (\gamma, \mathcal{O}, s - 2\rho_+ s_0, a - 2\rho_+ a_0)$. *Moreover the following estimates hold:*

$$\begin{aligned} C_{\vec{v},\mathfrak{p}_1}(\hat{G}) &\leq (1 + K_+^{\mu_1} \delta_{\mathfrak{p}_1}^{(2)}) C_{\vec{v},\mathfrak{p}_1}(G) \\ C_{\vec{v},\mathfrak{p}_2}(\hat{G}) &\leq (1 + K_+^{\mu_1} \delta_{\mathfrak{p}_1}^{(2)})(C_{\vec{v},\mathfrak{p}_2}(G) + K_+^{\mu_1}(\delta_{\mathfrak{p}_2}^{(1)} + \delta_{\mathfrak{p}_2}^{(2)}) C_{\vec{v},\mathfrak{p}_1}(G)), \end{aligned} \tag{6.4.76}$$

$$\|H_+^{(\theta,0)}\|_{\vec{v}_2,\mathfrak{p}_1} \leq (1 + K_1^{\mu_1} \delta_{\mathfrak{p}_1}^{(2)}) \Big[ K_1^{-(\mathfrak{p}_2 - \mathfrak{p}_1)} \delta_{\mathfrak{p}_1}^{(1)} + K_1^{\mu_1} \delta_{\mathfrak{p}_1}^{(2)} \delta_{\mathfrak{p}_1}^{(1)} \Big], \tag{6.4.77}$$

$$\begin{aligned} \max_{i=1,\dots,N} \{\|c_i^+\|_{\vec{v}_2,\mathfrak{p}_1}, \|d_i^+\|_{\vec{v}_2,\mathfrak{p}_1}\} &\leq (1 + K_+^{\mu_1} \delta_{\mathfrak{p}_1}^{(2)}) \max_{i=1,\dots,N} \{\|c_i\|_{\vec{v},\mathfrak{p}_1}, \|d_i\|_{\vec{v},\mathfrak{p}_1}\} \\ \max_{i=1,\dots,N} \{\|c_i^+\|_{\vec{v}_2,\mathfrak{p}_2}, \|d_i^+\|_{\vec{v}_2,\mathfrak{p}_2}\} &\leq (1 + K_+^{\mu_1} \delta_{\mathfrak{p}_1}^{(2)}) \Big( \max_{i=1,\dots,N} \{\|c_i\|_{\vec{v},\mathfrak{p}_2}, \|d_i\|_{\vec{v},\mathfrak{p}_2}\} \\ &\quad + K_+^{\mu_1}(\delta_{\mathfrak{p}_2}^{(1)} + \delta_{\mathfrak{p}_2}^{(2)}) \max_{i=1,\dots,N} \{\|c_i\|_{\vec{v},\mathfrak{p}_1}, \|d_i\|_{\vec{v},\mathfrak{p}_1}\} \Big). \end{aligned} \tag{6.4.78}$$

*Finally one has*

$$\begin{aligned} &\|\mathcal{T}_+ a_2 - a_2^+\|_{\vec{v},\mathfrak{p}_1}, \|\mathcal{T}_+ b_2 - b_2^+\|_{\vec{v},\mathfrak{p}_1} \leq K_+^{\mu_1} \delta_{\mathfrak{p}_1}^{(2)} \Big[ K_+^{-(\mathfrak{p}_2 - \mathfrak{p}_1 - 2\mathfrak{p}_0 - 2\tau - 2)} \delta_{\mathfrak{p}_2}^{(2)} + K_+^{2\mathfrak{p}_0 + 2\tau + 2} (\delta_{\mathfrak{p}_1}^{(2)})^2 \Big], \\ &\|\mathcal{T}_+ a_i - a_i^+\|_{\vec{v},\mathfrak{p}_1}, \|\mathcal{T}_+ b_i - b_i^+\|_{\vec{v},\mathfrak{p}_1} \leq K_+^{\mu_1} \delta_{\mathfrak{p}_1}^{(2)} \max\{\|a_i\|_{\vec{v},\mathfrak{p}_1}, \|b_i\|_{\vec{v},\mathfrak{p}_1}\}, \qquad i = 0, 1, \\ &\|\mathcal{T}_+ c_i - c_i^+\|_{\vec{v},\mathfrak{p}_1}, \|\mathcal{T}_+ d_i - d_i^+\|_{\vec{v},\mathfrak{p}_1} \leq K_+^{\mu_1} \delta_{\mathfrak{p}_1}^{(2)} \max\{\|c_i\|_{\vec{v},\mathfrak{p}_1}, \|d_i\|_{\vec{v},\mathfrak{p}_1}\}, \qquad i = 1, \dots, N \end{aligned} \tag{6.4.79}$$

**Proof of Lemma** (6.4.140)**.** The proof follows the scheme used in Section 4.2. The differences are due to the fact we want to reduce to constant coefficients only the linearized operator in the normal directions.

**Step 1.** The first step is to diagonalize the second order coefficient by eliminating the terms $b_2$ through a multiplication operator. We use a transformation of the form

$$\Phi_1 : \theta \to \theta, \quad y \to y, \quad w \to \Pi_S^\perp \mathcal{T}_A \Pi_S^\perp w. \tag{6.4.80}$$

The eigenvalues of

$$\begin{pmatrix} c + a_2(\theta, x) & b_2(\theta, x) \\ \bar{b}_2(\theta, x) & c + \bar{a}_2(\theta, x) \end{pmatrix}$$

are $\lambda_{1,2} = \sqrt{(c + a_2)^2 - |b_2|^2}$. Hence we set $\widetilde{a}_2(\varphi, x) = \lambda_1 - c$. We have that $\widetilde{a}_2 \in \mathbb{R}$ because $a_2 \in \mathbb{R}$ and $a_i, b_i$ are small. The diagonalizing matrix is

$$\frac{1}{2c} \begin{pmatrix} 2c + a_2 + \widetilde{a}_2 & -b_2 \\ -\bar{b}_2 & 2c + a_2 + \widetilde{a}_2 \end{pmatrix} := \mathbb{1} + A. \tag{6.4.81}$$

We define $\mathcal{T}_A^{-1} := \mathbb{1} + \Pi_{K_+} A$, hence

$$\|\Pi_{K_+} A\|_{\vec{v},p} \leq C_{\vec{v},p}(\Pi_{K_+} N^{(1)}). \tag{6.4.82}$$

The bound on the inverse $\mathcal{T}_1$ follows simply by the fact that

$$\det(\mathbb{1} + A) := \frac{(|b_2|^2 - (2c + a_2 + \widetilde{a}_2)^2)}{4c^2}. \tag{6.4.83}$$

The bounds on the truncated matrix is the same. One can also prove that

$$\|(\mathbb{1} + \Pi_{K_+} A)^{-1} - (\mathbb{1} + A)^{-1}\|_{\vec{v},p} \le \|\Pi_{K_+}^{\perp} A\|_{\vec{v},p} + \|A\|_{\vec{v},\mathfrak{p}_1}\|A\|_{\vec{v},p}. \tag{6.4.84}$$

We set $F^{(1)} = (\Phi_1)_* F = N_0 + G^{(1)} = (1+h)(\widetilde{N}_0 + N_1^{(1)} + N_1^{(2)} + H^{(1)})$ (see notations in (6.4.58)) and one has

$$C_{\vec{v},p}(G^{(1)}) \le C_{\vec{v},p}(G) + K_+^{\nu+1}\|A\|_{\vec{v},p} C_{\vec{v},\mathfrak{p}_1}(G). \tag{6.4.85}$$

In particular the term $(H^{(1)})^{(\theta,0)}$ satisfies a bounds like (6.4.85) with $G \rightsquigarrow (H)^{(\theta,0)}$. Moreover using equation (6.4.56) of Lemma 6.4.138 with $M = cE + \Pi_{K_+} A$ we obtain that

$$N_1^{(1)} = -iE\Pi_S^{\perp}\left[\begin{pmatrix} \widetilde{a}_2(\theta,x) & 0 \\ 0 & \widetilde{a}_2(\theta,x)\end{pmatrix}\partial_{xx}\right]\Pi_S^{\perp} - iE\Pi_S^{\perp}\left[\begin{pmatrix} a_2^{(1)}(\theta,x) & b_2^{(1)}(\theta,x) \\ \bar{b}_2^{(1)}(\theta,x) & a_2^{(1)}(\theta,x)\end{pmatrix}\partial_{xx}\right]\Pi_S^{\perp} \tag{6.4.86}$$

where

$$\|a_2^{(1)}\|_{\vec{v},p}, \|b_2^{(1)}\|_{\vec{v},p} \le C_{\vec{v},p}(\Pi_{K_+}^{\perp}(N^{(1)})) + C_{\vec{v},p}(N^{(1)})C_{\vec{v},\mathfrak{p}_1}(N^{(1)}), \tag{6.4.87}$$

and $N_1^{(2)} = \mathcal{K}^{(1)} + \mathscr{K}^{(1)}$ with $\mathcal{K}^{(1)}$ as in (6.4.62) and $\mathscr{K}^{(1)}$ as in (6.2.22). By reversibility one has that $a_2$ is an even function of $\theta$, hence the transformation is reversibility preserving. Finally by (6.4.82) one has

$$\begin{aligned} C_{\vec{v},p}(N_1^{(2)}) &\le C_{\vec{v},p}(N^{(2)}) + C_{\vec{v},p+2}(\Pi_{K_+}N^{(1)})C_{\vec{v},\mathfrak{p}_1}(N^{(2)}) + C_{\vec{v},p+2}(\Pi_{K_+}N^{(1)}) \\ &\quad + C_{\vec{v},\mathfrak{p}_1+2}(\Pi_{K_+}N^{(1)})C_{\vec{v},p}(H) + C_{\vec{v},p+2}(\Pi_{K_+}N^{(1)})C_{\vec{v},\mathfrak{p}_1}(H), \end{aligned} \tag{6.4.88}$$

and the same bounds holds on the single coefficients $a_i^{(1)}, b_i^{(1)}$ for $i = 0,1$. Hence

$$\begin{aligned} C_{\vec{v},\mathfrak{p}_1}(N_1^{(2)}) &\le (1 + K_+^2\delta_{\mathfrak{p}_1}^{(2)})C_{\vec{v},\mathfrak{p}_1}(N^{(2)}) + K_+^3 C_{\vec{v},\mathfrak{p}_1}(H)\delta_{\mathfrak{p}_1}, \\ C_{\vec{v},\mathfrak{p}_2}(N_1^{(2)}) &\le C_{\vec{v},\mathfrak{p}_2}(N^{(2)}) + C_{\vec{v},\mathfrak{p}_1}(N^{(2)})K_+^3\delta_{\mathfrak{p}_1}^2 + K_+^2(C_{\vec{v},\mathfrak{p}_2}(H)\delta_{\mathfrak{p}_1}^{(2)} + C_{\vec{v},\mathfrak{p}_1}(H)\delta_{\mathfrak{p}_2}^{(2)}) \end{aligned} \tag{6.4.89}$$

One the coefficients $c_i^{(1)}, d_i^{(1)}$ fo $\mathscr{K}^{(1)}$ the bound (6.4.57) holds.

**Step 2 - Change of the space variable** Now we want to eliminate the dependence on $x$ of the coefficients $\widetilde{a}_2$ of the field $F^{(1)}$. We use a change of variables $\Phi_2$ as in (6.4.48) of Lemma 6.4.136. We are looking for $\alpha(\theta,x)$ such that the coefficient of the second order differential operator does not depend on $y$, namely by equation (6.4.49)

$$\mathcal{T}_\alpha^{-1}[(1+\widetilde{a}_2)(1+\alpha_x)^2] = 1 + a_2^{(2)}(\theta), \tag{6.4.90}$$

for some function $a_2^{(2)}(\theta)$. Since $\mathcal{T}_2$ operates only on the space variables, the (6.4.90) is equivalent to

$$(1+\widetilde{a}_2(\theta,x))(1+\alpha_x(\theta,x))^2 = 1+m_2(\theta). \tag{6.4.91}$$

Hence we have to set

$$\alpha_x(\theta,x) = \rho_0, \qquad \rho_0(\theta,x) := (1+m_2)^{\frac{1}{2}}(\theta)(1+\widetilde{a}_2(\theta,x))^{-\frac{1}{2}} - 1, \tag{6.4.92}$$

that has solution $\alpha$ periodic in $x$ if and only if $\int_{\mathbb{T}} \rho_0 dy = 0$. This condition implies

$$m_2(\theta) = \left(\frac{1}{2\pi}\int_{\mathbb{T}}(1+\widetilde{a}_2(\theta,x))^{-\frac{1}{2}}\right)^{-2} - 1 \tag{6.4.93}$$

Then we have the “approximate” solution (with zero average) of (6.4.92)

$$\alpha(\theta,x) := (\partial_x^{-1}\Pi_{K_+}\rho_0)(\theta,x), \tag{6.4.94}$$

where $\partial_x^{-1}$ is defined by linearity as $\partial_x^{-1}e^{ikx} := \frac{e^{ikx}}{ik}, \quad \forall\, k\in\mathbb{Z}\backslash\{0\}, \quad \partial_x^{-1} = 0$. In other word $\partial_x^{-1}h$ is the primitive of $h$ with zero average in $x$. The function $\alpha$ (that is a trigonometric polynomial) satisfies

$$\|\alpha\|_{\vec{v},\mathfrak{p}_1+\mathfrak{p}_0} \le K_+^{\mathfrak{p}_0}\delta_{\mathfrak{p}_1}^{(2)}, \quad \|\alpha\|_{\vec{v},\mathfrak{p}_2+\mathfrak{p}_0} \le K_+^{\mathfrak{p}_0}\delta_{\mathfrak{p}_2}^{(2)} \tag{6.4.95}$$

For more details on the estimates on $\alpha$ we refer to [31].

With this definition of the function $\alpha$ and by Lemma A.171 one has that $\mathcal{T}_\alpha : \mathbb{T}_{a-(\rho/4)a_0} \to \mathbb{T}_\alpha$ if, by (6.4.95), $\delta_{\mathfrak{p}_1}^{(2)}$ is small enough. Setting

$$F^{(2)} := (\Phi_2)_* F^{(1)} = N_0 + G^{(2)} = (1+h)(\hat{N}_0 + N_2^{(1)} + N_2^{(2)} + H^{(2)}), \tag{6.4.96}$$

again one has

$$C_{\vec{v},p}(G^{(2)}) \le (1+\|\alpha\|_{\vec{v},\mathfrak{p}_1+\nu+1})\Big(C_{\vec{v},p}(G^{(1)}) + \|\alpha\|_{\vec{v},p+\nu+1}C_{\vec{v},\mathfrak{p}_1}(G^{(1)})\Big), \tag{6.4.97}$$

where we used $\|\alpha\|_{\vec{v},\mathfrak{p}_1+\nu+1} \le K^{\nu+1}\delta_{\mathfrak{p}_1}^{(2)} \le 1$. Moreover we have obtained

$$N_2^{(1)} = -iE\Pi_S^\perp\left[\begin{pmatrix} m_2(\theta) & 0 \\ 0 & m_2(\theta)\end{pmatrix}\partial_{xx}\right]\Pi_S^\perp - iE\Pi_S^\perp\left[\begin{pmatrix} a_2^{(2)}(\theta,x) & b_2^{(2)}(\theta,x) \\ \bar{b}_2^{(2)}(\theta,x) & a_2^{(2)}(\theta,x)\end{pmatrix}\partial_{xx}\right]\Pi_S^\perp \tag{6.4.98}$$

where

$$\|a_2^{(2)}\|_{\vec{v},p}, \|b_2^{(2)}\|_{\vec{v},p} \le C_{\vec{v},p}(\Pi_{K_+}^\perp(N^{(1)})) + C_{\vec{v},p}(N^{(1)})C_{\vec{v},\mathfrak{p}_1}(N^{(1)}), \tag{6.4.99}$$

with $N_2^{(2)} = \mathcal{K}^{(2)} + \mathscr{K}^{(2)}$ and $\mathcal{K}^{(2)}$ as in (6.4.62) and on the coefficients $a_i^{(2)}, b_i^{(2)}$ for $i = 0, 1$ bound (6.4.51) hold in term of the coefficients $a_i^{(1)}, b_i^{(1)}$ where $C_{\vec{v},p}(f)$ is replaced by $\|a_2\|_{\gamma,\mathcal{O},s,a-(\rho/4)a_0,p}$. More explicitly one has

$$\begin{aligned} &C_{\vec{v},\mathfrak{p}_1}(N_2^{(2)}) \leq (1 + K_+^{\mathfrak{p}_0+2}\delta_{\mathfrak{p}_1}^{(2)})C_{\vec{v},\mathfrak{p}_1}(N_1^{(2)}) + K_+^{\mathfrak{p}_0+2}\delta_{\mathfrak{p}_1}^{(2)}C_{\vec{v},\mathfrak{p}_1}(N_1^{(1)}), \\ &C_{\vec{v},\mathfrak{p}_1}(N_2^{(2)}) \leq C_{\vec{v},\mathfrak{p}_2}(N_1^{(2)}) + K_+^{\mathfrak{p}_0+2}\delta_{\mathfrak{p}_2}C_{\vec{v},\mathfrak{p}_1}(N_1^{(2)}) + K_+^{\mathfrak{p}_0+2}(C_{\vec{v},\mathfrak{p}_2}(N_1^{(1)})\delta_{\mathfrak{p}_1}^{(2)} + C_{\vec{v},\mathfrak{p}_1}(N_1^{(1)})\delta_{\mathfrak{p}_2}^{(2)}) \end{aligned} \tag{6.4.100}$$

Moreover the coefficients $c_i^{(2)}, d_i^{(2)}$ of $\mathscr{K}^{(2)}$ satisfy the bound (6.4.50). For more details on the estimates on $\alpha$ we refer to [31].

Note that in this two steps the function $h(\theta)$ did not change.

**Step 3 - Time reparameterization.** In order to eliminate the dependence on $\theta$ of $a_2^{(2)}$ we use a special diffeomorphism of the torus

$$\mathcal{T}_\beta : \theta \to \theta_+ = \theta + \omega\beta(\theta), \quad \theta \in \mathbb{T}_s^d, \quad \beta(\theta) \in \mathbb{R}, \tag{6.4.101}$$

where $\alpha$ is a small real valued function, $2\pi-$periodic in all its arguments. We consider a transformation $\Phi_3$ of the form (6.4.54) and we set

$$F^{(3)} := (\Phi_3)_* F^{(2)} = N_0 + G^{(3)} \tag{6.4.102}$$

with the usual notation $N^{(3)} := \Pi_{\mathcal{N}}(\Phi_3)_* F^{(2)}$, $X^{(3)} := \Pi_{\mathcal{A}}(\Phi_3)_* F^{(2)}$ and $R^{(3)} := \Pi_3(\Phi_3)_* F^{(2)}$. By the nature of the transformation $\Phi_3$ we also have that

$$\Pi_{\mathcal{N}}(\Phi_3)_* F^{(2)} := (\Phi_3)_* \Pi_{\mathcal{N}} F^{(2)}, \quad \Pi_{\mathcal{A}}(\Phi_3)_* F^{(2)} := (\Phi_3)_* \Pi_{\mathcal{A}} F^{(2)}, \quad \Pi_3(\Phi_3)_* F^{(2)} := (\Phi_3)_* \Pi_{\mathcal{R}} F^{(2)}, \tag{6.4.103}$$

Let us study in detail the from of $N^{(3)}$. First of all we have

$$((\Phi_3)_* \Pi_{\mathcal{N}} F^{(2)})^{(\theta)}(\theta_+) = \mathcal{T}_\beta^{-1}\Big[(1 + \omega \cdot \partial_\theta \beta)(1 + h)[\omega + (H^{(2)})^{(\theta,0)}]\Big](\theta_+), \tag{6.4.104a}$$

$$\begin{aligned} ((\Phi_3)_* \Pi_{\mathcal{N}} F^{(2)})^{(w)}(\theta_+, w_+) = \mathcal{T}_\beta^{-1}\Big[(1 + h)\Big(\hat{N}_0^{(w)} + N_2^{(1)} + N_2^{(2)} + d_w H^{(2)} w\Big)\Big] = \\ - iE\Pi_S^\perp \begin{pmatrix} \mathcal{T}_\beta^{-1}(1 + h)(c + m_2) & 0 \\ 0 & \mathcal{T}_\beta^{-1}(1 + h)(c + m_2) \end{pmatrix} \partial_{xx} \mathcal{T}_\beta \Pi_S^\perp \\ - iE\Pi_S^\perp \mathcal{T}_\beta^{-1} \left[ \begin{pmatrix} a_1^{(2)} & b_1^{(2)} \\ \bar{b}_1^{(2)} & \bar{a}_1^{(2)} \end{pmatrix} \partial_x + \begin{pmatrix} a_0^{(2)} & b_0^{(2)} \\ \bar{b}_0^{(2)} & \bar{a}_0^{(2)} \end{pmatrix} \right] \mathcal{T}_\beta \Pi_S^\perp + \mathcal{T}_\beta^{-1} \mathscr{K}^{(2)} \mathcal{T}_\beta \end{aligned} \tag{6.4.104b}$$

Our aim is to find $\beta$ in such a way the coefficients of the second order is proportional with respect to the coefficients of $\omega$. This is equivalent to require that

$$(1 + h(\theta))(c + m_2(\theta)) = c_+(1 + \omega \cdot \partial_\theta \beta(\theta))(1 + h(\theta)), \tag{6.4.105}$$

for some $c_+ := c + \mathfrak{c}$. By setting

$$\mathfrak{c} = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} (m_2(\theta)) d\theta, \tag{6.4.106}$$

we can find the (unique) approximate solution of (6.4.105) with zero average

$$\beta(\theta) := \frac{1}{c + \mathfrak{c}} (\omega \cdot \partial_\theta)^{-1} (c + \Pi_K(m_2) - c - \mathfrak{c})(\theta), \tag{6.4.107}$$

where $(\omega \cdot \partial_\theta)^{-1}$ is defined by linearity $(\omega \cdot \partial_\varphi)^{-1} e^{i\ell \cdot \varphi} := \frac{e^{i\ell\cdot\varphi}}{i\omega\cdot\ell}$, $\ell \neq 0$, $\quad (\omega \cdot \partial_\varphi)^{-1} 1 = 0$. Note that $\beta$ is trigonometric polynomial supported on $|\ell| \leq K$. As one can check (see also [31] for more details) the function $\beta$ in (6.4.107) satisfy the bound

$$\|\beta\|_{\vec{v}, p+\mathfrak{p}_0} \leq \gamma_0^{-1} \|\Pi_K m_2\|_{\vec{v}, p+\mathfrak{p}_0+2\tau+2}, \tag{6.4.108}$$

with $\vec{v} = (\gamma, \mathcal{O}, s - (\rho/4), a - (\rho/4)a_0)$ and $\gamma_0$ is the diophantine constant of $\omega$. Hence we have

$$\|\beta\|_{\vec{v}, \mathfrak{p}_1+\mathfrak{p}_0} \leq K_+^{\mathfrak{p}_0+2\tau+2} \delta^{(2)}_{\mathfrak{p}_1}, \qquad \|\beta\|_{\vec{v}, \mathfrak{p}_2+\mathfrak{p}_0} \leq K_+^{\mathfrak{p}_0+2\tau+2} \delta^{(2)}_{\mathfrak{p}_2} \tag{6.4.109}$$

This implies that the transformation $\mathcal{T}_3$ maps $\mathbb{T}^d_{s-(\rho/4)s_0} \to \mathbb{T}^d_s$ if $\delta_2$ is sufficiently smaller than $\rho$ (see condition (6.4.66)). We set

$$\widetilde{h}_+ := \mathcal{T}_\beta^{-1} (\omega \cdot \partial_\theta \beta + h + h\omega \cdot \partial_\theta \beta), \tag{6.4.110}$$

so that one has for $\vec{v} = (\gamma, \mathcal{O}, s - \rho s_0, a - \rho a_0)$

$$\|\widetilde{h}_+ - \mathcal{T}_3 h\|_{\vec{v},p} \leq \|\omega \cdot \partial_\theta \beta\|_{\vec{v},p} + \|\omega \cdot \partial_\theta \beta\|_{\vec{v}, 2\mathfrak{p}_0} \|h\|_{\vec{v},p} \leq K_+^{\mathfrak{p}_0+2\tau+2} (\|a_2^{(2)}\|_{\vec{v},p} + \|a_2\|_{\vec{v},\mathfrak{p}_0} \|h\|_{\vec{v},p}), \tag{6.4.111}$$

which for $\mathfrak{p}_1$ large enough, implies the bound (6.4.72). We have

$$F^{(3)} = N_0 + G^{(3)} = (1 + \widetilde{h}_+)(\widetilde{N}_0 + N_3^{(1)} + N_3^{(2)} + H^{(3)}), \tag{6.4.112}$$

with $\widetilde{N}_0 := \omega \cdot \partial_\theta + c_+ \Omega^{-1} w \cdot \partial_w$, with $c_+ = c + \mathfrak{c}$,

$$N_3^{(1)} = -iE\Pi_S^\perp \left[ \begin{pmatrix} a_2^{(3)}(\theta, x) & b_2^{(3)}(\theta, x) \\ \bar{b}_2^{(3)}(\theta, x) & a_2^{(3)}(\theta, x) \end{pmatrix} \partial_{xx} \right] \Pi_S^\perp \tag{6.4.113}$$

where

$$\begin{aligned} C_{\vec{v},p}(N_3^{(1)}) \leq{}& C_{\vec{v},p}(\Pi^\perp_{K_+}(N_2^{(1)})) + \|\beta\|_{\vec{v},p+\mathfrak{p}_0} C_{\vec{v},\mathfrak{p}_1}(\Pi^\perp_{K_+}(N_2^{(1)})) \\ &+ \max\{\|a_2^{(2)}\|_{\vec{v},p}, \|b_2^{(2)}\|_{\vec{v},p}\} + \|\beta\|_{\vec{v},p+\mathfrak{p}_0} \max\{\|a_2^{(2)}\|_{\vec{v},\mathfrak{p}_1}, \|b_2^{(2)}\|_{\vec{v},\mathfrak{p}_1}\} \end{aligned} \tag{6.4.114}$$

and the coefficients $a_2^3, b_2^3$ satisfies the same bound. and $N_3^{(2)} := (\Phi_3)_*(\hat{N}_0 + N_2^{(1)} + N_2^{(2)}) - (\widetilde{N}_0 + N_3^{(1)}) := \mathcal{K}^{(3)} + \mathscr{K}^{(3)}$ collects all the terms of order at most $O(\partial_x)$ and $\mathcal{K}^{(3)}$ as in (6.4.62) with coefficients $a_i^{(3)}, b_i^{(3)}$ for $i = 0, 1$ which satisfy the bound

$$\|a_i^{(3)}\|_{\vec{v},p}, \|b_i^{(3)}\|_{\vec{v},p} \leq \|a_i^{(2)}\|_{\vec{v},p} + \|\beta\|_{\vec{v},p+\mathfrak{p}_0} \|a_i^{(2)}\|_{\vec{v},\mathfrak{p}_1}, \tag{6.4.115}$$

More explicitly one has

$$\begin{aligned}
&C_{\vec{v},\mathfrak{p}_1}(N_3^{(1)}) \le (1 + K_+^{\mathfrak{p}_0+2\tau+2}\delta_{\mathfrak{p}_1}^{(2)})\Big[K_+^{-(\mathfrak{p}_2-\mathfrak{p}_1)}C_{\vec{v},\mathfrak{p}_2}(N_2^{(1)}) + \max\{\|a_2^{(2)}\|_{\vec{v},\mathfrak{p}_1}, \|b_2^{(2)}\|_{\vec{v},\mathfrak{p}_1}\}\Big],\\
&C_{\vec{v},\mathfrak{p}_2}(N_3^{(1)}) \le C_{\vec{v},\mathfrak{p}_2}(N_2^{(1)}) + K_+^{-(\mathfrak{p}_2-\mathfrak{p}_1)+\mathfrak{p}_0+2\tau+2}\delta_{\mathfrak{p}_2}^{(2)}C_{\vec{v},\mathfrak{p}_2}(N_2^{(1)})\\
&\qquad\qquad + K_+^{\mathfrak{p}_0+2\tau+2}\delta_{\mathfrak{p}_2}^{(2)}\max\{\|a_2^{(2)}\|_{\vec{v},\mathfrak{p}_1}, \|b_2^{(2)}\|_{\vec{v},\mathfrak{p}_1}\},\\
&C_{\vec{v},\mathfrak{p}_1}(N_3^{(2)}) \le (1 + K_+^{\mathfrak{p}_0+2\tau+2}\delta_{\mathfrak{p}_1}^{(2)})C_{\vec{v},\mathfrak{p}_1}(N_2^{(2)}),\\
&C_{\vec{v},\mathfrak{p}_2}(N_3^{(2)}) \le C_{\vec{v},\mathfrak{p}_2}(N_2^{(2)}) + K_+^{\mathfrak{p}_0+2\tau+2}\delta_{\mathfrak{p}_2}^{(2)}C_{\vec{v},\mathfrak{p}_1}(N_2^{(2)}).
\end{aligned} \tag{6.4.116}$$

By equation (6.4.106) we also obtain the bound $\sup_{\mathcal{O}}|c_+ - c| \le \delta_{\mathfrak{p}_1}^{(2)}$ on the constant $c_+$ hence (6.4.73) is satisfied. Using the estimates given by Lemmata 6.4.136, 6.4.137, 6.4.138 and collecting the estimates in (6.4.82), (6.4.95) and (6.4.109) we get on the field $G^{(3)}$

$$\begin{aligned}
&C_{\vec{v},\mathfrak{p}_1}(G^{(3)}) \le (1 + K_+^{(\mathfrak{p}_0+2\tau+2)}\delta_{\mathfrak{p}_1}^{(2)})C_{\vec{v},\mathfrak{p}_1}(G),\\
&C_{\vec{v},\mathfrak{p}_2}(G^{(3)}) \le (1 + K_+^{(\mathfrak{p}_0+2\tau+2)}\delta_{\mathfrak{p}_1}^{(2)})C_{\vec{v},\mathfrak{p}_2}(G) + (1 + K_+^{(\mathfrak{p}_0+2\tau+2)}\delta_{\mathfrak{p}_1}^{(2)})C_{\vec{v},\mathfrak{p}_1}(G)K_+^{\mathfrak{p}_0+2\tau+2}\delta_{\mathfrak{p}_2}^{(2)}.
\end{aligned} \tag{6.4.117}$$

Moreover the coefficients $c_i^{(3)}, d_i^{(3)}$ of $\mathscr{K}^{(3)}$ satisfy the bound (6.4.50) and finally

$$\begin{aligned}
&\|(H^{(3)})^{(\theta,0)}\|_{\vec{v},\mathfrak{p}_1} \le (1 + K_+^{(\mathfrak{p}_0+2\tau+2)}\delta_{\mathfrak{p}_1}^{(2)})\delta_{\mathfrak{p}_1}^{(1)},\\
&\|(H^{(3)})^{(\theta,0)}\|_{\vec{v},\mathfrak{p}_2} \le \|H^{(\theta,0)}\|_{\vec{v},\mathfrak{p}_2} + K_+^{\mathfrak{p}_0+2\tau+2}\delta_{\mathfrak{p}_2}^{(2)}\delta_{\mathfrak{p}_1}^{(1)}.
\end{aligned} \tag{6.4.118}$$

**Step 4 - Torus diffeomorphism.** The aim of the final step is reduce "quadratically" the size of the term $(H^{(3)})^{(\theta,0)}$. We define the trasformation

$$\Phi_4 : (\theta, y, w) \to (\theta + g(\theta), y, w), \tag{6.4.119}$$

and we call $\mathcal{T}_g$ the transformation $\mathcal{T}_g u = u(\theta + g(\theta), x)$. We set $\hat{F} = (\Phi_4)_* F^{(3)} = N_0 + \hat{G}$ and we study its projection on the subspace $\mathcal{N}$. By a direct calculation one can note that $\Pi_{\mathcal{N}}(\Phi_4)_* F^{(3)} = (\Phi_4)_*\Pi_{\mathcal{N}}F^{(3)}$, $\Pi_{\mathcal{A}}(\Phi_4)_* F^{(3)} = (\Phi_4)_*\Pi_{\mathcal{A}}F^{(3)}$ and $\Pi_{\mathcal{R}}(\Phi_4)_* F^{(3)} = (\Phi_4)_*\Pi_{\mathcal{R}}F^{(3)}$. For convenience we write

$$\begin{aligned}
(\Phi_4)_*\Pi_{\mathcal{N}}F^{(3)} &:= (\Phi_4)_*\Big((1 + \widetilde{h}_+)(\widetilde{N}_0 + (H^{(3)})^{(\theta,0)})\Big)\\
&\quad + (\Phi_4)_*\Big((1 + \widetilde{h}_+)(N_3^{(1)} + N_3^{(2)} + (\Pi_{\mathcal{N}}H^{(3)})^{(w)})\Big) =\\
&= (1 + \mathcal{T}_g^{-1}\widetilde{h}_+)\Big[(\Phi_4)_*(\widetilde{N}_0 + (H^{(3)})^{(\theta,0)}) + (\Phi_4)_*(N_3^{(1)} + N_3^{(2)} + (\Pi_{\mathcal{N}}H^{(3)})^{(w)})\Big]
\end{aligned} \tag{6.4.120}$$

We set $h_+ := \mathcal{T}_g^{-1}\widetilde{h}_+$. Moreover by Lemma B.180 one has

$$(\Phi_4)_* N^{(3)} = N^{(3)} + [g, N^{(3)}] + \int_0^1\int_0^t (\Phi^{(4)})_*^s[g,[g,N^{(3)}]]dsdt.$$

where $N^{(3)} := (\widetilde{N}_0 + (H^{(3)})^{(\theta,0)})$. We look for $g(\theta)$ such that the $\theta$ component of $N^{(3)} + [g, N^{(3)}]$ has the form in (6.4.71) with the size of $H_+^{(\theta,0)}$ "much smaller" of the size of $H^{(\theta,0)}$. More explicitly one has that

$$(N^{(3)} + [g, N^{(3)}])^{(\theta)} = (\omega_+ + \widetilde{H}^{(\theta,0)}(\theta)), \tag{6.4.121}$$

where , denoting by $\langle\cdot\rangle$ the average in the variable $\theta$,

$$\begin{aligned} \omega_+ &:= \omega + \langle (H^{(3)})^{(\theta,0)} \rangle, \\ \widetilde{H}^{(\theta,0)} &:= (H^{(3)})^{(\theta,0)}(\theta) - \langle (H^{(3)})^{(\theta,0)} \rangle + \omega \cdot \partial_\theta g(\theta) + (H^{(3)})^{(\theta,0)} \partial_\theta g(\theta) - \partial_\theta (H^{(3)})^{(\theta,0)} g(\theta) \end{aligned} \tag{6.4.122}$$

and we look for $g(\theta)$ such that

$$\Pi_K \left[ (H^{(3)})^{(\theta,0)}(\theta) - \langle (H^{(3)})^{(\theta,0)} \rangle \right] + \omega \cdot \partial_\theta g(\theta) = 0. \tag{6.4.123}$$

Equation (6.4.123) is satisfied by choosing

$$g(\theta) := (\omega \cdot \partial_\theta^{-1})^{-1} \Pi_K \left[ (H^{(3)})^{(\theta,0)}(\theta) - \langle (H^{(3)})^{(\theta,0)} \rangle \right]. \tag{6.4.124}$$

One has the following estimates hold

$$\begin{aligned} &\|g\|_{\vec{v},\mathfrak{p}_1} \le K_1^{2\tau+2} \gamma^{-1} C_{\vec{v},\mathfrak{p}_1}((H^{(3)})^{(\theta,0)}) \le K_1^{2\tau+2}(1 + K_+^{\mathfrak{p}_0+2\tau+2} \delta_{\mathfrak{p}_1}^{(2)}) \delta_{\mathfrak{p}_1}^{(1)}, \\ &\|g\|_{\vec{v},\mathfrak{p}_2} \le K_+^{2\tau+1}(\delta_{\mathfrak{p}_2}^{(1)} + K_+^{\mathfrak{p}_0+2\tau+2} \delta_{\mathfrak{p}_2}^{(2)} \delta_{\mathfrak{p}_1}^{(1)}). \end{aligned} \tag{6.4.125}$$

Moreover by the first of (6.4.122) we have that (6.4.73) holds. Now for $\delta_{\mathfrak{p}_1}^{(1)}, \delta_{\mathfrak{p}_1}^{(2)}$ small enough (see condition (6.4.66)) one can use Lemma A.171 to conclude that

$$\mathcal{T}_g : \mathbb{T}^d_{s-(\rho_+/4)s_0} \to \mathbb{T}^d_{s-(\rho_+/2)s_0}. \tag{6.4.126}$$

and hence

$$\hat{F} = (\Phi_4)_* F^{(3)} : D_{a-\rho_+ a_0, p+\nu}(r - \rho_+ r_0, r - \rho_+ r_0) \to V_{a-\rho_+ a_0, p}.$$

We set $\mathcal{T} := \Phi_4 \circ \Phi_3 \circ \Phi_2 \circ \Phi_1$. By the discussion above we have that

$$\mathcal{T}_+ : D_{a-\rho_+ a_0, p+\nu}(s - \rho_+ s_0, r_0 - \rho_+ r_0) \to D_{a,p+\nu}(s, r),$$

and moreover satisfies conditions 3.2.65, (3.2.66) and (3.2.67) by choosing $\mu \ge 2\mathfrak{p}_0 + 2\tau + 4$. Indeed we in the abstract theorem (3.2.39) we have fixed $\mathfrak{p}_1 := \mathfrak{p}_0 + \kappa_1$, and hence (see condition (3.2.62))

$$2\mathfrak{p}_0 + \kappa_1 + 2\tau + 4 \le \kappa_3 \le \kappa_1 + 3\mu. \tag{6.4.127}$$

By the discussion above we have

$$\hat{F} := N_0 + ((\Phi_4)_*\widetilde{N}_0 - \widetilde{N}_0) + (\Phi_4)_*(\widetilde{h}_+\widetilde{N}_0) + (\Phi^{(4)})_*\Big((1+\widetilde{h}_+)(N_3^{(1)} + N_3^{(2)} + (H^{(3)})^{(\theta,0)})\Big) \\ + (\Phi_4)_*((1+\widetilde{h}_+)\Pi_{\mathcal{N}^\perp}H^{(3)}). \tag{6.4.128}$$

Fix now $\vec{v} = (\gamma, \mathcal{O}, s - \rho_+ s_0, a - \rho_+ a_0)$. From this first splitting we have that $\hat{F} = \widetilde{N}_0 + \hat{G}$ and on $\hat{G}$ the following estimates hold:

$$C_{\vec{v},\mathfrak{p}_1}(\hat{G}) \leq (1 + K_+^\mu \delta_{\mathfrak{p}_1}^{(2)}) C_{\vec{v},\mathfrak{p}_1}(N^{(1)} + N^{(2)} + H) \\ C_{\vec{v},\mathfrak{p}_2}(\hat{G}) \leq (1 + K_+^\mu \delta_{\mathfrak{p}_1}^{(2)})(C_{\vec{v},\mathfrak{p}_2}(N^{(1)} + N^{(2)} + H) + K_+^{2\mathfrak{p}_0+2\tau+2}(\delta_{\mathfrak{p}_2}^{(1)} + \delta_{\mathfrak{p}_2}^{(2)})\gamma^{-1}C_{\vec{v},\mathfrak{p}_1}(G)), \tag{6.4.129}$$

which implies (6.4.76). To prove (6.4.129) one uses (6.4.117) and the smallness of $\delta$ and that

$$\begin{aligned} ((\Phi^{(4)})_*\widetilde{N}_0 - \widetilde{N}_0) &= \int_0^1 (\Phi_4)_*^s[g, \widetilde{N}_0] = \int_0^1 (\Phi_4)_*^s \Pi_{K_+}[g, \widetilde{N}_0] = \\ &= \int_0^1 (\Phi_4)_*^s \Pi_{K_+}[g, \widetilde{N}_0 + (H^{(3)})^{(\theta,0)}] - \int_0^1 (\Phi_4)_*^s \Pi_{K_+}[g, (H^{(3)})^{(\theta,0)}] \\ &= \int_0^1 (\Phi_4)_*^s r + \int_0^1 (\Phi_4)_*^s \Pi_{K_+}[g, (H^{(3)})^{(\theta,0)}], \end{aligned}$$

where by (6.4.123)

$$r = \Big(\Pi_{K_+}^\perp (H^{(3)})^{(\theta,0)} + (H^{(3)})^{(\theta,0)}\partial_\theta g(\theta) - \partial_\theta (H^{(3)})^{(\theta,0)} g(\theta)\Big) \cdot \partial_\theta \tag{6.4.130}$$

Clearly the estimates (6.4.129) follows. Trivially one has also

$$C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{N}^\perp}\hat{G}) \leq (1 + K_+^\mu \delta_{\mathfrak{p}_1}^{(2)}) C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{N}^\perp}G) \\ C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{N}^\perp}\hat{G}) \leq (1 + K_+^\mu \delta_{\mathfrak{p}_1})(C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{N}^\perp}G) + K_+^{\mathfrak{p}_0+2\tau+2}\mathfrak{A}_0\gamma^{-1}C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{N}^\perp}G)), \tag{6.4.131}$$

Now we want to rewrite the field $\hat{F}$ in a form more similar to (6.4.112) by using (6.4.120). We have

$$\begin{aligned} \hat{F} &:= (1 + \Phi_4^{-1}\widetilde{h}_+)\left[\hat{N}_0^+ + N_4^{(1)} + N_4^{(2)} + H^{(4)}\right], \\ \hat{N}_0^+ &:= (\omega + \langle (H^{(3)})^{(\theta,0)}\rangle)\cdot\partial_\theta + c_+\Omega^{-1}w\cdot\partial_w, \\ N_4^{(1)} &:= (\Phi_4)_* N_3^{(1)}, \quad N_4^{(2)} := (\Phi_4)_* N_3^{(1)}, \\ H^{(4)} &:= (\Phi_4)_* H^{(3)} + (\Phi_4)_*\widetilde{N}_0 - \hat{N}_0^+. \end{aligned} \tag{6.4.132}$$

This is another way to write (6.4.128). But now we give a precise estimates on the low norm of the component $\theta$ of the field $H^{(4)}$ on $\mathcal{N}$. First of all we have

$$H^{(4)} := \widetilde{N}_0 - \widetilde{N}_0^+ + [g, \widetilde{N}_0] + \int_0^1\int_0^t (\Phi_4)_*^{(s)}[g,[g,\widetilde{N}_0]] + H^{(3)} + \int_0^1 (\Phi_4)_*^{(s)}[g, H^{(3)}].$$

Now if we look at the the component $(H^{(4)})^{(\theta,0)}$, by using equation (6.4.123), we obtain

$$\begin{aligned}(H^{(4)})^{(\theta,0)} &:= \Pi^{\perp}_{K_+}(H^{(3)})^{(\theta,0)}(\theta) + \int_0^1\int_0^t (\Phi_4)^{(s)}_*[g,[g,\widetilde{N}_0]] + \int_0^1 (\Phi_4)^{(s)}_*[g,H^{(3)}] \\ &= \Pi^{\perp}_{K_+}(H^{(3)})^{(\theta,0)}(\theta) + \int_0^1 (\Phi_4)^{(s)}_*[g,H^{(3)}] + \int_0^1\int_0^t (\Phi_4)^{(s)}_*[g,-\Pi_{K_+}(H^{(3)})^{(\theta,0)}],\end{aligned} \tag{6.4.133}$$

and hence using (6.4.66) we get

$$\begin{aligned}\|(H^{(4)})^{(\theta,0)}\|_{\vec{v},\mathfrak{p}_1} &\le K_+^{-(\mathfrak{p}_2-\mathfrak{p}_1)} C_{\vec{v},\mathfrak{p}_2}((H^{(3)})^{(\theta,0)}) + K_+^{\nu+2}\|g\|_{\vec{v},\mathfrak{p}_1}\|(H^{(3)})^{(\theta,0)}\|_{\vec{v},\mathfrak{p}_1} \\ &\overset{(6.4.118),(6.4.125)}{\le} K_1^{-(\mathfrak{p}_2-\mathfrak{p}_1)}(1+K_1^{\mathfrak{p}_0+2\tau+1}\delta^{(2)}_{\mathfrak{p}_1})(\delta^{(1)}_{\mathfrak{p}_2}+\delta^{(1)}_{\mathfrak{p}_1}K_+^{\mathfrak{p}_0+2\tau+2}\delta^{(2)}_{\mathfrak{p}_1}) \\ &\quad + (1+K_+^{\mathfrak{p}_0+2\tau+1}\delta^{(2)}_{\mathfrak{p}_1})K_+^{\mathfrak{p}_0+2\tau+2}\delta^{(2)}_{\mathfrak{p}_1}\delta^{(1)}_{\mathfrak{p}_1},\end{aligned} \tag{6.4.134}$$

which implies (6.4.77). We conclude by noting that $N_4^{(2)} = \mathcal{K}^4 + \mathscr{K}^{(4)}$ with the coefficients $c_i^{(4)}, d_i^{(4)}$ of $\mathscr{K}^{(4)}$ that satisfy the bound (6.4.50). By collecting the bounds on $\mathscr{K}^{(1)}, \mathscr{K}^{(2)}, \mathscr{K}^{(3)}, \mathscr{K}^{(4)}$ we get the (6.4.78). Now by (6.4.132) we have, by using (6.4.116), (6.4.99)

$$\begin{aligned}C_{\vec{v},\mathfrak{p}_1}(N_4^{(1)}) &\le (1+K_+^{2\mathfrak{p}_0+2}\delta^{(2)}_{\mathfrak{p}_1})^3\Big[K_+^{-(\mathfrak{p}_2-\mathfrak{p}_1-2\mathfrak{p}_0-2\tau-2)}\delta^{(2)}_{\mathfrak{p}_2} + K_+^{2\mathfrak{p}_0+2\tau+2}(\delta^{(2)}_{\mathfrak{p}_1})^2\Big], \\ C_{\vec{v},\mathfrak{p}_2}(N_4^{(1)}) &\le C_{\vec{v},\mathfrak{p}_2}(N^{(1)}) + K_+^{2\mathfrak{p}_0+2\tau+4}C_{\vec{v},\mathfrak{p}_1}(N^{(1)}).\end{aligned} \tag{6.4.135}$$

which implies (6.4.75). In the same way , using (6.4.116), (6.4.100) and (6.4.88), the (6.4.74) follows. Bounds in (6.4.79) follow by the discussion above recalling the results of Lemmata 6.4.136, 6.4.137 and 6.4.138.

□

### 6.4.2 Inversion of the linearized operator in the normal directions

To prove our main Theorem 1.2.5 we said we want to use the abstract Theorem 3.2.39 on the field $F_0$ defined in (6.2.13) of Section 6.2. Theroem (3.2.39) is based on an iterative scheme and on the existence of sets of "good" parameters defined inductively (see Definition (3.2.37)). In this Section we give an explicit formulation of such sets in terms of the eigenvalues of the linearized operator in the normal directions. Moreover by Hypothesis 1 we have that our vector field $F_0$ has a particular structure of "reversibility" as explained in (6.2.17) and (6.2.18) of Section 6.2. In other words we want to work on a subspace (see Definition 3.2.36) of vector fields $\mathcal{E}$ that are reversible with respect to the involution $S$ in (6.2.17). Since the reversibility is given in terms of some parity conditions of the coefficients of the fields, one can easily check that our subspace $\mathcal{E}$ of reversible fields has the property $\mathcal{P}$ of Definition 3.2.41. Hence, by Lemma 3.2.43, we can study the sets of good parameters defined in 3.2.42 instead of those defined in 3.2.37.

Now we inductively assume that at some step of the iteration the vector field $F_0$ has been transformed into a field $F = N_0 + G$ of the form (6.4.58) with all the properties in equations (6.4.59)-(6.4.65).

In the following we will also see that we can assume that hypotheses of Lemma 6.4.140 are satisfied. We introduce a further notation. We set

$$\delta_p := \gamma^{-1} C_{\vec{v},p}(\Pi_{\mathcal{A}} H). \tag{6.4.136}$$

Now we apply Lemma 6.4.140 to the field $F$ and we obtain the field $\hat{F} = (\mathcal{T}_+)_* F = (1+h_+)(\hat{N}_0^+ + \hat{N}^{(1)} + \hat{N}^{(2)} + \hat{H})$ in (6.4.71). We want to study the form of some set $\mathcal{O}' \in \mathcal{M}_{K,\mathfrak{r},\gamma}(\hat{F})$ with $\mathfrak{r} := \delta K^{-\kappa_1}$ (see definition 3.2.42). The parameters $K, \gamma$ are the same of Lemma 6.4.140. One can note that the conditions (3.2.79) and (3.2.80) on the operator $\mathcal{W}$ are equivalent to find an "approximate" solution $g \in \hat{\mathcal{B}}$ of the equation

$$\Pi_{K_+} \Pi_{\mathcal{A}}[g, N] = \Pi_{K_+} X, \qquad X \in \mathcal{A} \cap \mathcal{E}, \tag{6.4.137}$$

and where

$$N = (1+h_+)\Big(\hat{N}_0^+ + \hat{N}^{(1)} + \hat{N}^{(2)} + \Pi_{\mathcal{N}} \hat{H}\Big). \tag{6.4.138}$$

In Indeed by an explicit calculation equation 6.4.137 defining $\omega(\theta) := \hat{F}^{(\theta,0)}$ and $\Omega(\theta) := \hat{F}^{(w,w,)}[\cdot]$, becomes

$$\Pi_{K_+}\Big(\omega(\theta) \cdot \partial_\theta g^{(y)} + g^{(y,w)} \Omega(\theta) w\Big) \cdot \partial_y + \Pi_{K_+}\Big(\omega(\theta) \cdot \partial_\theta g^{(w,0)} - \Omega(\theta) g^{(w,0)}\Big) \cdot \partial_w = \Pi_{K_+} X. \tag{6.4.139}$$

We recall that $\omega(\theta) := (1+h_+)(\omega_+ + \hat{H}^{(\theta,0)})$ and $\Omega(\theta) := (1+h_+)(c_+\Omega^{-1} + \hat{N}^{(1)} + \hat{N}^{(2)} + (\Pi_{\mathcal{N}}\hat{H})^{(w)})$. By construction one of the effect of the map $\mathcal{T}_+$ is that the size of the terms $\hat{H}^{(\theta,0)})$ and $\hat{N}^{(1)}$ is "much" smaller with respect to the size of $H^{(\theta,0)})$ and $N^{(1)}$ (see equations (6.4.75) and (6.4.77)). Hence we claim that in order to find an approximate solution of (6.4.139) we just look for a solution of

$$\begin{aligned}
&\omega_+ \cdot \partial_\theta g^{(y,0)}(\theta) = \Pi_{K_+} X^{(y,0)} \frac{1}{1 + \Pi_{K_+} h_+}, \\
&\omega_+ \cdot \partial_\theta g^{(y,y)}(\theta) y = \Pi_{K_+} X^{(y,y)} \frac{1}{1 + \Pi_{K_+} h_+} y, \\
&\omega_+ \cdot \partial_\theta g^{(w,0)}(\theta) - \widetilde{\Omega}(\theta) g^{(w,0)} = \Pi_{K_+} X^{(w,0)} \frac{1}{1 + \Pi_{K_+} h_+}, \\
&\omega_+ \cdot \partial_\theta g^{(w,w)}(\theta) w + g^{(w,w)}(\theta) \widetilde{\Omega}(\theta) w = \Pi_{K_+} X^{(w,w)} w \frac{1}{1 + \Pi_{K_+} h_+},
\end{aligned} \tag{6.4.140}$$

with

$$\Omega(\theta) := (1+h_+)(c_+\Omega^{-1} + \hat{N}^{(2)} + (\Pi_{\mathcal{N}}\hat{H})^{(w)}), \tag{6.4.141}$$

To solve the first two equations it is enough to ask that

$$|\omega_+ \cdot k| \geq \frac{\gamma}{\langle k \rangle^\tau}, \quad k \in \mathbb{Z}^d, \ |k| \leq K_+. \tag{6.4.142}$$

In other word thanks to the bound (6.4.187) we are able to estimate the operator $W_0 := (\omega_+ \cdot \partial_\theta)^{-1}$. Moreover thanks to Remark B.176 in Appendix B one can see that if one is able to solve the two equations

$$\mathcal{L}_\pm u = \Big(\omega_+ \cdot \partial_\theta \pm \widetilde{\Omega}(\theta)\Big) u = f, \tag{6.4.143}$$

for $u = u(\theta, x), f = f(\theta, x)$ maps

$$u, f : \mathbb{T}^d_s \to \Pi^\perp_S \mathbf{h}^{a,p}_{\mathrm{odd}}$$

then one can solve the third and the fourth equation in (6.4.189). To conserve some coherence of notation with [31] we rename the functional spaces one which we work in the following way.

**Definition 6.4.141.** *Consider the spaces $X, Y, Z$ in* (2.2.23). *We define*

$$\mathbf{G} := \Big\{ \mathbf{u} := (u, \bar{u}) \ : \ u \in G \ \text{ s.t. } \ u = \sum_{l \in \mathbb{Z}^d, j \in S^c} u_j(l) e^{\mathrm{i} l \cdot \theta} e^{\mathrm{i} j \cdot x} \Big\}, \tag{6.4.144}$$

*for $G = X, Y, Z$ of $G = H^s(\mathbb{T}^d_s \times \mathbb{T} : \mathbb{C})$ endowed with the norm $\| \cdot \|_{s,a,p}$ defined in*

Hence what we want to study is the invertibility of the operators $\mathcal{L}_\pm$. We analyze the operator $\mathcal{L}_+ =: \mathcal{L}$ but the analysis of $\mathcal{L}_-$ is completely similar. Explicitly we have that

$$\mathcal{L} := \Pi^\perp_S \omega_+ \cdot \partial_\theta + \Pi^\perp_S \left( -\mathrm{i}E \begin{pmatrix} c_+ & 0 \\ 0 & c_+ \end{pmatrix} \partial_{xx} - \mathrm{i}E \begin{pmatrix} a_1 & b_1 \\ -\bar{b}_1 & \bar{a}_1 \end{pmatrix} \partial_x - \mathrm{i}E \begin{pmatrix} a_0 & b_0 \\ -\bar{b}_0 & \bar{a}_0 \end{pmatrix} \right) \Pi^\perp_S + \hat{\mathscr{K}}, \tag{6.4.145}$$

where we rename $\hat{a}_i, \hat{b}_i \rightsquigarrow a_i, b_i$ the coefficients of $\hat{N}^{(2)}$. The inversion of $\mathcal{L}$ stands on two fundamental results. The first is the following:

**Lemma 6.4.142.** *Fix $\delta = \xi^{\frac{1}{4}}$ with $\xi \in \varepsilon^2 \Lambda$ (see* (6.2.10)*) and recall that $\gamma_0 := \boldsymbol{c}\xi$ (see* (6.3.28)*). Consider a reversible, tame linear operator $\mathcal{L}$ defined for $\xi \in \mathcal{O}_0$ of the form*

$$\begin{aligned} \mathcal{L} = \Pi^\perp_S \omega_+ \cdot \partial_\theta \mathbb{1} + \Pi^\perp_S \left( -\mathrm{i}E \begin{pmatrix} c_+ & 0 \\ 0 & c_+ \end{pmatrix} \partial_{xx} \right) \Pi^\perp_S + \hat{\mathscr{K}} \\ + \Pi^\perp_S \left( -\mathrm{i}E \begin{pmatrix} a_1 & b_1 \\ \bar{b}_1 & \bar{a}_1 \end{pmatrix} \partial_x + -\mathrm{i}E \begin{pmatrix} a_0 & b_0 \\ \bar{b}_0 & \bar{a}_0 \end{pmatrix} \right) \Pi^\perp_S, \end{aligned} \tag{6.4.146}$$

*with $|c_+ - 1| \le C\xi$, and, for $\xi \in \mathcal{O} \subseteq \mathcal{O}_0$*

$$a_i = a_i^{(0)} + a_i', \quad b_i := b_i^{(0)} + b_i', \ \ i = 0, 1, \tag{6.4.147}$$

*where the coefficients $a_i^{(0)}, b_i^{(0)}$ for $i = 0, 1$ are given by formulæ*(6.3.25) *while*

$$\begin{aligned} \|a_i'\|_{\vec{v}, \mathfrak{p}_1}, \|b_i'\|_{\vec{v}, \mathfrak{p}_1} \le C\xi\delta, \quad i = 0, 1, \\ \|c_i\|_{\vec{v}, \mathfrak{p}_1}, \|f_i\|_{\vec{v}, \mathfrak{p}_1} \le C\xi\delta, \quad i = 1, \ldots, N, \end{aligned} \tag{6.4.148}$$

*for some constant* $C = C(\mathfrak{p}_2, d) > 0$ *and where* $c_i, d_i$ *are the coefficients of* $\hat{\mathscr{K}}$*. Then there exists a tame and reversibility preserving map*

$$\mathcal{S} = \mathbb{1} + \Psi : \mathbf{H} \to \mathbf{H} \tag{6.4.149}$$

*with*

$$C_{\vec{v},\mathfrak{p}_1}(\Psi) \leq C\xi, \tag{6.4.150}$$

*such that the conjugated* $\mathcal{L}_+ := \mathcal{S}^{-1}\mathcal{L}\mathcal{S}$ *has the form*

$$\begin{aligned}\mathcal{L}_+ = \Pi_S^\perp \omega_+ \cdot \partial_\theta + \Pi_S^\perp \left( -\mathrm{i}E \begin{pmatrix} c_+ & 0 \\ 0 & c_+ \end{pmatrix} \partial_{xx} - \mathrm{i}E \begin{pmatrix} \mathrm{diag}_{j\in\mathbb{Z}_+} r_0^j & 0 \\ 0 & \mathrm{diag}_{j\in\mathbb{Z}_+} r_0^j \end{pmatrix} \right) \Pi_S^\perp + \hat{\mathscr{K}}^+ \\ + \Pi_S^\perp \left( -\mathrm{i}E \begin{pmatrix} 0 & b_1^+ \\ \bar{b}_1^+ & 0 \end{pmatrix} \partial_x + -\mathrm{i}E \begin{pmatrix} a_0^+ & b_0^+ \\ \bar{b}_0^+ & \bar{a}_0^+ \end{pmatrix} \right) \Pi_S^\perp,\end{aligned} \tag{6.4.151}$$

*where* $r_0 \in \mathbb{R}$ *is such that* $|r_0|_\gamma \leq C\xi$ *and the coefficients* $a_0^+, b_j^+$ *with* $j = 0, 1$*, and the coefficients* $c_i^+, d_i^+$ *for* $i = 1, \ldots, N$ *of* $\hat{\mathscr{K}}^+$ *satisfy the bound*

$$\|a_0^+\|_{\vec{v},\mathfrak{p}_1} \leq C\xi\delta. \tag{6.4.152}$$

*Proof.* We divide the proof into two steps.

**Step 1 - Descent Method.** In this step we want to eliminate the term $a_1 := a_1^{(0)} + a_1'$ in the operator of order $O(\delta_x)$. We follows the strategy used in Step 4 of Section 3 in [31]. We introduce a change of coordinates of the form

$$\mathcal{S}_1 := \mathbb{1} + \Psi_1 := \mathbb{1} + \Pi_S^\perp \begin{pmatrix} s(\theta, x) & 0 \\ 0 & \bar{s}(\theta, x) \end{pmatrix} \Pi_S^\perp \tag{6.4.153}$$

for a function $s$ small enough in such a way $S_1$ is invertible. By a direct calculation we have that the coefficients

$$\begin{aligned} a_1^{(1)} &:= 2c_+ \frac{s_x}{1+s} + a_1, \qquad a_0^{(1)} := \frac{-i(\omega_+ \cdot \partial_\varphi s) + c_+ s_{xx}}{1+s} + a_0, \\ b_1^{(1)} &:= b_1 \frac{1+\bar{s}}{1+s}, \qquad b_0^{(1)} := b_0 \frac{1+\bar{s}}{1+s}. \end{aligned} \tag{6.4.154}$$

We look for $s$ such that $a_1^{(1)} \equiv 0$. Recall that by the reversibility one has on $\mathcal{U}$ that $a_1$ has zero average in $x$. Hence, by setting $1 + s = \exp(q(\theta, x))$, one has that $a_1^{(1)} = 0$ becomes

$$\mathrm{Re}(q_x) = -\frac{1}{2c_+}\mathrm{Re}(a_1), \qquad \mathrm{Im}(q_x) = -\frac{1}{2c_+}\mathrm{Im}(a_1), \tag{6.4.155}$$

that have unique (with zero average in $x$) solution

$$\mathrm{Re}(q) = -\frac{1}{2c_+}\partial_x^{-1}\mathrm{Re}(a_1), \qquad \mathrm{Im}(q) = -\frac{1}{2c_+}\partial_x^{-1}\mathrm{Im}(a_1). \tag{6.4.156}$$

One has that the solution $q$ is satisfies the estimates

$$\begin{aligned} \|q\|_{\vec{v},p} &\leq C\|a_1\|_{\vec{v},p}, \\ \|q\|_{\vec{v},\mathfrak{p}_1} &\leq C\xi, \end{aligned} \tag{6.4.157}$$

where we used the estimate $|c_+ - 1| \leq C\xi$. Clearly the function $s$ satisfies the same estimates in (6.4.157). Hence we have obtained

$$\begin{aligned} \mathcal{L}_1 = \omega_+ \cdot \partial_\theta + \Pi_S^\perp \left( -\mathrm{i}E \begin{pmatrix} c_+ & 0 \\ 0 & c_+ \end{pmatrix} \partial_{xx} \right) \Pi_S^\perp + \hat{\mathscr{K}}^{(1)} \\ + \Pi_S^\perp \left( -\mathrm{i}E \begin{pmatrix} 0 & b_1^{(1)} \\ \bar{b}_1^{(1)} & 0 \end{pmatrix} \partial_x + -\mathrm{i}E \begin{pmatrix} a_0^{(1)} & b_0^{(1)} \\ \bar{b}_0^{(1)} & \bar{a}_0^{(1)} \end{pmatrix} \right) \Pi_S^\perp. \end{aligned} \tag{6.4.158}$$

Now since the transformation $\mathcal{S}_1 = \mathbb{1} + O(\xi)$ trivially (see Lemma (6.4.138)) one has again that

$$a_0^{(1)} = a_0^{(0)} + a_0'', \quad b_i^{(1)} := b_i^{(0)} + b_i'', \ \ i = 0, 1, \tag{6.4.159}$$

the coefficients $a_0^{(1)}, b_j^{(1)}$ with $j = 0, 1$, and the coefficients $c_i^{(1)}, d_i^{(1)}$ for $i = 1, \ldots, N$ of $\hat{\mathscr{K}}^{(1)}$ satisfy the bounds (6.4.148). Moreover by equation (6.4.156) one has that $q$ is even in $x$ and hence the transformation $\mathcal{S}$ does not change the parity of the coefficients, i.e. it is reversibility preserving.

**Step 2 - Linear Birkhoff Normal Form.** In this step we look for a reversibility preserving map

$$\widetilde{\mathcal{S}}_2 := \mathbb{1} + \Psi_1 := \mathbb{1} + \begin{pmatrix} (\Psi_1)_1^1 & (\Psi_1)_1^{-1} \\ (\Psi_1)_{-1}^1 & (\Psi_1)_{-1}^{-1} \end{pmatrix} : \Pi_S^\perp \mathbf{h}_{\mathrm{odd}}^{a,p} \to \Pi_S^\perp \mathbf{h}_{\mathrm{odd}}^{a,p}, \tag{6.4.160}$$

that eliminates the coefficients $a_0^{(0)}, b_i^{(0)}$ for $i = 0, 1$. First we write

$$\mathcal{L}_1 := \omega_+ \cdot \partial_\theta + \Pi_S^\perp \left( -\mathrm{i}E \begin{pmatrix} c_+ & 0 \\ 0 & c_+ \end{pmatrix} \partial_{xx} + B \right) \Pi_S^\perp + R, \tag{6.4.161}$$

where

$$\begin{aligned} B &:= \Pi_S^\perp \left( -\mathrm{i}E \begin{pmatrix} 0 & b_1^{(0)} \\ \bar{b}_1^{(0)} & 0 \end{pmatrix} \partial_x - \mathrm{i}E \begin{pmatrix} a_0^{(0)} & b_0^{(0)} \\ \bar{b}_0^{(0)} & \bar{a}_0^{(0)} \end{pmatrix} \right) \Pi_S^\perp, \\ R &:= \Pi_S^\perp \left( -\mathrm{i}E \begin{pmatrix} 0 & b_1'' \\ \bar{b}_1'' & 0 \end{pmatrix} \partial_x - \mathrm{i}E \begin{pmatrix} a_0'' & b_0'' \\ \bar{b}_0'' & \bar{a}_0'' \end{pmatrix} \right) \Pi_S^\perp + \hat{\mathscr{K}}^{(1)} \end{aligned} \tag{6.4.162}$$

We have that

$$\begin{aligned} \mathcal{L}_1 \widetilde{\mathcal{S}}_1 - \widetilde{\mathcal{S}}_1 \Pi_S^\perp (\omega_+ \cdot \partial_\theta - \mathrm{i}E c_+ \partial_{xx}) \Pi_S^\perp = \\ = \Pi_S^\perp \left[ \omega_+ \cdot \partial_\theta \mathbb{1} + [-\mathrm{i}E c_+ \partial_{xx}, \Psi_1] + B \right] \Pi_S^\perp + \widetilde{R} \end{aligned} \tag{6.4.163}$$

where

$$\widetilde{R} := B\Psi_1 + R(\mathbb{1} + \Psi). \tag{6.4.164}$$

We look for $\Psi_1$ such that

$$\omega_+ \cdot \partial_\theta \mathbb{1} + [-\mathrm{i}Ec_+\partial_{xx}, \Psi_1] + B = 0. \tag{6.4.165}$$

In Fourier space, using the exponential basis both in time and space, equation (6.4.165) reads

$$i\omega_+ \cdot l - i\sigma c_+(j^2 - (\sigma\sigma')k^2)(\Psi_1)_{\sigma,j}^{\sigma',k}(l) = -B_{\sigma,j}^{\sigma',k}(l), \quad l \in \mathbb{Z}^d, \;\; j,k \in S^c, \;\; \sigma,\sigma' = \pm 1. \tag{6.4.166}$$

Now by (6.4.162) we have that the opeator $B$ depends only on the terms defined in (6.3.25). Moreover by (6.2.9) and (6.2.12) we have that the function $v(\theta,x)$ has the form

$$\begin{gathered} v(\theta,x) := \sum_{i=1}^{d} \sqrt{\xi_i + y_i} e^{\mathrm{i}\ell(\mathtt{v}_i)\cdot\theta} \sin(\mathtt{v}_i x), \\ \ell : S_+ \in \mathbb{Z}^d, \quad \ell(\mathtt{v}_i) = e_i, \end{gathered} \tag{6.4.167}$$

where $e_i = (0\ldots,1,\ldots,0)$ is the $i-th$ vector of the canonical basis of $\mathbb{R}^d$. Definitions in (6.2.9) and (6.2.12) are given in the sine basis in space, since we deal with odd function of $x$. On the other hand in this case it is more convenient to use the exponential basis also in $x$. It is sufficient to change the definition in (6.2.9) by recalling that for $u$ odd in space then the Fourier coefficients (in space) has the property $u_j = -u_{-j}$. Hence $\mathbf{v} = (v^+, v^-)$ and $w = (z^+, z^-)$ are

$$\begin{aligned} v^\sigma &= \sum_{\mathtt{v}_i \in S} \mathrm{sign}(\mathtt{v}_\mathtt{i}) e^{\mathrm{i}\sigma\mathtt{v}_i\cdot x} \sqrt{\xi_{|\mathtt{v}_i|} + y_{|\mathtt{v}_i|}} e^{\mathrm{i}\sigma\ell(v_i)\cdot\theta}, \xi_{\mathtt{v}_i} = \xi_i, y_{\mathtt{v}_i} = y_i, \ell(-|\mathtt{v}_i|) = -\ell(|\mathtt{v}_i|), \\ z^\sigma &= \sum_{j\in S^c} u_j^\sigma e^{\mathrm{i}\sigma j\cdot x}, \;\; u_{-j}^\sigma = -u_j^\sigma \end{aligned} \tag{6.4.168}$$

that are equivalent to definitions in (6.2.9). With this formalism we have that

$$\begin{gathered} B_{\sigma,j}^{\sigma,k}(l) := -\mathrm{i}\sigma(2a_1 - j_1^2 a_2 - j_1 j_2 a_3 - a_5(j_1^2 j_2^2 + j_2^2))\sqrt{\xi_{j_1}}\sqrt{\xi_{j_2}}, \\ \text{for } \; j_1,\, j_2,\, j-k \in S, \quad j_1 - j_2 + k - j = 0, \;\; l = \ell(j_1) - \ell(j_2), \end{gathered} \tag{6.4.169}$$

and $B_{\sigma,j}^{\sigma,k}(l) = 0$ otherwise, and

$$\begin{gathered} B_{\sigma,j}^{-\sigma,k}(l) := -\mathrm{i}\sigma\left(a_1 - j_1^2 a_2 - (a_3 j_1 - a_4 j_1^2 j_2)(-\sigma)k\right), \\ \text{for } \; j_1,\, j_2,\, j+k \in S, \quad j_1 - j_2 + k - j = 0, \;\; l = \ell(j_1) - \ell(j_2), \end{gathered} \tag{6.4.170}$$

and $B_{\sigma,j}^{-\sigma,k}(l) = 0$ otherwise. We define the solution of equation (6.4.165) as

$$(\Psi_1)_{\sigma,j}^{\sigma',k}(l) := \begin{cases} \dfrac{B_{\sigma,j}^{\sigma',k}(l)}{i(\omega_+ \cdot l - \sigma c_+(j^2 - (\sigma\sigma')k^2))}, & \text{if } \; \omega^{-1}\cdot l + \sigma j^2 - \sigma' k^2 \neq 0, \\ 0 \;\; \text{otherwise} \end{cases} \tag{6.4.171}$$

Actually the operator $\Psi_1$ in (6.4.171) is the solution of

$$\omega_+ \cdot \partial_\theta \mathbb{1} + [-\mathrm{i}Ec_+\partial_{xx}, \Psi_1] + B = [B], \qquad [B] := \begin{cases} B^{\sigma,j}_{\sigma,j}(0), l = 0, j = k, \sigma, = \sigma' \\ 0 \;\text{ otherwise,} \end{cases} \tag{6.4.172}$$

indeed for such values of $l, j, k$ the denominators in (6.4.171) are zero. Now $\Psi_1$ is well posed and solves (6.4.172). Indeed by (6.4.169) and (6.4.170) we have that $B^{\sigma',k}_{\sigma,j}(\ell) = 0$ unless $|l| \leq 2$, hence since $l = \ell(j_1) - \ell(j_2)$ one has

$$|\omega^{-1} \cdot l + k^2 - j^2| = |j_1^2 - j_2^2 + k^2 - j^2| \geq 1, \text{ for } j \neq \pm k, j_1 \neq \pm j_2,$$

where one uses Lemma 6.1.123. Moreover for $\sigma = -\sigma'$ one has

$$|\omega^{-1} \cdot l + k^2 + j^2| \geq 1, \text{ for any } j, k, j_1, j_2.$$

By using that $\omega_+ = \omega^{-1} + O(\xi)$ and that $c_+ = 1 + O(\xi)$ we have that the denominators in (6.4.171) satisfy

$$|\omega_+ \cdot l - \sigma c_+(j^2 - (\sigma\sigma')k^2)| \geq |c_+|\omega^{-1} \cdot l + \sigma(j^2 - k^2)| - |l||\omega_+ - c_+\omega^{-1}| \geq \frac{1}{2}, \tag{6.4.173}$$

for $\xi$ small enough.

**Lemma 6.4.143.** *Set $d^{\sigma',k}_{\sigma,j}(l) = (\omega_+ \cdot l - \sigma c_+(j^2 - (\sigma\sigma')k^2))$. One has that*

$$|d^{\sigma',k}_{\sigma,j}(l)| \geq \begin{cases} C(j^2 + k^2), & \sigma = -\sigma', \\ C(|j| + |k|), & \text{if } \; \sigma = \sigma' \; j \neq k, \\ C, \quad \text{if } \; \sigma = \sigma' \; j = \pm k, \; l \neq 0 \end{cases} \tag{6.4.174}$$

*Proof.* If one assume $j^2 + k^2 > \widetilde{C} > 0$ then, since $|\omega_+| \leq |\omega^{-1}| + 1$, one has

$$|d^{-\sigma,k}_{\sigma,j}(l)| \geq \frac{1}{4}|j^2 + k^2| - 2|\omega_+| \geq \frac{1}{8}(j^2 + k^2). \tag{6.4.175}$$

If $j^2 + k^2 \leq \widetilde{C}$ but $j - k \in S$ then one can use equation (6.4.173) to obtain the result. Finally if $j - k \in S^c$ then one has $B^{-\sigma,k}_{\sigma,j}(l) = 0$. The second bound is obtained following the same reasoning and using the fact that $|j^2 - k^2| = |(j-k)(j+k)| \geq |j| + |k|$. The last bound is trivial. □

The following properties is a consequence of Lemma 6.4.2.

**Lemma 6.4.144.** *Let us define the operator $A := \Psi_1 - \{\Psi_1\}$ where $\{\Psi_1\}^{\sigma',k}_{\sigma,j}(l) = \Psi^{\sigma,j}_{\sigma,j}(l)$ for $\sigma = \sigma'$, $j = k$ and $l \neq 0$. Then one has that $|A\partial_x|_{\vec{v},p} + |\partial_x A|_{\vec{v},p} \leq C(p)\xi$ where $|\cdot|^{\mathrm{dec}}_{\vec{v},p}$ is the decay norm extended obtained by extending the norm defined in* (4.3.86) *with $j \in \mathbb{Z}$ instead of $\mathbb{Z}_+$.*

*Proof.* One has that

$$|\Psi_1\partial_x|_{s,a,p} = \sup_{\sigma,\sigma'=\pm}\sum_{h\in\mathbb{Z},l\in\mathbb{Z}^d}\langle j,l\rangle^{2p}e^{2|l|s}e^{2|j|a}\sup_{j-k=k}|(\Psi_1)^{\sigma',k}_{\sigma,j}(l)k|^2 \le C(p)\xi \tag{6.4.176}$$

since $(\Psi_1)^{\sigma',k}_{\sigma,j}(l)=0$ outside the set $|l|\le 2$ and $|j-k|\le C_S$ and the decay norm of $B$ is controlled by the norm of its coefficients $a_0^{(0)}, b_1^{(0)}, b_0^{(0)}$. In particular note that we used Lemma 6.4.143 in the following way. For instance we have the estimate

$$|(\Psi_1)^{-\sigma,k}_{\sigma,j}(l)k| \le C\frac{1}{(j^2+k^2)}|B^{-\sigma,k}_{\sigma,j}(l)k| \tag{6.4.177}$$

and one uses the gain of two derivatives of the denominator to control the two derivatives in the numerator. Hence one control the coefficients using $\|b_1^{(0)}\|_{\vec{v},p}+\|b_0^{(0)}\|_{\vec{v},p}$ The bounds second term and the Lipschitz estimate follows in the same way. □

By Lemma 6.4.2 follows that for $\xi$ small the map $\widetilde{\mathcal{S}}_1$ is invertible. Moreover we have the following Lemma

**Lemma 6.4.145.** *Consider a linear operator $A=(A)^{\sigma'}_{\sigma}$ for $\sigma,\sigma'=\pm1$ on the spaces $\mathbf{G}:=G\times G$ where $G=X,Y,Z$ the spaces defined in* (4.1.47)*. One has that $A$ is reversibility preserving if and only if for any $\sigma,\sigma'=\pm1$, $l\in\mathbb{Z}^d$, $j,k\in\mathbb{Z}$*

$$A^{\sigma',k}_{\sigma,j}(l)=\overline{A^{\sigma',k}_{\sigma,j}(l)},\quad A^{\sigma',-k}_{\sigma,-j}(l)=A^{\sigma',k}_{\sigma,j}(l),\quad \overline{A^{\sigma',k}_{\sigma,j}(-l)}=A^{-\sigma',k}_{-\sigma,j}(l). \tag{6.4.178}$$

*An operator $B$ is reversible, i.e. maps $\mathbf{X}\to\mathbf{Z}$ if and only if*

$$B^{\sigma',k}_{\sigma,j}(l)=-\overline{B^{\sigma',k}_{\sigma,j}(l)},\quad B^{\sigma',-k}_{\sigma,-j}(l)=B^{\sigma',k}_{\sigma,j}(l),\quad \overline{B^{\sigma',k}_{\sigma,j}(-l)}=B^{-\sigma',k}_{-\sigma,j}(l). \tag{6.4.179}$$

The proof of Lemma 6.4.145 is similar to the proof of Lemma (4.36) in [31]. Clearly in this case the differences stands in the fact that we developed in Fourier coefficients using the exponential basis in $x$. By this Lemma and an explicit computation, we have that $\Psi_1$ is reversibility preserving since $B$ is reversible. Now we can define the map

$$\mathcal{S}_1 := \exp(\Psi_1), \tag{6.4.180}$$

the time$-1$ flow of the field $\Psi_1$. Clearly again $\mathcal{S}_1-\mathbb{1}=O(\xi)$. Hence by equation (6.4.172) we obtain

$$\mathcal{L}_+ := \mathcal{S}_1^{-1}\mathcal{L}_1\mathcal{S}_1 = \Pi_S^{\perp}(\mathrm{i}E\Omega_+^{-1}(\xi)+R_+)\Pi_S^{\perp}$$

with

$$\Omega_+^{-1} := \mathrm{diag}_{j\in\mathbb{Z}_+}\left(\begin{pmatrix} c_+j^2 & 0\\ 0 & c_+j^2\end{pmatrix}+\begin{pmatrix} B^{1,j}_{1,j}(0) & B^{1,-j}_{1,j}(0)\\ B^{1,j}_{1,-j}(0) & B^{1,-j}_{1,-j}(0)\end{pmatrix}\right) \tag{6.4.181}$$

and $R_+$ as the form (6.4.151). Note that we have defined $\Omega_+^{-1}$ as infinite dimensional matrix with index $\ell \in \mathbb{Z}^d$ and $j \in \mathbb{Z}_+$. It is an operator one the space of sequences $\{z_j\}_{j\in\mathbb{Z}}$. But by our condition of reversibility we work on sequences such that $z_j = -z_{-j}$. Hence we can rewrite the matrix $\Omega_+^{-1}$ as an operator acting on the space of "odd" sequences as a diagonal matrix

$$\Omega_+^{-1} := \mathrm{diag}_{j\in\mathbb{Z}_+}\left(c_+ j^2 + r_0^j\right), \qquad r_0^j := B_{1,j}^{1,j}(0) - B_{1,j}^{1,-j}(0), \tag{6.4.182}$$

and $r_0^j$ is real by the reversibility of the field $B$. Hence the Lemma is proved. □

**Remark 6.4.146.** *The terms $r_0^j$ are of order $O(\xi)$. In particular they are the integrable terms that cannot be cancelled through a Birkhoff transformation. Moreover such terms are the corrections of order $O(\xi)$ to $j^2$ that we have considered in* (6.6.275) *of Section 6.6.*

The following Lemma is the last important abstract result we will use in order to invert the operator of the type $\mathcal{L}$ in (6.4.145).

**Lemma 6.4.147.** *Fix $\delta = \xi^{\frac{1}{4}}$ with $\xi \in \varepsilon^2\Lambda := \mathcal{O}_0$ (see* (6.2.10)*) and recall that $\gamma_0 := \boldsymbol{c}\xi$ (see* (6.3.28)*). Consider a reversible, tame linear operator $\mathcal{L}$ defined for $\xi \in \mathcal{O}_0$ of the form*

$$\mathcal{L} = \Pi_S^{\perp}(\omega_+ \cdot \partial_\theta + \mathcal{D} + \mathcal{R})\Pi_S^{\perp} \tag{6.4.183}$$

*where*

$$\begin{aligned} \mathcal{D} &:= -\mathrm{i}E\,\mathrm{diag}_{j\in\mathbb{Z}_+\cap S^c}\Big(c_+ j^2 + r_0^j\Big), \\ \mathcal{R} &:= E_1 D + E_0 = E_1(\mathcal{L})D + E_0(\mathcal{L}) \end{aligned} \tag{6.4.184}$$

*with $D := \mathrm{diag}_{j\in\mathbb{Z}_+\cap S^c}\{j\}$, and where, if we write $k = (\sigma, j, p) \in \boldsymbol{C} \times \mathbb{N} \times \mathbb{Z}^d$,with $\boldsymbol{C} := \{+1, -1\}$ $q = 0, 1$,*

$$\begin{aligned} E_q = \Big((E_q)_k^{k'}\Big)_{k,k'\in\boldsymbol{C}\times\mathbb{N}\times\mathbb{Z}^d} &= \Big((E_q)_{\sigma,j}^{\sigma',j'}(l)\Big)_{k,k'\in\boldsymbol{C}\times\mathbb{N}\times\mathbb{Z}^d}, \\ (E_1)_{\sigma,j}^{\sigma,j'}(l) \equiv 0, &\quad \forall\, j, j' \in \mathbb{Z}_+ \cap S^c,\ \ l \in \mathbb{Z}^d. \end{aligned} \tag{6.4.185}$$

*Assume that $|c_+ - 1| \le C\xi$, and $|r_+^j| \le C\xi$. Fix parameters*

$$\kappa_4 = 7\tau + 3, \quad \kappa_5 = 7\tau + 5, \tag{6.4.186}$$

*and take an arbitrary $N > 0$ large. Assume that*

$$\begin{aligned} &|E_1|_{\vec{v},\mathfrak{p}_1} + |E_0|_{\vec{v},\mathfrak{p}_1} \le C\xi\delta, \qquad \mathfrak{p}_0 + \kappa_5 \le \mathfrak{p}_1 \le \mathfrak{p}_2, \\ &|E_1|_{\vec{v},\mathfrak{p}_2} + |E_0|_{\vec{v},\mathfrak{p}_2} \le C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{N}^\perp} G_n) \end{aligned} \tag{6.4.187}$$

*There exists a constant* $C_0 = C_0(\mathfrak{p}_2, d) > 0$ *such that, if*

$$K_0^{C_0}\gamma^{-1}C\xi\delta \leq \epsilon \tag{6.4.188}$$

*and* $\epsilon = \epsilon(d, \mathfrak{p}_2)$ *is small enough then the following holds. There exists a sequence of purely imaginary numbers*

$$\mu^N_{\sigma,j}(\xi) := -\mathrm{i}\sigma(c_+ j^2 + r^N_j), \quad \sigma = \pm 1,\ j \in \mathbb{Z}_+ \cap S^c, \tag{6.4.189}$$

*with*

$$|r^N_j| \leq C\xi, \quad |r^N_j - r^j_0| \leq C\xi\delta$$

*defined on* $\mathcal{O}_0$ *and such that for any* $\xi \in \Lambda^{2\gamma}_N$, *defined as*

$$\Lambda^{2\gamma}_N := \left\{ \begin{array}{ll} \xi \in \mathcal{O}: & |\omega_+ \cdot l + \mu^N_{\sigma,j}(\xi) - \mu^N_{\sigma',j'}(\xi)| \geq \frac{2\gamma|\sigma j^2 - \sigma' j'^2|}{\langle l \rangle^\tau} \\ & \forall l \in \mathbb{Z}^d,\ |l| \leq N \ \ \forall (\sigma, j), (\sigma', j') \in \boldsymbol{C} \times (\mathbb{Z}_+ \cap S^c) \end{array} \right\}, \tag{6.4.190}$$

*there exists a bounded, reversibility preserving, linear operator* $\Phi_N = \Phi_N(\xi)$ *such that*

$$\mathcal{L}_N := \Phi_N^{-1} \circ \mathcal{L} \circ \Phi_N := \omega_+ \cdot \partial_\theta + \mathcal{D}_N + \mathcal{R}_N, \tag{6.4.191}$$

*where*

$$\begin{aligned} \mathcal{D}_N &:= \mathrm{diag}_{\sigma \in \boldsymbol{C}, j \in \mathbb{Z}_+}(\mu^N_{\sigma,j}), \\ \mathcal{R}_N &: = E^N_1 D + E^N_0, \end{aligned} \tag{6.4.192}$$

$$\begin{aligned} |E^N_1|_{\vec{v}_1,p} + |E^N_0|_{\vec{v}_1,p} &\leq \Big(|E_1|_{\vec{v},p+\kappa_5} + |E_0|_{\vec{v},p+\kappa_5}\Big) N^{-\kappa_4}, \quad \vec{v}_1 := (\gamma, \Lambda^{2\gamma}_N, s, a), \\ |E^N_1|_{\vec{v}_1,p+\kappa_5} + |E^N_0|_{\vec{v}_1,p+\kappa_5} &\Big(|E_1|_{\vec{v},p+\kappa_5} + |E_0|_{\vec{v},p+\kappa_5}\Big) N, \quad \mathfrak{p}_0 \leq p \leq \mathfrak{p}_2 - \kappa_5. \end{aligned} \tag{6.4.193}$$

*Moreover one has that*

$$|\Phi^{\pm 1}_N - \mathbb{1}|_{\vec{v}_1,p} \leq \gamma^{-1}\Big(|E_1|_{\vec{v},p} + |E_0|_{\vec{v},p}\Big). \tag{6.4.194}$$

Before giving the proof of the Lemma we make some remarks. This Lemma essentially can be applied to operators $\mathcal{L}$ of the form (6.4.151). Indeed our strategy is to use Lemma 6.4.142 as a preliminary step before using a *KAM*-like scheme in order to diagonalized the linear operator $\mathcal{L}$. Lemma 6.4.147 provides an approximate diagonalization, but anyway the order of the approximation $N$ is arbitrary large. The conditions on the parameters in (6.4.190) are the Second order Melnikov conditions. Clearly such conditions depends on $N$ (see formula (6.4.190)). In particular to obtain a partial diagonalization one can ask for the conditions (6.4.190) only for $|l| \leq N$. On the contrary in order to completely diagonalize one has to ask the the lower bounds in (6.4.190) for any $l \in \mathbb{Z}^d$. Our choice is less restrictive but it is sufficient we are just looking for an approximate inverse of $\mathcal{L}$. Lemma (6.4.147) is the equivalent result of Theorem 4.27 in Section 4 of [31]. The proof of the Lemma above is based on the following Iterative Lemma.

Take $\mathcal{L}$ as in (6.4.183) and define

$$\varepsilon_p^0 := |E_1^0|_{\vec{v},p} + |E_0^0|_{\vec{v},p}, \quad \text{for} \quad p \geq 0. \tag{6.4.195}$$

**Lemma 6.4.148 (KAM iteration).** *There exist constant $C_0 > 0$, $K_0 \in \mathbb{N}$ large, such that if*

$$K_0^{C_0} \gamma^{-1} \varepsilon^0_{\mathfrak{p}_0+\kappa_5} \leq 1, \tag{6.4.196}$$

*then, for any $\nu \geq 0$, one has:*
$(\mathbf{S1}.)_\nu$ *Set $\Lambda_0^\gamma := \mathcal{O}_0$ and for $\nu \geq 1$*

$$\Lambda_\nu^\gamma := \left\{ \begin{array}{ll} \xi \in \Lambda_{\nu-1}^\gamma : & |\omega \cdot \ell + \mu_{\sigma,j}^{\nu-1}(\xi) - \mu_{\sigma',j'}^{\nu-1}(\xi)| \geq \frac{\gamma|\sigma j^2 - \sigma' j'^2|}{\langle \ell \rangle^\tau}, \\ & \forall |\ell| \leq K_{\nu-1}, (\sigma,j), (\sigma',j') \in \boldsymbol{C} \times \mathbb{N} \end{array} \right\}, \tag{6.4.197}$$

*For any $\xi \in \Lambda_\nu^\gamma = \Lambda_\nu^\gamma(\mathcal{L})$, there exists an invertible map $\Phi_{\nu-1}$ of the form $\Phi_{-1} = \mathbb{1}$ and for $\nu \geq 1$, $\Phi_{\nu-1} := \mathbb{1} + \Psi_{\nu-1} : \mathbf{H}^s \to \mathbf{H}^s$, with the following properties.*

*The maps $\Phi_{\nu-1}$, $\Phi_{\nu-1}^{-1}$ are reversibility-preserving according to Definition 4.1.47, moreover $\Psi_{\nu-1}$ is Töplitz in time, $\Psi_{\nu-1} := \Psi_{\nu-1}(\varphi)$ (see (4.3.98)) and satisfies the bounds:*

$$|\Psi_{\nu-1}|_{\vec{v}_\nu,p} \leq \varepsilon^0_{p+\kappa_5} K_{\nu-1}^{2\tau+1} K_{\nu-2}^{-\alpha}, \quad \vec{v}_\nu := (\gamma, \Lambda_\nu^\gamma, s, a). \tag{6.4.198}$$

*Setting, for $\nu \geq 1$, $\mathcal{L}_\nu := \Phi_{\nu-1}^{-1} \mathcal{L}_{\nu-1} \Phi_{\nu-1}$, we have:*

$$\begin{aligned} &\mathcal{L}_\nu = \omega \cdot \partial_\varphi \mathbb{1} + \mathcal{D}_\nu + \mathcal{R}_\nu, \qquad \mathcal{D}_\nu = \mathrm{diag}_{(\sigma,j)\in \boldsymbol{C}\times\mathbb{N}}\{\mu_{\sigma,j}^\nu\}, \\ &\mu_{\sigma,j}^\nu = \mu_{\sigma,j}^0(\xi) + r_{\sigma,j}^\nu(\xi), \quad \mu_{\sigma,j}^0(0) = -\sigma \mathrm{i}(c_+ j^2 + r_0^j) \end{aligned} \tag{6.4.199}$$

*and*

$$\mathcal{R}_\nu = E_1^\nu(\xi) D + E_0^\nu(\xi), \tag{6.4.200}$$

*where $\mathcal{R}_\nu$ is reversible and the matrices $E_q^\nu$ satisfy (6.4.185) for $q = 1, 2$. For $\nu \geq 0$ one has $r_{\sigma,j}^\nu \in i\mathbb{R}$, $r_{\sigma,j}^\nu = -r_{-\sigma,j}^\nu$ and the following bound holds:*

$$|r_{\sigma,j}^\nu|_\gamma := |r_{\sigma,j}^\nu|_{\Lambda_\nu^\gamma,\gamma} \leq \xi \delta C. \tag{6.4.201}$$

*Finally, if we define*

$$\varepsilon_p^\nu := |E_1^\nu|_{\vec{v}_\nu,p} + |E_0^\nu|_{\vec{v}_\nu,p}, \quad \forall p \geq 0, \tag{6.4.202}$$

*one has* $\forall\, p \in [\mathfrak{s}_0, \mathfrak{p}_2 - \kappa_5]$ *($\kappa_5$ is defined in* (6.4.186)*) and* $\nu \geq 0$

$$\begin{aligned} \varepsilon_p^\nu &\leq \varepsilon^0_{p+\kappa_5} K_{\nu-1}^{-\kappa_4}, \\ \varepsilon^\nu_{p+\kappa_5} &\leq \varepsilon^0_{p+\kappa_5} K_{\nu-1}. \end{aligned} \tag{6.4.203}$$

$(\mathbf{S2})_\nu$ *For all* $j \in \mathbb{N}$ *there exists Lipschitz extensions* $\widetilde{\mu}^\nu_h(\cdot) : \mathcal{O}_0 \to \mathbb{R}$ *of* $\mu^\nu_h(\cdot) : \Lambda^\gamma_\nu \to \mathbb{R}$*, such that for* $\nu \geq 1$*,*

$$|\widetilde{\mu}^\nu_{\sigma,j} - \widetilde{\mu}^{\nu-1}_{\sigma,j}|_\gamma \leq \varepsilon^{\nu-1}_{\mathfrak{s}_0}, \qquad \forall\, k \in \boldsymbol{C} \times \mathbb{N}. \tag{6.4.204}$$

$(\mathbf{S3})_\nu$ *Let* $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *as in* (6.4.183)*, defined on* $\mathcal{O}_0$ *such that* (6.4.188) *and* (6.4.196) *hold. Then for* $\nu \geq 0$*, for any* $\xi \in \Lambda^{\gamma_1}_\nu(\mathcal{L}_1) \cap \Lambda^{\gamma_2}_\nu(\mathcal{L}_2)$*, with* $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$*, one has*

$$|E^\nu_0(\mathcal{L}_1) - E^\nu_0(\mathcal{L}_2)|_{\vec{v},\mathfrak{p}_0} \leq N^{-\kappa_5}_{\nu-1} |E^0_0(\mathcal{L}_1) - E^0_0(\mathcal{L}_2)|_{\vec{v},\mathfrak{p}_0}, \tag{6.4.205a}$$

$$|E^\nu_0(\mathcal{L}_1) - E^\nu_0(\mathcal{L}_2)|_{\vec{v},\mathfrak{p}_0+\kappa_5} \leq N_{\nu-1} |E^0_0(\mathcal{L}_1) - E^0_0(\mathcal{L}_2)|_{\vec{v},\mathfrak{p}_0+\kappa_5} \tag{6.4.205b}$$

*with* $\vec{v} := (\gamma, \Lambda^{\gamma_1}_\nu(\mathcal{L}_1) \cap \Lambda^{\gamma_2}_\nu(\mathcal{L}_2), s, a)$*, and moreover, for* $\nu \geq 1$*, for any* $p \in [\mathfrak{p}_0, \mathfrak{p}_0 + \kappa_5]$*, for any* $(\sigma, j) \in \boldsymbol{C} \times \mathbb{N}$ *and for any* $\xi \in \Lambda^{\gamma_1}_\nu(\mathcal{L}_1) \cap \Lambda^{\gamma_2}_\nu(\mathcal{L}_2)$*,*

$$|(r^\nu_{\sigma,j}(\mathcal{L}_2) - r^\nu_{\sigma,j}(\mathcal{L}_1)) - (r^{\nu-1}_{\sigma,j}(\mathcal{L}_2) - r^{\nu-1}_{s,j}(\mathcal{L}_1))| \leq |E^{\nu-1}_0(\mathcal{L}_1) - E^{\nu-1}_0(\mathcal{L}_2)|_{s,a,\mathfrak{p}_0}, \tag{6.4.206a}$$

$$|(r^\nu_{\sigma,j}(\mathcal{L}_2) - r^\nu_{\sigma,j}(\mathcal{L}_1))| \leq C |E^0_0(\mathcal{L}_1) - E^0_0(\mathcal{L}_2)|_{\vec{v},\mathfrak{p}_0}. \tag{6.4.206b}$$

$(\mathbf{S4})_\nu$ *Let* $\mathcal{L}_1, \mathcal{L}_2$ *be as in* $(\mathbf{S3})_\nu$ *and* $0 < \rho < \gamma/2$*. For any* $\nu \geq 0$ *one has*

$$C K^\tau_{\nu-1} \sup_{\xi \in \mathcal{O}_0} |E^0_0(\mathcal{L}_1) - E^0_0(\mathcal{L}_2)|_{\vec{v},\mathfrak{p}_0} \leq \rho \quad \Rightarrow \quad \Lambda^\gamma_\nu(\mathcal{L}_1) \subset \Lambda^{\gamma-\rho}_\nu(\mathcal{L}_2), \tag{6.4.207}$$

*Proof of Lemmata 6.4.147 and 6.4.148.* The proof is the same that the one of Theorem 4.3.60 in Section 4.3 which is based on the iterative scheme in Lemma 4.3.71. Here Lemma 6.4.147 follows from Lemma 6.4.148. The proof of Lemma 6.4.148 is similar to the one of Lemma 4.3.71. Indeed by hypothesis the operator $\mathcal{L}$ in (6.4.183) has the same class of operators defined in Definition 4.3.70 and moreover smallness condition in (6.4.187) is the equivalent of the smallness condition of $\gamma^{-1}\varepsilon$ in Theorem 4.3.60. One difference is that here the frequency $\omega_+$ deends on parameters $\xi \in \mathbb{R}^d$, while in Chapter 4 there is only one-dimensional parameters $\lambda \in \mathbb{R}$ that modulate $\omega$. Anyway there are no differences in the proof since Kirszbraun's Theorem on Lipschitz extension of functions holds in $\mathbb{R}^d$ (see Lemma A.2 in [50]). The proofs of items $(S3)_\nu, (S4)_\nu$ of Lemma 6.4.148 are the same of those of items $(S3)_\nu, (S4)_\nu$ of Lemma 4.3.71. The difference is in the fact that in Chapter 4 one considers the same linear operator $\mathcal{L}$ that is the linearized of the same non linear operator on two different points $u_1$ and $u_2$. Moreover the difference of $\mathcal{L}(u_1)$ and $\mathcal{L}(u_2)$ is given by the difference of $u_1$ and $u_2$. In other words the operators are close to each other if the points $u_i$ are close. Here one gives the estimates on the differences of $r^{\nu}_{\sigma,j}(\mathcal{L}_1)$ and $r^{\nu}_{\sigma,j}(\mathcal{L}_2)$ directly in terms of the differences of the two operators $\mathcal{L}_1$ and $\mathcal{L}_2$.

Another difference is that in Theorem 4.3.60 one get a complete diagonalization. This is obtained by applying infinitely many changes of coordinates that approximatively diagonalize as one can see in Lemma 4.3.60 in Section 4.3. In this case, to prove formula (6.4.191) it is enough to consider $\Phi_N$ the composition of a *finite* number of changes of coordinates. That is why the set of parameters in (6.4.190) is defined for $|l| \leq N$. The last difference is that here the sites $j \in S^c$ instead of $\mathbb{Z}_+$. This cannot create problem by Remark 6.4.133. □

**Remark 6.4.149. Approximate eigenvalues** *In Theorem 4.3.60 in Section 4.3 given an operator $\mathcal{L}$ one construct the eigenvalues $\mu^{\infty}_{\sigma,j}$ as limit of some "approximate" eigenvalues $\mu^{\nu}_{\sigma,j}$, for $\nu \geq 1$. Here we have that we stop the sequence of $\mu^{\nu}_{\sigma,j}$ after the number of steps one need in order to get the approximation of order $N$ in* (6.4.192),(6.4.193) *and the one defines $\mu^{N}_{\sigma,j}$ as the last term of such sequence. Moreover in Chapter 4 the operator $\mathcal{L}$ is the linearize operator of a field $F_0$ in some point $u$. Theorem 4.3.60 provides also Lipschitz dependence of the approximate eigenvalues $\mu^{\nu}_{\sigma,j}(u)$ with respect the point $u$. This is used in measure estimate in Section 4.5 Here the situation is different. As we will see the operator $\mathcal{L}$ comes form the linearization in zero of some vector field $F_1$. Hence while in Section 4.5 one has to control the difference between the eigenvalues of $\mathcal{L}(u_1)$ and $\mathcal{L}(u_2)$, i.e the linearized operator of $F_0$ in two different functions $u_1, u_2$, here we need to control the differences between the eigenvalues of the linearized operators of two different fields $F_1, F_2$. If one knows that $\mathcal{L}_1$ is "close" to $\mathcal{L}_2$ (clearly one has to explain the meaning of "close") then the bounds on the eigenvalues follows trivially.*

**Corollary 6.4.150.** *For $\mathbf{g} \in \mathbf{Z}$, consider the equation*

$$\mathcal{L}_N \mathbf{u} = \mathbf{g}. \tag{6.4.208}$$

*Let us define*

$$P^{2\gamma}_N(\mathbf{u}) := \left\{ \begin{array}{ll} \xi \in \mathcal{O}: & |\omega_+ \cdot l + \mu^N_{\sigma,j}(\xi)| \geq \frac{2\gamma j^2}{\langle l \rangle^{\tau}}, \\ & \forall \in \mathbb{Z}^d, \;\; |l| \leq N \;\; \forall\, (\sigma, j) \in \boldsymbol{C} \times \mathbb{Z}_+ \cap S^c \end{array} \right\}. \tag{6.4.209}$$

*If $\xi \in \Lambda^{2\gamma}_N \cap P^{2\gamma}_N$ (defined respectively in* (6.4.187) *and* (6.4.209)*), then there exists $\mathbf{h} = (h, \bar{h}) \in \mathbf{X}$ such that*

$$\begin{aligned} &\|\mathcal{L}_N \mathbf{h} - \mathbf{g}\|_{\vec{v},\mathfrak{p}_0} \leq C \|\mathbf{g}\|_{\vec{v},\mathfrak{p}_1} \delta N^{-\kappa_4}, \qquad v := (\gamma, \Lambda^{2\gamma}_N \cap P^{2\gamma}_N, a, s). \\ &\|\mathcal{L}_N \mathbf{h} - \mathbf{g}\|_{\vec{v},p} \leq \gamma^{-1} \Big( |E_1|^{\rm dec}_{\vec{v},p+\kappa_5} + |E_0|^{\rm dec}_{\vec{v},p+\kappa_5} \Big) N^{-\kappa_4} \|\mathbf{g}\|_{\vec{v},\mathfrak{p}_1} + \gamma^{-1} C \xi \delta N^{-\kappa_4} \|\mathbf{g}\|_{\vec{v},p+2\tau+1}, \end{aligned} \tag{6.4.210}$$

*Proof.* First of all we can define

$$\mathbf{h} := (\omega_+ \cdot \partial_\theta + \mathcal{D}_N)^{-1} \mathbf{g}, \tag{6.4.211}$$

since $\mathcal{D}_N$ is diagonal and hence it is trivial to define its inverse. Let us check estimate (6.4.210). Following the same strategy of Lemma 5.44 in [31] one get the bound

$$\|\mathbf{h}\|_{\vec{v},p} \overset{(6.4.209)}{\leq} \gamma^{-1}\|\mathbf{g}\|_{\vec{v},p+2\tau+1} \tag{6.4.212}$$

and that $\mathbf{h} \in \mathbf{X}$. Eq. (6.4.211) implies $\mathcal{L}_N\mathbf{h} - \mathbf{g} = \mathcal{R}_N\mathbf{h}$ and moreover one has

$$\|\mathcal{R}_N\mathbf{h}\|_{\vec{v},p} \overset{(4.3.94c)}{\leq} (|E_1^N|^{\rm dec}_{\vec{v},p} + |E_0^N|^{\rm dec}_{\vec{v},p})\|\mathbf{h}\|_{\vec{v},\mathfrak{p}_0} + (|E_1^N|^{\rm dec}_{\vec{v},\mathfrak{p}_0} + |E_0^N|^{\rm dec}_{\vec{v},\mathfrak{p}_0})\|\mathbf{h}\|_{\vec{v},p}. \tag{6.4.213}$$

By using (6.4.212), (6.4.193) and (6.4.187) we have that (4.3.94c) implies (6.4.210). □

## 6.5 The sets of "good" parameters

In this Section we conclude the proof of Theorem 1.2.5. In Sections 6.1 and 6.2 essentially we rewrite the (1.2.16) as a infinite dynamical system given by the vector field in (6.2.14). In this way we are allowed to apply Theorem (3.2.39) to the vector field $F_0$ defined in (6.2.14). The analysis performed in Section 6.3 guarantees that one can satisfies hypotheses (3.2.69) and (3.2.70) of the Abstract theorem. In order to apply Theorem (3.2.39) one need to identify the sequences of maps $\mathcal{T}_n$ with properties (3.2.65),(3.2.66) and (3.2.67). and give a more explicit formulation of the sets of "good" parameters defined in (3.2.72) in order to estimate the measure of such sets.

On the field $F_0$ we cannot apply directly Lemma (6.4.140) just because $N^{(1)}$ is not "small enough" and we are not able to prove that $\mathcal{T}$ is close to the identity. We overcome this problem using an algebraic arguments. We strictly follows the strategy of Lemma (6.4.140), we will underline the fundamental differences. Roughly speaking the aim of the following Lemma is to conjugate $F_0$ to a vector field for which the term $N^{(1)}$ has constant coefficients of order $O(\xi)$ plus terms of order at least $O(\xi^{\frac{3}{2}})$. Clearly a procedure like this cannot be iterated infinitely many times. We just perform it one time in order to fullfil the smallness hypothesis of Lemma 6.4.140. In such Lemma one reduces the size of the term in $N^{(1)}$ "quadratically". We have the following result.

**Lemma 6.5.151.** *Consider the field $F_0$ defined in* (6.3.26)*. Fix $K_1 = K_0^{\frac{3}{2}}$ and take $\rho_1$ of definition* (3.2.61)*. Assume*

$$\rho_1^{-1}\delta K_0^C \leq \epsilon \tag{6.5.214}$$

*for some $C$ depending only on $d, \tau$. Then, if $\epsilon$ is small enough, there exists a tame, reversibility preserving map*

$$\mathcal{T}_1 = \mathbb{1} + f : \mathcal{O}_0 \times D_{a_0+\rho_1 a_0, p+\nu}(s_0 - \rho_1 s_0, r - \rho_1 r_0) \to D_{a_0,p+\nu}(s_0, r_0), \tag{6.5.215}$$

*with*

$$C_{\vec{v},p}(f) \le C\gamma_0^{-1}K^{\eta}\max\{\|a_2\|_{\vec{v},p}, \|b_2\|_{\vec{v},p}, \|H^{(\theta,0)}\|_{\vec{v},p}, \}, \ \ p \le \mathfrak{p}_2 \tag{6.5.216}$$

*that satisfies* (3.2.65), (3.2.66) *and* (3.2.67) *with* $n = 0$ *and with* $\mathfrak{p}_1 := \mathfrak{p}_0 + \mu$, $\mathfrak{p}_2 = \mathfrak{p}_0 + \kappa_2$ *defined in* (3.2.61) *and* (3.2.63) *provided* $\mu \ge 2\mathfrak{p}_0 + 2\tau + 4$. *We set*

$$\hat{F}_0 := (\mathcal{T}_1)_* F_0 = N_0 + \hat{G} : D_{a_0 - 2\rho_1 a_0, p+\nu}(s_0 - 2\rho_1 s_0, r - 2\rho_1 r_0) \to V_{a_0 - 2\rho_1 a_0, p},$$

*and moreover* $\hat{F}$ *has the form*

$$\hat{F} := (1 + h_1)\Big(\hat{N}_0^1 + N_1^{(1)} + N_1^{(2)} + H_1\Big). \tag{6.5.217}$$

*On the set of* $\xi \in \mathcal{O}_0$ *such that* $|\widetilde{\omega} \cdot l| \ge \gamma/\langle l\rangle^{\tau}$ *for* $|l| \le K_1$ *one has the following. The function* $h_1$ *satisfies bounds* (6.4.72) *with* $\delta^{(3)}_{\mathfrak{p}_2} \rightsquigarrow \mathfrak{M}_0\xi$ *and* $\delta^{(2)}_{\mathfrak{p}_2} \rightsquigarrow \mathfrak{M}_0\xi$. *One has that* $\hat{N}_0^1 = \omega_1 \cdot \partial_\theta + \widetilde{\Omega}_1^{-1} w \cdot \partial_w$ *with* $\widetilde{\Omega}_1^{-1} = c_0\Omega^{-1}$, *and*

$$|\omega_1 - \widetilde{\omega}| \le \mathfrak{M}_0\xi^2, \quad |c_0 - 1| \le \mathfrak{M}_0\xi, \quad c_0 := \frac{1}{(2\pi)^{d+1}}\int_{\mathbb{T}^{d+1}} a_2 dx d\theta \tag{6.5.218}$$

*and that* $N_+^{(1)}$ *as the form* (6.4.61) *with coefficients* $a_2^1, b_2^1$, $N_1^{(2)}$ *has the form* (6.4.62) *with coefficients* $a_i^1, b_i^1$ *for* $i = 0, 1$, $\mathscr{K}_1$ *of the form* (6.2.22) *with the same* $N$ *of* $\mathscr{K}_0$. *On the field* $N_1^{(2)}$ *the bounds* (6.4.74) *hold with* $K_+ \rightsquigarrow K_1$, $C_p(N^{(2)}) \rightsquigarrow \mathfrak{M}_0\xi$ *and* $\delta^{(1)}_{\mathfrak{p}_2} \rightsquigarrow \gamma_0^{-1}\|H_0^{(\theta,0)}\|_{\vec{v},p} = \mathfrak{A}_0\delta$. *With the same convention also the bounds* (6.4.75), (6.4.76), (6.4.77) *and* (6.4.78) *hold.*

*Proof.* Recall that , using the notations of Lemma (6.4.140), for the field $F_0$ the function $h \equiv 0$ and the constant $c = 1$. Note that the definition of $\delta_p^{(1)}$ is the same of Lemma 6.4.140. It controls the norm of the $H_0^{(\theta,0)}$ divided by the size of the small divisor $\gamma_0$. The definition of $\delta_p^{(2)}$ has changed. Indeed in this case we have set $\delta_p^{(2)} \approx C_{\vec{v},p}(N_0^{(1)})$ without divide by $\gamma_0$. This is due to the fact that $\gamma_0^{-1}C_{\vec{v},\mathfrak{p}_1}(N_0^{(1)}) = \mathfrak{A}_0$ that is not small. In Lemma 6.4.140 we used the smallness of $\delta^{(2)}_{\mathfrak{p}_1}$ in order to prove that $\mathcal{T}$ is close to the identity. In this case to get the result we need to use different arguments. However we follows the same strategy used in Lemma 6.4.140 and we perform the same four steps of that Lemma. Concerning step 1 and step 2 we apply the same transformations defined exactly in the same way. In this case there is no small divisors in the equations that define transformation $\Phi_1$ and $\Phi_2$. Hence the same estimates of Lemma 6.4.140 hold with the convention $\delta^{(2)}_{\mathfrak{p}_2} \rightsquigarrow \mathfrak{M}_0\xi$ and $\delta^{(1)}_{\mathfrak{p}_2} \rightsquigarrow \gamma_0^{-1}\|H_0^{(\theta,0)}\|_{\vec{v},p} = \mathfrak{A}_0\delta$. The step 3 has to be analyzed more carefully. Indeed if one looks at the equation (6.4.107) for $\beta$ one sees that one has to control the inverse of the operator $\widetilde{\omega} \cdot \partial_\theta$. By using the diophantine condition (6.3.28), one get by (6.4.73) that

$$\|\beta\|_{\vec{v},\mathfrak{p}_1+\mathfrak{p}_0} \le K_1^{\mathfrak{p}_0+2\tau+1}\gamma_0^{-1}\|m_2\|_{\vec{v},\mathfrak{p}_1} \le K_1^{\mathfrak{p}_0+2\tau+1}\gamma_0^{-1}C_{\vec{v},\mathfrak{p}_1}(N_0^{(1)}) \le K_+^{\mathfrak{p}_0+2\tau+1}\mathfrak{A}_0,$$

that is not small. We need to estimates $\beta$ in a different way. We first analyze the form of the coefficient $a_2^{(2)}$ . By equation (6.4.93) we have

$$m_2(\theta) = \left(\frac{1}{2\pi}\int_{\mathbb{T}}(1+\widetilde{a}_2(\theta,x))^{-\frac{1}{2}}\right)^{-2} - 1, \tag{6.5.219}$$

where $\widetilde{a}_2 = \sqrt{(1+a_2)^2 - |b_2|^2} - 1$ , with $a_2$ and $b_2$ the coefficients of $N_0^{(1)}$. We have that we can write

$$\widetilde{a}_2 = a_2(\theta,x) + b(\theta,x), \qquad b(\theta,x) := \sqrt{(1+a_2)^2 - |b_2|^2} - (1+a_2) \tag{6.5.220}$$

and we note also that

$$\|b\|_{\vec{v},\mathfrak{p}_1} \leq (C_{\vec{v},\mathfrak{p}_1}(N_0^{(1)}))^2 \leq C\xi^2, \tag{6.5.221}$$

for some constant $C$. Moreover by an explicit computation we can write

$$\begin{aligned}
m_2 &:= \frac{1}{2\pi}\int_{\mathbb{T}}\widetilde{a}_2 dx + \frac{d}{2\pi(1-\int_{\mathbb{T}}\widetilde{a}_2 dx + \hat{c})},\\
d &:= 2\pi\hat{c} - \left(\int_{\mathbb{T}}\widetilde{a}_2\right)^2 + \left(\int_{\mathbb{T}}\widetilde{a}_2\right)\hat{c},\\
\hat{c} &:= \frac{1}{\pi}\int_{\mathbb{T}} c + \frac{1}{4\pi^2}\left(\int_{\mathbb{T}}\widetilde{a}_2\right)^2 - \frac{1}{2\pi}\left(\int_{\mathbb{T}}\widetilde{a}_2\right)\left(\int_{\mathbb{T}} c\right) + \frac{1}{4\pi}\left(\int_{\mathbb{T}} c\right)^2,\\
c &:= \frac{1}{2}\left((1+\widetilde{a}_2)^{-\frac{1}{2}} - 1 + \frac{1}{2}\widetilde{a}_2\right).
\end{aligned} \tag{6.5.222}$$

Clearly one has

$$\begin{aligned}
\|m_2 - \frac{1}{2\pi}\int_{\mathbb{T}}\widetilde{a}_2\|_{\vec{v}_0,p} &\leq C(p)\|a_2\|_{\vec{v}_0,\mathfrak{p}_1}\|a_2\|_{\vec{v}_0,p},\\
\|m_2 - \frac{1}{2\pi}\int_{\mathbb{T}}\widetilde{a}_2\|_{\vec{v}_0,\mathfrak{p}_1} &\leq C\xi^2.
\end{aligned} \tag{6.5.223}$$

Clearly using (6.5.220) one has

$$m_2 = \frac{1}{2\pi}\int_{\mathbb{T}} a_2 dx + s(\theta,x), \qquad \|s\|_{\vec{v},\mathfrak{p}_1} \overset{(6.5.223),(6.5.221)}{\leq} C\xi^2. \tag{6.5.224}$$

Roughly speaking this implies that in low norm $\mathfrak{p}_1$ one has $m_2 \approx a_2 + O(\xi^2)$. Now equation (6.4.93) becomes

$$\beta(\theta) := \frac{1}{1+\mathfrak{c}}(\widetilde{\omega}\cdot\partial_\theta)^{-1}(1+\Pi_{K_1}\left(\frac{1}{2\pi}\int_{\mathbb{T}} a_2 dx + s\right) - 1 - \mathfrak{c})(\theta), \quad \mathfrak{c} = \frac{1}{(2\pi)^{d+1}}\int_{\mathbb{T}^{d+1}} m_2(\theta,x)d\theta dx, \tag{6.5.225}$$

Now we have to estimate $\beta$. The critical term is obviously the term of $O(\xi)$ because one cannot use estimate (6.3.28) since $\gamma_0 \approx \xi$. One can use an algebraic arguments. First we recall that by (6.3.24) one has $\widetilde{\omega} = \omega^{-1} + \omega^{(0)}(\xi)$ with $\omega^{-1} = j^2,\ \ j \in S^+$. On the other hand the term of order $\xi$ of $\beta$ (6.5.225) depends only on the coefficients $a_2$ given in (6.3.25). Hence in formula (6.5.225) we need to estimate $\widetilde{\omega} \cdot k$ with $k \in (S^+)^d$ but with only two components different from zero and not for $k \in \mathbb{Z}^d$ as in (6.3.28). This implies, using Lemma 6.1.123, in the term $a_2 z_{xx} \cdot \partial_z$ there are only trivial resonances, and hence

$$\|\beta\|_{\vec{v}_0, p+\mathfrak{p}_0} \leq \|\Pi_{K_1} a_2\|_{\vec{v}_0, p+\mathfrak{p}_0}, \quad \|\beta\|_{\vec{v}_0, \mathfrak{p}_1} \leq C\xi, \tag{6.5.226}$$

for some constant $C$, that is a better estimates with respect to the one in (6.4.108). In this way we get that the transformation is $\xi-$close to the identity. Now the last step can be performed exactly as in Lemma 6.4.140 because there are no other differences and one can estimate the transformation $\Phi_4$ as done in (6.4.124) and (6.4.125). Thanks to the perturbative arguments in (6.5.222) and (6.5.220) one can fix

$$c_0 := \frac{1}{(2\pi)^{d+1}} \int_{\mathbb{T}^{d+1}} a_2 dx d\theta$$

and the Lemma follows.

**Remark 6.5.152.** *Note that the coefficient $c_0$ in* (6.5.218) *gives the correction of order $O(\xi j^2)$ to the eigenvalues $j^2$ as we will see in Section 6.6 (see equation 6.6.275). This term will remain the same at each step of our iteration since all the further correction will be of higher order in $\xi$.*

□

The main result of this Section is the following:

**Lemma 6.5.153.** *There exists a sequence of maps $\mathcal{T}_n$ $n \geq 1$ that satisfies* (3.2.65), (3.2.66) *and* (3.2.67), *such that the $n-th$ vector field $F_n$ is defined on $\mathcal{O}_0$ and on $\mathcal{O}_n$ in* (3.2.72) *satisfies bounds* (3.2.74). *Moreover $F_n$ is pseudo differential of the Schrödinger type (see Definition 6.2.127) and can be written in the form* (6.4.58)

$$F_n := (1 + h_n)\Big(N_0^{(n)} + N_n^{(1)} + N_n^{(2)} + H_n\Big) \tag{6.5.227}$$

*where $N_0^{(n)} = \omega_n \cdot \partial_\theta + \Omega_n^{-1} w \cdot \partial_w$ with $\Omega_n^{-1} = c_n \Omega^{-1}$, $\omega_n \in \mathbb{R}^d$ is diophantine and*

$$|\omega_n - \widetilde{\omega}|_\gamma \leq C\xi\delta, \quad |c_n - c_0|_\gamma \leq C\xi\delta, \tag{6.5.228}$$

*where $c_0$ is defined in* (6.5.218),

$$d_w H^{(w)}(u)[\cdot] : \Pi_S^\perp \mathbf{h}_{\text{odd}}^{a,p+1} \to \Pi_S^\perp \mathbf{h}_{\text{odd}}^{a,p}. \tag{6.5.229}$$

*In particular $N_n^{(1)}, N_n^{(2)}$ have the form* (6.4.61) *and* (6.4.62) *and the following estimates hold:*

$$C_{\vec{v}_n,p}(H_n) \leq C_{\vec{v}_n,p}(\Pi_{\mathcal{N}^\perp} G_n), \quad \|h_n\|_{\vec{v}_n,p},\ C_{\vec{v}_n,p}(N_n^{(2)}) \leq C_{\vec{v}_n,p}(G_n). \tag{6.5.230}$$

$$\gamma_n^{-1}\|H_n^{(\theta,0)}\|_{\vec{v}_n,\mathfrak{p}_1} \le \delta K_{n-1}^{-\kappa_1+\mu+4}, \quad C_{\vec{v},\mathfrak{p}_1}(N_n^{(1)}) \le \delta K_{n-1}^{-\kappa_1+\mu+4}. \tag{6.5.231}$$

*We also have that the coefficients $a_i^{(n)}, b_i^{(n)}, c_j^{(n)}, d_j^{(n)}$ of the field $N_n^{(2)}$ for $i = 0, 1$, $j = 1, \dots, N$ have the form*

$$a_i^{(n)} = a_i^{(0)} + \widetilde{a}_i^{(n)}, \quad b_i := b_i^{(0)} + \widetilde{b}_i^{(n)}, \;\; i = 0, 1, \tag{6.5.232}$$

*where the coefficients $a_i^{(0)}, b_i^{(0)}$ for $i = 0, 1$ are given by formulæ(6.3.25), while*

$$\|\widetilde{a}_i^{(n)}\|_{\vec{v},\mathfrak{p}_1}, \|\widetilde{b}_i^{(n)}\|_{\vec{v},\mathfrak{p}_1} \le C\xi\delta, \quad i = 0, 1, \tag{6.5.233}$$

*and*

$$\|c_i^{(n)}\|_{\vec{v},\mathfrak{p}_1}, \|d_i^{(n)}\|_{\vec{v},\mathfrak{p}_1} \le (1 + \delta K_{n-1}^{-1}) C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{N}^\perp}(G_0)). \tag{6.5.234}$$

*In particular one has that*

$$\begin{aligned} &\|a_i^{(n)} - \mathcal{T}_n a_i^{(n-1)}\|_{\vec{v},\mathfrak{p}_1}, \|b_i^{(n)} - \mathcal{T}_n b_i^{(n-1)}\|_{\vec{v},\mathfrak{p}_1} \le \delta K_{n-1}^{-\kappa_1+\mu+4}, \quad i = 0, 1, 2, \\ &\|c_i^{(n)} - \mathcal{T}_n c_i^{(n-1)}\|_{\vec{v},\mathfrak{p}_1}, \|d_i^{(n)} - \mathcal{T}_n d_i^{(n-1)}\|_{\vec{v},\mathfrak{p}_1} \le \delta K_{n-1}^{-\kappa_1+\mu+4}, \quad i = 1, \dots, N. \end{aligned} \tag{6.5.235}$$

*Proof.* We prove the Lemma by induction on $n$. If one assume that we already constructed the map $\mathcal{T}_n$ such that all the properties above are satisfied then we proceed as follows. First of all by (6.5.231) one note that hypotheses (6.4.66) of Lemma 6.4.140 are satisfied. Then we apply the Lemma to the field $F_n$. We set $\mathcal{T}_{n+1}$ the map given by the Lemma (6.5.215). It satisfies (3.2.65), (3.2.66) and (3.2.67) by (6.5.216). We set $\hat{F}_n := (\mathcal{T}_{n+1})_* F_n$ (see (6.4.70)) that has the form

$$\hat{F}_n := N_0 + \hat{G}_n = (1 + h_{n+1})\Big(\hat{N}_0^{(n)} + \hat{N}_n^{(1)} + \hat{N}_n^{(2)} + \hat{H}_n\Big) \tag{6.5.236}$$

that clearly has all the bounds (6.4.72),(6.4.73), (6.4.74), (6.4.75), (6.4.76), (6.4.77) and (6.4.78) hold and these bounds together the inductive hypotheses implies that the field $\hat{F}_n$ satifies bounds like (6.5.228)-(6.5.234) except for (6.5.231). Actually one proves better bounds on $\hat{N}_n^{(1)}, \hat{H}_n^{(\theta,0)}$.

Indeed we have $\hat{N}_0^{(n)} = \omega_{n+1} \cdot \partial_\theta + \hat{\Omega}_n^{-1} w \cdot \partial_w$ with $\hat{\Omega}_n^{-1} = c_{n+1}\Omega^{-1}$, $\omega_{n+1} \in \mathbb{R}^d$ is diophantine and

$$|\omega_{n+1} - \widetilde{\omega}|_\gamma \le C\xi\delta, \quad |c_{n+1} - 1|_\gamma \le C\xi, \tag{6.5.237}$$

and

$$\begin{aligned} C_{\vec{v},\mathfrak{p}_1}(\hat{N}_n^{(2)}) &\le (1 + K_n^{\mu_1}\delta K_{n-1}^{-\kappa_1})\mathfrak{A}_n \\ C_{\vec{v},\mathfrak{p}_2}(\hat{N}_n^{(2)}) &\le (1 + K_n^{\mu_1}\delta K_{n-1}^{-\kappa_1})\mathfrak{A}_0(K_{n-1}^{\kappa_1} + K_n^{\mu_1}\delta K_{n-1}^{\kappa_1}), \end{aligned} \tag{6.5.238}$$

$$\begin{aligned} &C_{\vec{v},\mathfrak{p}_1}(\hat{N}_n^{(1)}) \le (1 + K_n^{\mu_1}\delta K_{n-1}^{\kappa_1})\Big[K_n^{-(\mathfrak{p}_2-\mathfrak{p}_1-\mu_1)}\delta K_{n-1}^{\kappa_1} + K_n^{\mu_1} K_{n-1}^{-2\kappa_1}\delta^2\Big], \\ &C_{\vec{v},\mathfrak{p}_2}(\hat{N}_n^{(1)}) \le \delta\mathfrak{A}_0\Big(K_{n-1}^{\kappa_1} + K_n^{\mu_1} K_{n-1}^{\kappa_1}\Big), \end{aligned} \tag{6.5.239}$$

$$\begin{aligned} C_{\vec{v},\mathfrak{p}_1}(\hat{G}_n) &\leq (1+K_n^{\mu_1}\delta K_{n-1}^{-\kappa_1})\mathfrak{A}_n \\ C_{\vec{v},\mathfrak{p}_2}(\hat{G}_n) &\leq (1+K_n^{\mu_1}\delta K_{n-1}^{-\kappa_1})\mathfrak{A}_0(K_n^{\kappa_1}+K_n^{\mu_1}\delta), \end{aligned} \tag{6.5.240}$$

$$\|\hat{H}_n^{(\theta,0)}\|_{\vec{v}_2,\mathfrak{p}_1} \leq (1+K_n^{\mu_1}\delta K_{n-1}^{-\kappa_1})\Big[K_n^{-(\mathfrak{p}_2-\mathfrak{p}_1)}\delta K_{n-1}^{-\kappa_1} + K_n^{\mu_1}\delta^2 K_{n-1}^{-2\kappa_1}\Big], \tag{6.5.241}$$

while the finite rank operator $\hat{\mathscr{K}}_n$ in $\hat{N}_n^{(2)}$ has coefficients

$$\begin{aligned} \max_{i=1,\dots,N}\{\|\hat{c}_i^{(n)}\|_{\vec{v}_2,\mathfrak{p}_1}, \|\hat{d}_i^{(n)}\|_{\vec{v}_2,\mathfrak{p}_1}\} &\leq (1+K_n^{\mu_1}\delta K_{n-1}^{-\kappa_1}) \max_{i=1,\dots,N}\{\|c_i^{(n)}\|_{\vec{v},\mathfrak{p}_1}, \|d_i^{(n)}\|_{\vec{v},\mathfrak{p}_1}\} \\ \max_{i=1,\dots,N}\{\|\hat{c}_i^{(n)}\|_{\vec{v}_2,\mathfrak{p}_2}, \|\hat{d}_i^{(n)}\|_{\vec{v}_2,\mathfrak{p}_2}\} &\leq (1+K_n^{\mu_1}\delta K_{n-1}^{-\kappa_1}) \Bigg(\max_{i=1,\dots,N}\{\|c_i^{(n)}\|_{\vec{v},\mathfrak{p}_2}, \|d_i^{(n)}\|_{\vec{v},\mathfrak{p}_2}\} \\ &\quad + K_n^{\mu_1}\delta K_{n-1}^{\kappa_1} \max_{i=1,\dots,N}\{\|c_i^{(n)}\|_{\vec{v},\mathfrak{p}_1}, \|d_i^{(n)}\|_{\vec{v},\mathfrak{p}_1}\}\Bigg). \end{aligned} \tag{6.5.242}$$

Clearly one has that bounds in (6.5.230), (6.5.233) and (6.5.234) holds also for $\hat{F}_n$. The more critical conditions are those in (6.5.231). Using (3.2.62) by (6.5.239) and (6.5.241) one gets

$$C_{\vec{v},\mathfrak{p}_1}(\hat{N}_n^{(1)}) \leq \delta K_n^{-\kappa_1}, \qquad \|\hat{H}_n^{(\theta,0)}\|_{\vec{v},\mathfrak{p}_1} \leq \delta K_n^{-\kappa_1} \tag{6.5.243}$$

that are bounds even better than (6.5.231).

By the definition in Theorem 3.2.39, we have that the field $F_{n+1}$ is given by $F_{n+1} = (\Phi_{n+1})_*\hat{F}_n$, where the map $\Phi_{n+1}$ is generated by the field $g_{n+1}$ of Definition (3.2.37) used with $F \rightsquigarrow \hat{F}_n$. We have to show that the map $\Phi$ does not affect seriously the coefficients of $\hat{N}_n^{(1)}, N_n^{(2)}$ in such a way the estimates on $F_{n+1}$ remains essentially the same of those on $\hat{F}_n$. First we note that, by the form of the map $\Phi_{n+1}$ one has

$$F_{n+1} := (\Phi_{n+1})_*\hat{F}_n = (1+h_{n+1})\Big((\Phi_{n+1})_*(\hat{N}_0^{(n)}) + (\Phi_{n+1})_*(\hat{N}_n^{(1)} + \hat{N}_n^{(2)} + \hat{H}_n)\Big). \tag{6.5.244}$$

Let us study first the term that does not contains the constant coefficients term $\hat{N}_0^{(n)}$. Again by the form of the map $\Phi_{n+1}$, that is generated by $g_{n+1} \in \hat{\mathcal{B}}$, we have that $\Phi_{n+1}$ preserves the pseudo-differential structure of the vector fields. By setting

$$\mathcal{F} = (\Phi_{n+1})_*(\hat{N}_n^{(1)} + \hat{N}_n^{(2)} + \hat{H}_n)$$

we have that the coefficients of $(\Pi_{\mathcal{N}}\mathcal{F})^{(w)}$ comes form $\Pi_{\mathcal{N}}(\Phi_{n+1})_*(\hat{N}_n^{(1)} + \hat{N}_n^{(2)} + \Pi_{\mathcal{N}}\hat{H}_n)$ or from $\Pi_{\mathcal{N}}(\Phi_{n+1})_*(\Pi_{\mathcal{N}^\perp}\hat{H}_n)$. Obviously the first coefficients satisfies (6.5.231) using (6.5.239) and the fact that $\Phi_{n+1} \approx \mathbb{1} + O(\delta K_n^{-\kappa_1+\mu})$. The second terms satisfies (6.5.231) because by Lemma B.180 one has

$$\Pi_{\mathcal{N}}(\Phi_{n+1})_*(\Pi_{\mathcal{N}^\perp}\hat{H}_n) := \Pi_{\mathcal{N}}\int_0^1 (\Phi_{n+1})_*^s[g_{n+1}, \Pi_{\mathcal{N}^\perp}\hat{H}_n]ds, \tag{6.5.245}$$

and hence one gets estimates (6.5.231) by using the estimate (3.2.73) on $g_{n+1}$. Let us analyse the first term. We note that Lemma B.184 holds also with $\hat{N}_0^{(n)}$, since it has only constant coefficients. Hence we have

$$\begin{aligned}\Pi_{\mathcal{N}}(\Phi_{n+1})_*(\hat{N}_0^{(n)}) &\overset{(B.31)}{=} \Pi_{\mathcal{N}}\int_0^1 ds(\Phi_{n+1})_*^s[g_{n+1},\hat{N}_0^{(n)}] = \Pi_{\mathcal{N}}\int_0^1 ds(\Phi_{n+1})_*^s\Pi_{K_{n+1}}\Pi_{\mathcal{A}}[g_{n+1},\hat{N}_0^{(n)}]\\ &= \Pi_{\mathcal{N}}\int_0^1 ds(\Phi_{n+1})_*^s\Pi_{K_{n+1}}\Pi_{\mathcal{A}}\left[g_{n+1},(1+h_{n+1})\Big(\hat{N}_0^{(n)}+\hat{N}_n^{(1)}+\hat{N}_n^{(2)}+\Pi_{\mathcal{N}}\hat{H}_n\Big)\right]\\ &- \Pi_{\mathcal{N}}\int_0^1 ds(\Phi_{n+1})_*^s\Pi_{K_{n+1}}\Pi_{\mathcal{A}}[g_{n+1},h_{n+1}\hat{N}_0^{(n)}]\\ &- \Pi_{\mathcal{N}}\int_0^1 ds(\Phi_{n+1})_*^s\Pi_{K_{n+1}}\Pi_{\mathcal{A}}\left[g_{n+1},(1+h_{n+1})\Big(\hat{N}_n^{(1)}+\hat{N}_n^{(2)}+\Pi_{\mathcal{N}}\hat{H}_n\Big)\right]\end{aligned} \tag{6.5.246}$$

Now we use the definition of $g_{n+1}$ and by item *(iii)* in Definition 3.2.37 we obtain that

$$\Pi_{K_{n+1}}\Pi_{\mathcal{A}}\left[g_{n+1},(1+h_{n+1})\Big(\hat{N}_0^{(n)}+\hat{N}_n^{(1)}+\hat{N}_n^{(2)}+\Pi_{\mathcal{N}}\hat{H}_n\Big)\right] = \Pi_{K_{n+1}}\Pi_{\mathcal{A}}\hat{F}_n + r_n \tag{6.5.247}$$

where $r_n$ satisfies bounds (3.2.60) and $\Pi_{K_{n+1}}\Pi_{\mathcal{A}}\hat{F}_n = (1+h_{n+1})\Pi_{\mathcal{A}}\hat{H}_n$. Equation (6.5.231) simply follows by applying Lemma B.182 and the inductive hypotheses.

In order to prove the inductive basis we reason as follows. First we note that if $n=0$ then we cannot apply Lemma 6.4.140 in order to define the map $\mathcal{T}_1$ and the field $\hat{F}_0$. On the other hand we apply Lemma 6.5.151 that provides the same result. Then on can reason as done in above using the map $\Phi_1$.

□

Now our aim is to give an explicit form to the sets $\mathcal{O}' \in \mathcal{M}_{K_{n+1},\delta K_n^{-\kappa_1},\gamma_n}(\hat{F}_n)$. We have the following result.

**Lemma 6.5.154.** *Consider the vector field $\hat{F}_n$ given in* (6.5.236) *and assume the Inductive hypothesis. Fix $\nu \geq 0$ as*

$$\nu := n + n^*, \qquad n^* := \log_{\frac{3}{2}}\frac{\kappa_1}{\kappa_4}, \quad \kappa_4 := 7\tau + 3. \tag{6.5.248}$$

*Then there exist a vector $\omega_{n+1} \in \mathbb{R}^d$ and a sequence of purely imaginary numbers*

$$\mu_{\sigma,j}(\xi) := \mu_{\sigma,j}^{(n)}(\xi) := -\mathrm{i}\sigma(c_{n+1}j^2 + r_j), \quad \sigma = \pm 1,\ j \in \mathbb{Z}_+ \cap S^c, \tag{6.5.249}$$

*with*

$$|r_j|_{\gamma_n} \leq C\xi, \quad |r_j - r_0^j|_{\gamma_n} \leq C\xi\delta, \qquad |c_{n+1} - c_n|_{\gamma_n} \leq \xi\delta K_n^{-\kappa_1}, \qquad |\omega_{n+1} - \omega_n|_{\gamma_n} \leq \xi\delta K_n^{-\kappa_1}, \tag{6.5.250}$$

*with $r_0^j$ defined in* (6.4.182)*, such that, by defining*

$$
\begin{aligned}
\Lambda_\nu^{2\gamma_n}(n+1) &:= \left\{ \begin{array}{ll} \xi \in \mathcal{O}_n : & |\omega_{n+1}\cdot l + \mu_{\sigma,j}^{(n)}(\xi) - \mu_{\sigma',j'}^{(n)}(\xi)| \geq \frac{2\gamma_n |\sigma j^2 - \sigma' j'^2|}{\langle l \rangle^\tau} \\ & \forall l \in \mathbb{Z}^d,\ |l| \leq K_\nu\ \ \forall (\sigma,j),(\sigma',j') \in \boldsymbol{C} \times (\mathbb{Z}_+ \cap S^c) \end{array} \right\}, \\
P_\nu^{2\gamma_n}(n+1) &:= \left\{ \begin{array}{ll} \xi \in \mathcal{O}_n : & |\omega_{n+1}\cdot l + \mu_{\sigma,j}^{(n)}(\xi)| \geq \frac{2\gamma j^2}{\langle l \rangle^\tau}, \\ & \forall \in \mathbb{Z}^d,\ \ |l| \leq K_\nu\ \ \forall\ (\sigma,j) \in \boldsymbol{C} \times \mathbb{Z}_+ \cap S^c \end{array} \right\}. \\
\mathcal{S}_\nu^{2\gamma_n}(n+1) &:= \left\{ \begin{array}{ll} \xi \in \mathcal{O}_n : & |\omega_{n+1}\cdot l| \geq \frac{2\gamma_n}{\langle l \rangle^\tau} \\ & \forall l \in \mathbb{Z}^d,\ |l| \leq K_\nu \end{array} \right\},
\end{aligned}
\tag{6.5.251}
$$

*one has that*

$$
\Lambda_\nu^{2\gamma_n} \cap P_\nu^{2\gamma_n} \cap \mathcal{S}_\nu^{2\gamma_n} \in \mathcal{M}_{K_{n+1},\delta K_n^{-\kappa_1},\gamma_n}(\hat{F}_n). \tag{6.5.252}
$$

*Proof.* We show that for any parameter $\xi \in \Lambda_\nu^{2\gamma_n} \cap P_\nu^{2\gamma_n} \cap \mathcal{S}_\nu^{2\gamma_n}$ there exist operators $W_0^{(n)}, W_\pm^{(n)}$ that satisfy all the properties of Definition 3.2.42 with $\gamma \rightsquigarrow \gamma_n$, $K \rightsquigarrow K_{n+1}$, $\mathfrak{r} \rightsquigarrow \delta K_n^{-\kappa_1}$ and $F \rightsquigarrow \hat{F}_n$. As we said at the beginning of Section 6.4.2, we look for an "approximate solution of the equation (6.4.137) where $K_+ \rightsquigarrow K_{n+1}$ and

$$
N \rightsquigarrow (1+h_{n+1})\Big(\hat{N}_0^{(n)} + \hat{N}_n^{(1)} + \hat{N}_n^{(2)} + \Pi_{\mathcal{N}}\hat{H}_n\Big).
$$

First of all we need that $\omega_{n+1}$ is a diophantine vector of $\mathbb{R}^d$. As first approximation we does not consider in the commutator the terms $\hat{N}_n^{(1)}, (\Pi_{\mathcal{N}}\hat{H}_n)^{(\theta)}$. This is possible because one of the consequences of Lemma 6.4.140 is to reduce the size of such terms. We will see that this approximation is sufficients to get a good approximate solutions that satisfies the requirements in Definition (3.2.42). Hence we study equation

$$
(\mathcal{L}_n)_\pm u = \Big(\omega_{n+1}\cdot\partial_\theta \pm \hat{\Omega}_n(\theta)\Big)u = f, \tag{6.5.253}
$$

with $f \in \mathbf{Z}$ and $u \in \mathbf{X}$ (see Definition 6.4.141) and where

$$
\hat{\Omega}_n := (\Pi_{\mathcal{N}}\hat{N}_0^{(n)})^{(w)} + \hat{N}_n^{(2)} + (\Pi_{\mathcal{N}}\hat{H}_n)^{(w)}. \tag{6.5.254}
$$

The following strategy can be applied indifferently to $(\mathcal{L}_n)_+$ of $(\mathcal{L}_n)_-$. Hence we will simply write $\mathcal{L}_n$. By the construction of the field $\hat{F}_n$ one has that the operator $\mathcal{L}_n$ satisfies the hypotheses of Lemma 6.4.142. Indeed the first smallness condition in (6.4.148) comes from the hypothesis (6.5.233) and the fact that the map $\mathcal{T}_{n+1}$ in Lemma 6.4.140 is of the form $\mathbb{1} + O(\delta K_n^{-1})$. The second bound on the coefficients $c_i^{(n)}, d_i^{(n)}$ is obtained in the following way. By inductive hypothesis we have the control of the $\mathfrak{p}_1-$norm in terms of $\Pi_{\mathcal{N}^\perp}G_0$. Now by Lemma 6.4.130 we have

$$
|\mathscr{K}_n|_{\vec{v},\mathfrak{p}_1}^{\mathrm{dec}} \leq (1 + K_n^{\mu_1}\delta K_{n-1}^{-\kappa_1}) C_{\vec{v},\mathfrak{p}_1+2d+1}(\Pi_{\mathcal{N}^\perp}G_0) \leq C\xi\delta. \tag{6.5.255}
$$

By applying Lemma 6.4.142 to the operator $\mathcal{L}_n$ in (6.5.253) we get the operator

$$\widetilde{\mathcal{L}}_n := \Pi_S^{\perp}(\omega_{n+1}\cdot\partial_\theta + \mathcal{D}_n + \mathcal{R}_n)\Pi_S^{\perp} \tag{6.5.256}$$

of the form (6.4.183) where

$$\mathcal{D}_n := -\mathrm{i}E\mathrm{diag}_{j\in\mathbb{Z}_+\cap S^c}(c_{n+1}j^2 + r_0^j), \qquad \mathcal{R}_n := E_1^{(n)}D + E_0^{(n)}, \tag{6.5.257}$$

as in (6.4.151). Again by construction we have that $\widetilde{\mathcal{L}}_n$ in (6.5.256) satisfies the hypotheses of Lemma 6.4.147. In order to prove a bound like (3.2.80) we fix the number $N > 0$ in Lemma 6.4.147 in such a way one has

$$N^{-\kappa_4} \leq K_n^{-\kappa_1}. \tag{6.5.258}$$

By the construction in the Lemma, one has that $N$ has the form $K_0^{\left(\frac{3}{2}\right)^\nu}\kappa_4$ for some $\nu \geq 0$. Hence to get the (6.5.258) it is enough to fix

$$\nu := n + n^*, \qquad n^* := \log_{\frac{3}{2}} \frac{\kappa_1}{\kappa_4}. \tag{6.5.259}$$

As we already seen, Lemma 6.4.147 is based on an KAM-like scheme. In other words, if we are at step $n$ of the abstract algorithm in Theorem 3.2.39, hence we have to perform $n + n^*$ Kam steps in Lemma 6.4.147.

By applying Lemma 6.4.147 with $N = K_0^{(\frac{3}{2})^\nu} := K_\nu$ to the truncated operator

$$\mathcal{L}_n^+ := \Pi_S^{\perp}(\omega_{n+1}\cdot\partial_\theta + \mathcal{D}_n + (\Pi_{K_{n+1}}E_1^{(n)})D + \Pi_{K_{n+1}}(E_0^{(n)}))\Pi_S^{\perp} \tag{6.5.260}$$

we have that, for the parameters $\xi \in \Lambda_\nu^{2\gamma_n} := \Lambda_{K_\nu}^{2\gamma_n}$ defined in (6.4.190), there is a map $\Phi_\nu := \Phi_{K_\nu}$ that satisfies (6.4.194) and conjugates $\mathcal{L}_n^+$ with the operator

$$(\mathcal{L}_n^+)_\nu := \Pi_S^{\perp}\Big(\omega_{n+1}\cdot\partial_\theta + \mathcal{D}_n^\nu + \mathcal{R}_n^\nu\Big)\Pi_S^{\perp} \tag{6.5.261}$$

where

$$\begin{aligned} \mathcal{D}_n^\nu &:= \mathrm{diag}_{\sigma\in\boldsymbol{C}, j\in\mathbb{Z}_+}(\mu_{\sigma,j}^\nu), \\ \mathcal{R}_n^\nu &: = (E_1^{(n)})^\nu D + (E_0^{(n)})^\nu, \end{aligned} \tag{6.5.262}$$

given by equation (6.4.192). Moreover using estimates (6.4.193), (6.4.194) and the Inductive Hypothesis we get the bounds:

$$|(E_1^{(n)})^\nu|_{\vec{v}_1,p}^{\mathrm{dec}} + |(E_0^{(n)})^\nu|_{\vec{v}_1,p}^{\mathrm{dec}} \leq K_{n+1}^{\kappa_5}\Big(|E_1^{(n)}|_{\vec{v},p}^{\mathrm{dec}} + |E_0^{(n)}|_{\vec{v},p}^{\mathrm{dec}}\Big)K_\nu^{-\kappa_4}, \quad \vec{v}_1 := (\gamma, \Lambda_N^{2\gamma}, s, a), \tag{6.5.263}$$

Moreover one has that

$$|\Phi_\nu^{\pm1} - \mathbb{1}|_{\vec{v}_1,p}^{\mathrm{dec}} \leq \gamma^{-1}\Big(|E_1^{(n)}|_{\vec{v},p}^{\mathrm{dec}} + |E_0^{(n)}|_{\vec{v},p}^{\mathrm{dec}}\Big). \tag{6.5.264}$$

In particular this means that

$$|(E_1^{(n)})^\nu|_{\vec{v}_1,\mathfrak{p}_1}^{\rm dec} + |(E_0^{(n)})^\nu|_{\vec{v}_1,\mathfrak{p}_1}^{\rm dec} \le K_{n+1}^{\kappa_5} K_n^{-\kappa_1}\xi\delta \tag{6.5.265}$$

Now consider the equation

$$\Big((\omega_+ + \hat{H}_n^{(\theta,0)})\cdot\partial_\theta + (\Pi_{\mathcal{N}}\hat{F}_n)^{(w)}\Big)u = f \tag{6.5.266}$$

We claim that, if we set

$$u := W_-^{(n)} f := \Phi_\nu^{-1}(\omega_+\cdot\partial_\theta + \mathcal{D}_n^\nu)^{-1}\Phi_\nu f, \tag{6.5.267}$$

then equation (3.2.80) holds. First of all we have, by (6.5.267), that (6.5.266) reads

$$Af := \Phi_\nu^{-1}\Big(\hat{N}_n^{(1)} + \hat{H}_n^{(\theta,0)}\cdot\partial_\theta + (\Pi_{K_{n+1}}^\perp E_1^{(n)})D + (\Pi_{K_{n+1}}^\perp E_0^{(n)})\Big)(\omega_+\cdot\partial_\theta + \mathcal{D}_n^\nu)^{-1}\Phi_\nu f = 0, \tag{6.5.268}$$

Hence using (3.2.61), Corollary 6.4.150, (6.4.194) to estimate $\Phi_\nu^\pm$, (6.4.193) to estimate $\mathcal{R}_n^\nu$, (6.4.75) and (6.4.77) we have (3.2.80) for the $w-$component and in particular

$$\begin{aligned} C_{\vec{v},\mathfrak{p}_1}(Af) &\le \gamma_n^{-1}\delta K_n^{-\kappa_1}K_{n+1}^\eta C_{\vec{v},\mathfrak{p}_1}(f),\\ C_{\vec{v},\mathfrak{p}_2}(Af) &\le \gamma_n^{-1}K^\eta\Big(C_{\vec{v},\mathfrak{p}_2}(f) + C_{\vec{v},\mathfrak{p}_1}(f)\mathfrak{A}_0 K_n^{\kappa_1+\kappa_5}\Big)\end{aligned} \tag{6.5.269}$$

Note that (3.2.80) follows even with $\alpha = 0$. The operator $W_+^{(n)}$ is constructed in the same way. Clearly by setting $W_0 := (\omega_+\cdot\partial_\theta)^{-1}$ and using that $\xi \in \mathcal{S}_\nu^{2\gamma_n}(n)$, one get the (3.2.80) by defining $\mathcal{W}^{(n)}$ as in (3.2.78). □

**Remark 6.5.155.** *Cosider $F_n$ the sequence of vector fields given by Theorem 1.2.5 and consider the approximate eigenvalues $\mu_{\sigma,j}^{(n)}(\xi)$ given by Lemma 6.5.154. By Remark 6.4.146 and 6.5.152 one has that*

$$\mu_{\sigma,j}^{(n)}(\xi) = c_0 j^2 + r_0^j + o(\xi)j^2 + o(\xi) = \Omega_j^{\rm int} + o(\xi)j^2 + o(\xi), \tag{6.5.270}$$

*where $\Omega^{\rm int}$ is defined in Section 6.6.*

## 6.6 Measure estimates

In this last Section we prove that the measure of the set of "good" parameters is large as $\xi \to 0$. In particular in Section 6.5 we have seen that Theorem 1.2.5 holds in the set

$$\mathcal{C}_\varepsilon := \bigcap_{n\ge1}\mathcal{O}_n, \qquad \mathcal{O}_n := \Lambda_\nu^{2\gamma_n}\cap P_\nu^{2\gamma_n}\cap\mathcal{S}_\nu^{2\gamma_n}, \tag{6.6.271}$$

with $\nu$ defined in (6.5.259) (see Lemma 6.5.154). Before performing such measure estimates we first prove that the map which link the parameters $\xi$ to the frequency $\omega(\xi)$ and $\xi \to \mu_{\sigma,j}(\xi)$ is a diffeomorphism.

### 6.6.1 The "twist" condition

Recall that by definition we can write $F = N_0 + G$ where $N_0 := (\omega^{-1} + \omega^{(0)}(\xi)) \cdot \partial_\theta + \Omega^{-1} w \partial_\theta$ where $\omega_j^{(-1)} := j^2, \omega_j^{(0)}(\xi) := -(\mathcal{M}\xi)_j$, $j \in S^+$ and $(\Omega^{-1})_\sigma^\sigma = \mathrm{i}\sigma \mathrm{diag} j^2$, $(\Omega^{-1})_\sigma^{-\sigma} = 0$. Hence by equation (6.3.24) we have

$$\begin{aligned}
&\Pi_{\mathcal{N}} F = \omega(\xi,\theta)\partial_\theta + \Omega(\theta,\xi) w \partial_w, \\
&\omega(\xi,\theta) = \omega^{(-1)} + \omega^{(0)}(\xi) + G^{(\theta,0)}(\xi,\theta), \\
&\Omega(\theta,\xi) = \Omega^{(-1)} + \Omega^{(0)}(\theta,\xi) = d_w F^{(w)}(\theta,0,0)[\cdot] = \Omega^{-1} + d_w G^{(w)}(\theta,0,0)[\cdot]
\end{aligned} \tag{6.6.272}$$

Let us study in particular the linear operator $\Omega$ on $\Pi_S^\perp \mathbf{h}_{\mathrm{odd}}^{a_0,p}$. We have that $\Omega^{-1} := -iE\partial_{xx} : \Pi_S^\perp \mathbf{h}_{\mathrm{odd}}^{a_0,0} \to \Pi_S^\perp \mathbf{h}_{\mathrm{odd}}^{a_0,p-2}$ where $E := \mathrm{diag}\{1,-1\}$. Moreover one has that $\Omega^{(0)} = ((\Omega^{(0)})_\sigma^{\sigma'})_{\sigma,\sigma'=\pm 1}$ can be seen as a 2 times 2 matrix whose components are operator on $H^{a_0,p}$. In particular it has the form given by (6.3.25).

We have the vector field

$$\begin{cases}
\dot\theta = \omega(\xi,\theta) + \Pi_{\mathcal{N}^\perp} G^{(\theta)}(\theta,y,w) \\
\dot y = G^{(y)}(\theta,y,w) \\
\dot w = \Omega(\xi,\theta) w + \Pi_{\mathcal{N}^\perp} G^{(w)}(\theta,y,w)
\end{cases} \tag{6.6.273}$$

where $G$ is small. In order to run our algorithm we need to reduce the matrix $\Omega(\omega(\xi)t;\xi)$. In order to do this perturbatively we need to impose second Melnikov conditions. The minimal requirement (so that the reduction algorithm runs at least at a formal level) is that the difference of the eigenvalues is not identically zero as function of $\xi$, namely

- *Twist.* Denote by $\mu_j(\xi)$ for $j \in S^c$ the *eigenvalue functions* of $\Omega(\theta,\xi)$. For all $l, j, k, \sigma_1, \sigma_2$ such that: if $\sigma_1 = \sigma_2$ then $(l,j,k) \neq (0,j,j)$ and moreover $\sum_i l_i + \sigma_1 = \sigma_2$ consider the map

$$\xi \to \widetilde\omega(\xi) \cdot l + \sigma_1 \mu_j(\xi) - \sigma_2 \mu_k(\xi) \tag{6.6.274}$$

  where we defined $\omega(\theta,\xi) := \widetilde\omega(\xi) + \Pi_{\mathcal{N}} G^{(\theta)} = \omega^{-1} + \omega^{(0)}(\xi) + \Pi_{\mathcal{N}} G^{(\theta)}$. We require that these maps are never identically zero.

This is the reason why we needed to introduce $\omega^{(0)}, \Omega^{(0)}$ since clearly

$$\omega^{(-1)} \cdot l + \Omega_j^{(-1)} \pm \Omega_k^{(-1)} \equiv 0$$

for infinitely many choices of $l, j, k$.

We split:

$$\Omega(\theta,\xi) = \Omega^{\mathrm{int}}(\xi) + \widetilde\Omega^{(0)}(\theta,\xi)$$

where

$$\Omega^{(\text{int})}(\xi) = \Omega^{(-1)} + \text{diag}(\mathtt{m}_j \cdot \xi)_{j \in S^c \cap \mathbb{Z}_+} = \text{diag}(j^2 + \mathtt{m}_j \cdot \xi)_{j \in S^c \cap \mathbb{Z}_+}, \tag{6.6.275}$$

$$\mathtt{m}^i_j = \frac{1}{4}\big(C_j^{\mathtt{v}_i} + C_j^{-\mathtt{v}_i}\big) = \frac{1}{2}\Big(2\mathtt{a}_1 - \mathtt{a}_2(j^2 + \mathtt{v}_i^2) + \mathtt{a}_3\mathtt{v}_i^2 - 2\mathtt{b}_2\mathtt{v}_i^2 - \mathtt{b}_3\mathtt{v}_i^2 - \mathtt{a}_4\mathtt{v}_i^2 j^2 - 2\mathtt{a}_6\mathtt{v}_i^4 j^2\Big)$$

$$\mathtt{m}_j = \frac{1}{2}\Big(2\mathtt{a}_1 - \mathtt{a}_2(j^2 + \mathtt{V}^2) + (\mathtt{a}_3 - 2\mathtt{b}_2 - \mathtt{b}_3)\mathtt{V}^2 - \mathtt{a}_4 j^2\mathtt{V}^2 - 2\mathtt{a}_6 j^2\mathtt{V}^4\Big)\vec{1},$$

where $\mathtt{V} := \text{diag}_i(\mathtt{v}_i)$. Note that $\widetilde{\Omega}$ is of the same order as $\Omega^{(\text{int})}$, however it turns out that for *generic* choices of $\mathtt{a}_1, \dots, \mathtt{a}_6, \mathtt{b}_2, \mathtt{b}_3, \mathtt{b}_4, \mathtt{v}_1, \dots, \mathtt{v}_d$:

- $\widetilde{\omega}, \Omega^{(0)}_{\text{int}}(\xi)$ satisfy the *twist conditions*, namely $\forall l, j, k$ $\sigma_1, \sigma_2 = 0, \pm 1$ the affine maps

$$\xi \to \widetilde{\omega}(\xi) \cdot l + \sigma_1 \Omega_j^{(\text{int})}(\xi) - \sigma_2 \Omega_k^{(\text{int})}(\xi)$$

  are not identically zero (with the usual restrictions on $(l, j, k)$).

- The twist condition above implies the corresponding twist condition for the $\mu_j$ (see (6.6.274)).

Now we prove that our normal form satisfies the twist condition. First we introduce the following non-resonance condition.

**Definition 6.6.156.** *We say that* $(\mathtt{a}, \mathtt{b}) := (\mathtt{a}_1, \mathtt{a}_2, \mathtt{a}_3, \mathtt{a}_4, \mathtt{a}_6, \mathtt{b}_2, \mathtt{b}_3, \mathtt{b}_4)$ *is* non-resonant *if one of the following occurs:*

1. $\mathtt{a}_6 \neq 0$,

2. $\mathtt{a}_6 = 0$ *and* $\mathtt{a}_1 \neq 0$,

3. $\mathtt{a}_6 = \mathtt{a}_1 = 0$, $-\mathtt{a}_4 + \mathtt{b}_4 \neq 0$ *and one of the following holds:*

   - $\mathtt{a}_4 = 0$, *or*

   - $\mathtt{a}_4 \neq 0$ *and* $(2d-1)\mathtt{a}_4 - \mathtt{b}_4 \neq 0$ *or*

4. $\mathtt{a}_6 = \mathtt{a}_1 = -\mathtt{a}_4 + \mathtt{b}_4 = 0$, $\mathtt{a}_3 - \mathtt{a}_2 - \mathtt{b}_2 - \mathtt{b}_3 \neq 0$ *and one of the following holds:*

   - $\mathtt{a}_2 = 0$ *and* $\mathtt{a}_3 - 3\mathtt{b}_2 - \mathtt{b}_3 = 0$, *or*

- $\mathtt{a}_2 \neq 0$, $\mathtt{a}_3 - \mathtt{a}_2 - 3\mathtt{b}_2 - \mathtt{b}_3 \neq 0$, *or*

- $\mathtt{a}_2 = 0$, $\mathtt{a}_3 - 3\mathtt{b}_2 - \mathtt{b}_3 \neq 0$ *and* $\mathtt{a}_3 - \frac{6d+1}{2d+1}\mathtt{b}_2 - \mathtt{b}_3 \neq 0$, *or*

- $\mathtt{a}_2 \neq 0$, $\mathtt{a}_3 - \mathtt{a}_2 - 3\mathtt{b}_2 - \mathtt{b}_3 = 0$ *and* $d\mathtt{a}_2 \neq \mathtt{b}_2$.

Note that a non-resonant vector $(\mathtt{a}, \mathtt{b})$ is "generic" in the sense of Definition 1.2.4.

**Lemma 6.6.157.** *For all non-resonant choices of* $(\mathtt{a}, \mathtt{b})$ *there exists a "generic" choice of the tangential sites* $S^+ = \{\mathtt{v}_1, \dots, \mathtt{v}_d\} \subset \mathbb{N}$ *such that the map*

$$\varepsilon^2 \Lambda \ni \xi \to \widetilde{\omega}(\xi) = \omega^{(-1)}(\xi) - \mathcal{M}\xi \tag{6.6.276}$$

*is a affine diffeomorphism.*

*Proof.* Since $\widetilde{\omega}$ is affine we only need to show that $\mathcal{M}$ is invertible. Recalling that $\mathcal{M}_{ij} = (1/4)\big(C_{\mathtt{v}_i}^{\mathtt{v}_j} + C_{\mathtt{v}_i}^{-\mathtt{v}_j}\big)$ for $i, j = 1, \dots, d$. It is convenient to represent

$$\mathcal{M} = \frac{1}{4}\sum_{k=0}^{3} \mathcal{M}^{(2k)} \tag{6.6.277}$$

where the matrix elements $\mathcal{M}^{(2k)}$ are homogeneous of degree $2k$ in the variables $\mathtt{v}_1, \dots, \mathtt{v}_d$. More precisely setting $V = \mathrm{diag}(\mathtt{v}_i)$, $A_{ij} = 1$, forall $i, j$, we have

$$\mathcal{M}^{(0)} = \mathtt{a}_1(4A - \mathbb{1}), \quad \mathcal{M}^{(6)} = -\mathtt{a}_6 \mathtt{V}^2(4A - \mathbb{1})\mathtt{V}^4$$

$$\mathcal{M}^{(2)} := -(\mathtt{a}_3 - \mathtt{a}_2 - \mathtt{b}_2 - \mathtt{b}_3)\mathtt{V}^2 + (\mathtt{a}_3 - \mathtt{a}_2 - 3\mathtt{b}_2 - \mathtt{b}_3)2A\mathtt{V}^2 - 2\mathtt{a}_2\mathtt{V}^2 A$$

$$\mathcal{M}^{(2)}_{ik} = \begin{cases} (\mathtt{a}_3 - \mathtt{a}_2)\mathtt{v}_i^2 - 2\mathtt{a}_2\mathtt{v}_i^2 - 3\mathtt{b}_2\mathtt{v}_i^2 - \mathtt{b}_3\mathtt{v}_i^2 & \text{if} \quad i = k \\ 2(\mathtt{a}_3 - \mathtt{a}_2)\mathtt{v}_k^2 - 2\mathtt{a}_2\mathtt{v}_i^2 - 4\mathtt{b}_2\mathtt{v}_k^2 - 2\mathtt{b}_3\mathtt{v}_k^2 & \text{if} \quad i \neq k \end{cases}$$

$$\mathcal{M}^{(4)}_{ik} = \begin{cases} (-\mathtt{a}_4 - \mathtt{b}_4)\mathtt{v}_i^4 & \text{if} \quad i = k \\ -2\mathtt{a}_4\mathtt{v}_i^2\mathtt{v}_k^2 & \text{if} \quad i \neq k \end{cases}$$

$$\mathcal{M}^{(4)} := -(-\mathtt{a}_4 + \mathtt{b}_4)\mathtt{V}^4 - 2\mathtt{a}_4\mathtt{V}^2 A\mathtt{V}^2$$

We now compute $P(\mathtt{a}, \mathtt{b}, \mathtt{v}) :=\det(\mathcal{M})$ which is a non trivial polynomial in $(\mathtt{a}_1, \dots, \mathtt{a}_6, \mathtt{b}_2, \mathtt{b}_3, \mathtt{b}_4, \mathtt{v}_1, \dots, \mathtt{v}_d)$. Indeed $P(\mathtt{a}, 0) = \det(\mathcal{M}^{(0)}) = \mathtt{a}_1^d(2d - 1)$, so *for any* $\mathtt{a}$ such that $\mathtt{a}_1 \neq 0$ we impose $P(\mathtt{a}, \mathtt{v}) \neq 0$ as genericity condition on the $\mathtt{v}$.

In the same way, the term of highest degree in $\mathtt{v}$ is $\det(\mathcal{M}^{(6)}) = \mathtt{a}_6(2d-1)\prod_i \mathtt{v}_i^6$, so again *for any* $\mathtt{a}$ such that $\mathtt{a}_6 \neq 0$ we impose $P(\mathtt{a},\mathtt{v}) \neq 0$ as genericity condition on the $\mathtt{v}$.

We are left with the case $\mathtt{a}_1 = \mathtt{a}_6 = 0$. Now the term of minimal degree is $\det(\mathcal{M}^{(2)})$ while term of maximal degree is $\det(\mathcal{M}^{(4)})$. Now we show that for generic choices of $\mathtt{a}_i, \mathtt{b}_i$ then $\det(\mathcal{M}^{(2)})$ is not identically zero as function of the $\mathtt{v}_i$. First of all we set

$$\lambda := -(\mathtt{a}_3 - \mathtt{a}_2 - \mathtt{b}_2 - \mathtt{b}_3), \quad \alpha := (\mathtt{a}_3 - \mathtt{a}_2 - 3\mathtt{b}_2 - \mathtt{b}_3)2, \quad \beta := -2\mathtt{a}_2. \tag{6.6.278}$$

Hence we can write

$$\mathcal{M}^{(2)} = \lambda \mathtt{V}^2 + \alpha A \mathtt{V}^2 + \beta \mathtt{V}^2 A.$$

Assume that $\lambda \neq 0$. Now if $\beta = 0$ and $\alpha \neq 0$ then one has

$$\mathcal{M}^{(2)} := (\mathbb{1} + \frac{\alpha}{\lambda} A)\mathtt{V}^2.$$

The first matrix in the product is invertible if has all the eigenvalues different from zero. Hence we impose that

$$1 + \frac{\alpha}{\lambda} d \neq 0, \quad \text{i.e.} \quad \mathtt{a}_3 - \mathtt{a}_2 - \frac{6d+1}{2d+1}\mathtt{b}_2 - \mathtt{b}_3 \neq 0. \tag{6.6.279}$$

If on the contrary $\alpha = 0$ and $\beta \neq 0$ then

$$\mathcal{M}^{(2)} := \mathtt{V}^2(\mathbb{1} + \frac{\beta}{\lambda} A),$$

which is invertible if

$$1 + \frac{\beta}{\lambda} d \neq 0, \quad \text{i.e.} \quad \mathtt{a}_3 - (2d+1)\mathtt{a}_2 - \mathtt{b}_2 - \mathtt{b}_3 \neq 0. \tag{6.6.280}$$

Consider the case $\alpha \neq 0$ and $\beta \neq 0$. Then we have

$$\mathcal{M}^{(2)} = (\mathbb{1} + \frac{\alpha}{\lambda} A + \frac{\beta}{\lambda} \mathtt{V}^2 A \mathtt{V}^{-2})\mathtt{V}^2.$$

The invertibility of $\mathcal{M}^{(2)}$ relies on the invertibility of the matrix $\mathbb{1} + R := \mathbb{1} + \frac{\alpha}{\lambda} A + \frac{\beta}{\lambda}\mathtt{V}^2 A \mathtt{V}^{-2}$. We have that $R$ has at most rank 2, hence has at most two eigenvalues different from zero. Say that $\mu_{1,2} = \mu_{1,2}(\mathtt{v}_i)$ is such eigenvalues that in principle depends on the $\mathtt{v}_i$. Now one has that $\mathbb{1} + R$ has $d-2$ eigenvalues equals to 1 and two equals to $1 + \mu_{1,2}(\mathtt{v}_i)$. One must have that $1 + \mu_{1,2}(\mathtt{v}_i) \neq 0$. Hence if $\mu(\mathtt{v}_i)$ is not a trivial polynomial in the variables $\mathtt{v}_i$ then one get the invertibility of $\mathcal{M}^{(2)}$ as genericity condition on $\mathtt{v}_i$. Otherwise one has to exclude some values of $\frac{\alpha}{\lambda}$ and $\frac{\beta}{\lambda}$ by imposing a genericity condition on $\mathtt{a}_2, \mathtt{a}_3, \mathtt{b}_2, \mathtt{b}_3$ (as done in equations (6.6.279) and (6.6.280)) and then taking a generic choice of $\mathtt{v}_i$. This second option does not occur. Indeed one note that the vector $\vec{w}_1 := (1, \dots, 1) \in \mathbb{R}^d$ is orthogonal to the kernel of the matrix $\frac{\alpha}{\lambda} A$. Moreover the vector $\vec{w}_2 := \vec{v}$, where $\vec{v} := (\mathtt{v}_1^2, \dots, \mathtt{v}_d^2)$, is orthogonal to the kernel of the matrix $y\mathtt{V}^2 A \mathtt{V}^{-2}$. Hence the range of the matrix $R$ is generated by $\{\vec{w}_1, \vec{w}_2\}$. One can note that

$$\begin{aligned}
&(\frac{\alpha}{\lambda}A + \frac{\beta}{\lambda}\mathtt{V}^2 A \mathtt{V}^{-2})\vec{w}_1 = \frac{\alpha}{\lambda}d\vec{w}_1 + \frac{\beta}{\lambda}C_1\vec{w}_2,\\
&(\frac{\alpha}{\lambda}A + \frac{\beta}{\lambda}\mathtt{V}^2 A \mathtt{V}^{-2})\vec{w}_2 = \frac{\alpha}{\lambda}C_2\vec{w}_1 + \frac{\beta}{\lambda}d\vec{w}_2,\\
&\text{where} \quad C_1 := \sum_{i=1}^{d}\frac{1}{\mathtt{v}_i^2}, \quad C_2 := \sum_{i=1}^{d}\mathtt{v}_i^2.
\end{aligned} \tag{6.6.281}$$

The $2 \times 2$ matrix which represent the matrix $R$ has eigenvalues given by

$$\mu_{1,2} := \frac{1}{2\lambda}\left(d(\alpha+\beta) \pm \sqrt{d^2(\beta-\alpha)^2 + 4\alpha\beta C_1 C_2}\right) \tag{6.6.282}$$

The dimension of the range of $R$, for any $\alpha/\lambda \neq 0$ and $\beta/\lambda \neq 0$, depends only on the $\mathtt{v}_i$ for $i = 1, \dots, d$.

The same reasoning holds verbatim if $\mathtt{a}_1 = \mathtt{a}_6 = 0$, $\lambda = 0$ (see (6.6.278)) but

$$\lambda_1 := -(\mathtt{b}_4 - \mathtt{a}_4), \tag{6.6.283}$$

indeed one can write

$$\mathcal{M}^{(4)} := \mathtt{V}^4 + \frac{-2\mathtt{a}_4}{\lambda_1}\mathtt{V}^2 A \mathtt{V}^2.$$

Here, as in the case of $\mathcal{M}^{(2)}$, we get some additional conditions on $\mathtt{a}_i$ and $\mathtt{b}_i$: if $\mathtt{a}_4 = 0$, then $\mathcal{M}^{(4)}$ is invertible, otherwise we have the invertibility of the matrix if

$$\mathtt{a}_4 \neq 0 \quad \text{and} \quad \mathtt{a}_4(2d-1) - \mathtt{b}_4 \neq 0. \tag{6.6.284}$$

Suppose that $\lambda = \lambda_1 = 0$ then

$$\mathcal{M} = (\mathtt{a}_3 - \mathtt{a}_2 - 3\mathtt{b}_2 - \mathtt{b}_3)2A\mathtt{V}^2 - 2\mathtt{a}_2\mathtt{V}^2 A - 2\mathtt{a}_4\mathtt{V}^2 A \mathtt{V}^2,$$

which has at most rank 2. □

**Lemma 6.6.158.** *For all non-resonant choices of* $(\mathtt{a}_1, \mathtt{a}_2, \mathtt{a}_3, \mathtt{a}_4\mathtt{a}_6, \mathtt{b}_2, \mathtt{b}_3, \mathtt{b}_4)$ *there exists a no-trivial polynomial in the* $\mathtt{v}_i$ *such that for all choices of* $(\mathtt{v}_1, \dots, \mathtt{v}_d)$ *with* $\mathtt{v}_i$ *"generic" with respect to the polynomial the following holds. For all* $\ell, j, k, \sigma_1, \sigma_2$ *such that: if* $\sigma_1 = \sigma_2$ *then* $(l, j, k) \neq (0, j, j)$ *and moreover* $\sum_i l_i + \sigma_1 = \sigma_2$ *the affine map*

$$\xi \to \widetilde{\omega}(\xi) \cdot l + \sigma_1 \Omega_j^{(\text{int})}(\xi) - \sigma_2 \Omega_k^{(\text{int})}(\xi) \tag{6.6.285}$$

*is not identically zero.*

*Proof.*

$$\widetilde{\omega}(\xi)\cdot l+\sigma_1\Omega_j^{(\mathrm{int})}(\xi)-\sigma_2\Omega_k^{(\mathrm{int})}(\xi)=$$

$$\omega^{(-1)}\cdot\ell+\sigma_1 j^2-\sigma_2 k^2+\big(\mathcal{M}^T\ell+\sigma_1\mathtt{m}_j-\sigma_2\mathtt{m}_k\big)\cdot\xi$$

then if $\omega^{(-1)}\cdot\ell+\sigma_1 j^2-\sigma_2 k^2=0$, and using (6.6.275), we look at the vector

$$\mathcal{M}^T\ell+\frac{1}{2}(\sigma_1-\sigma_2)\big(2\mathtt{a}_1+(\mathtt{a}_3-\mathtt{a}_2-2\mathtt{b}_2-\mathtt{b}_3)\mathtt{V}^2\big)\vec{1}+$$

$$+\frac{1}{2}(\sigma_1 j^2-\sigma_2 k^2)\big(-\mathtt{a}_2-\mathtt{a}_4)\mathtt{V}^2-2\mathtt{a}_6\mathtt{V}^4\big)\vec{1}=$$

$$\Big(\mathcal{M}^T+\frac{1}{2}\Big(\mathtt{a}_2+\mathtt{a}_4\mathtt{V}^2+2\mathtt{a}_6\mathtt{V}^4\Big)A\mathtt{V}^2-\frac{1}{2}\big(2\mathtt{a}_1+(\mathtt{a}_3-\mathtt{a}_2-2\mathtt{b}_2-\mathtt{b}_3)\mathtt{V}^2\big)A\Big)\ell$$

since $(\sigma_1 j^2-\sigma_2 k^2)\vec{1}=-A\mathtt{V}^2\ell$ and $(\sigma_1-\sigma_2)\vec{1}=-A\ell$. Hence, by using (6.6.277), we say that a list $(\mathtt{a},\mathtt{v})$ is acceptable if for all $\ell,j,k$ such that $\sum_i\ell_i=-\sigma_1+\sigma_2$ one has

$$-\frac{1}{4}\Big(\mathtt{a}_1\mathbb{1}+(\mathtt{a}_3-\mathtt{a}_2-\mathtt{b}_2-\mathtt{b}_3)\mathtt{V}^2+(-\mathtt{a}_4+\mathtt{b}_4)\mathtt{V}^4-\mathtt{a}_6\mathtt{V}^6\Big)\ell\neq 0 \tag{6.6.286}$$

then one only needs to require that none of the $\mathtt{v}_i$ satisfy

$$p(x):=\mathtt{a}_1+(\mathtt{a}_3-\mathtt{a}_2-\mathtt{b}_2-\mathtt{b}_3)x^2+(-\mathtt{a}_4+\mathtt{b}_4)x^4-\mathtt{a}_6x^6=0.$$

The hypothesis of non resonance implies that $p$ is not identically zero. □

**Remark 6.6.159.** *Just to fix the ideas we give some examples of cubic non linearity (see* (1.2.18)*) for which the extraction of parameters give the twist condition on the tangential sites. The classical cubic NLS with* $\mathtt{a}_1=1$*,* $\mathtt{a}_i\equiv 0$ *for* $i=2,\dots,6$ *and* $\mathtt{b}_i=0$ *for* $i=2,3,4,5$ *. The derivative NLS* $\mathtt{a}_3=1$*,* $\mathtt{a}_i=0$ *for* $i=1,2,4,6$ *and* $\mathtt{b}_i=0$ *for* $i=2,3,4$*. (this case has been studied in [61]).*

### 6.6.2 The estimates of "good" parameters

We prove the following Proposition.

**Proposition 6.6.160.** *Consider the set* $\mathcal{C}_\varepsilon$ *defined in* (6.6.271)*. One has that*

$$|\mathcal{O}_0\backslash\mathcal{C}_\varepsilon|\to 0,\quad \text{as}\quad \gamma_0\to 0. \tag{6.6.287}$$

For simplicity we set $G_n^{(1)}:=\Lambda_\nu^{2\gamma_n}(n), G_n^{(2)}:=\mathcal{P}_\nu^{2\gamma_n}(n), G_n^{(3)}:=\mathcal{S}_\nu^{2\gamma_n}(n)$ (see equation (6.5.251)). In order to prove (6.6.287) we prove prove by induction that, for any $n\geq 0$, one has

$$|G_0^{(i)}\backslash G_1^{(i)}|\leq C_\star\gamma,\quad |G_n^{(i)}\backslash G_{n+1}^{(i)}|\leq C_\star\gamma K_n^{-1},\quad n\geq 1,\quad i=1,2,3. \tag{6.6.288}$$

We follows the same strategy used in Section 6 of [31] and we bounds the measure only of the sets $G_n^{(1)}$ which is the more difficult case. The other estimates can be obtained in the same way. First of all write, dropping the index 1,

$$G_n\backslash G_{n+1} := \bigcup_{\substack{\sigma,\sigma'\in \boldsymbol{C}, j,j'\in\mathbb{Z}_+\\ l\in\mathbb{Z}^n}} R^{\sigma,\sigma'}_{ljj'}(n) \tag{6.6.289}$$

$$R^{\sigma,\sigma'}_{ljj'}(n) \quad := \left\{\lambda\in G_n : |\mathrm{i}\omega_{n+1}\cdot\ell + \mu^{(n)}_{\sigma,j} - \mu^{(n)}_{\sigma',j'}| < \tfrac{2\gamma_n|\sigma j^2-\sigma' j'^2|}{\langle l\rangle^\tau}\right\}.$$

By (6.3.28) we have $R^{\sigma,\sigma}_{ljj}(n)=\emptyset$ and moreover recalling (6.5.259) for $|l|\leq K_{n+n^*}$ one has $R^{\sigma,\sigma'}_{ljj'}(n)=\emptyset$. In the following we assume that if $\sigma=\sigma'$, then $j\neq j'$. Important properties of the sets $R^{\sigma,\sigma'}_{ljj'}(n)$ are the following. The proofs are quite standard and follow very closely Lemmata 5.2 and 5.3 in [4].

**Lemma 6.6.161.** *For any $n\geq 0$, $|\ell|\leq K_{n+n^*}$, one has, for $|\xi|$ small enough,*

$$R^{\sigma,\sigma'}_{ljj'}(n)\subseteq R^{\sigma,\sigma'}_{ljj'}(n-1). \tag{6.6.290}$$

*Moreover,*

$$\text{if}\qquad R^{\sigma,\sigma'}_{ljj'}\neq\emptyset,\qquad \text{then}\qquad |\sigma j^2-\sigma' j'^2|\leq 8|\widetilde{\omega}\cdot l|. \tag{6.6.291}$$

*Proof.* We first prove the (6.6.291); note that if $(\sigma,j)=(\sigma',j')$ then it is trivially true. If $R^{\sigma,\sigma'}_{ljj'}(n)\neq\emptyset$, then, by definition (6.6.289), there exists a $\xi\in\mathcal{O}_0$ such that

$$|\mu^{(n)}_{\sigma,j}-\mu^{(n)}_{\sigma',j'}| < 2\gamma_n|\sigma j^2-\sigma' j'^2|\langle l\rangle^{-\tau} + 2|\omega_{n+1}\cdot l|. \tag{6.6.292}$$

On the other hand, for $\xi$ small and since $(\sigma,j)\neq(\sigma',j')$,

$$|\mu^{(n)}_{\sigma,j}-\mu^{(n)}_{\sigma',j'}| \overset{(6.5.250)}{\geq} \frac{1}{2}(1-C\xi)|\sigma j^2-\sigma' j'^2| - C\xi \geq \frac{1}{3}|\sigma j^2-\sigma' j'^2|. \tag{6.6.293}$$

By the (6.6.292), (6.6.293) and $\gamma_n\leq 2\gamma$ follows

$$2|\omega_{n+1}\cdot l| \geq \left(\frac{1}{3}-\frac{4\gamma}{\langle\ell\rangle^\tau}\right)|\sigma j^2-\sigma' j'^2| \geq \frac{1}{4}|\sigma j^2-\sigma' j'^2|, \tag{6.6.294}$$

since $\gamma\leq\gamma_0$, by choosing $\gamma_0$ small enough. It is sufficient $\gamma_0<1/48$. Then, the (6.6.291) hold.
In order to prove the (6.6.290) we need to understand the variation of the eigenvalues $\mu^{(n)}_{\sigma,j}$ with respect to $n$. In other word the eigenvalues of the linearized operator of the field $F_n$. If we assume that

$$|(\mu^{(n)}_{\sigma,j}-\mu^{(n)}_{\sigma',j'}) - (\mu^{(n-1)}_{\sigma,j}-\mu^{(n-1)}_{\sigma',j'})| \leq C\xi|\sigma j^2-\sigma' j'^2|K^{-\kappa_4}_{n+n^*}, \tag{6.6.295}$$

then, for $j\neq j'$, $|l|\leq K_{n+n^*}$, and $\xi\in G_n$, we have

$$\begin{aligned} |\mathrm{i}\omega_{n+1}\cdot l+\mu^{(n)}_{\sigma,j} \quad &- \quad \mu^{(n)}_{\sigma',j'}| \overset{(6.6.295)}{\geq} 2\gamma_{n-1}|\sigma j^2-\sigma' j'^2|\langle l\rangle^{-\tau} \\ &- \quad C\xi|\sigma j^2-\sigma' j'^2|K^{-\kappa_4}_{n+n^*} \geq 2\gamma_n|\sigma j^2-\sigma' j'^2|\langle\ell\rangle^{-\tau}, \end{aligned} \tag{6.6.296}$$

because $C\delta K_{n+n^*}^{\tau-\kappa_4}2^{n+1} \leq 1$. We complete the proof by verifying (6.6.295).

By Lemma 6.4.148, using the $(\mathbf{S4})_\nu$ (for $\nu \leq n+n^*$) with $\gamma = \gamma_{n-1}$ and $\gamma - \rho = \gamma_n$, and with $\mathcal{L}_1 = \mathcal{L}_{n-1}$, $\mathcal{L}_2 = \mathcal{L}_n$, (where $\mathcal{L}_{n-1}$ and $\mathcal{L}_n$ are the linearized operator of the vector fields $F_{n-1}$ and $F_n$ rispectively) we have

$$\Lambda_\nu^{\gamma_{n-1}}(n-1) \subseteq \Lambda_\nu^{\gamma_n}(n), \tag{6.6.297}$$

since, for $|\xi|$ small enough,

$$\varepsilon C N_{n+n^*}^\tau \sup_{\xi\in\mathcal{O}_0} |E_0^0(\mathcal{L}_{n-1}) - E_0^0(\mathcal{L}_n)|_{\vec{v},\mathfrak{p}_0} \overset{(6.5.235)}{\leq} \delta C C_\star K_{n+n^*}^\tau K_n^{-\kappa_1+\mu_1+4} \leq \gamma_{n-1} - \gamma_n =: \rho = \gamma 2^{-n}.$$

We also note that,

$$G_n \overset{(6.5.251)}{\subseteq} \Lambda_\nu^{2\gamma_{n-1}}(n) \overset{(6.6.297)}{\subseteq} \Lambda_\nu^{\gamma_n}(n+1). \tag{6.6.298}$$

This means that $\xi \in G_n \subset \Lambda_\nu^{\gamma_{n-1}}(n) \cap \Lambda_\nu^{\gamma_n}(n+1)$, and hence, we can apply the $(\mathbf{S3})_\nu$, with $\nu = n+1$, in Lemma 6.4.148 to get

$$|r_{\sigma,j}^\nu(n) - r_{\sigma,j}^{\nu-1}(n-1)| \leq \varepsilon K_{n+n^*}^{-\alpha}. \tag{6.6.299}$$

Then, by (6.5.228) and (C.12), one has that the (6.6.295) hold and the proof of Lemma (6.6.161) is complete.

□

The next Lemma is fundamental. It is the equivalent of Lemma 4.5.84 in Section 4.5 and its proof is very similar. Anyway in the autonomous cases it is slightly more difficult. This is due to the fact that if one "move" the parameters $\xi$, then $\omega(\xi)$ and $\mu_{\sigma,j}$ moves together. This is why one need to prove that the entire map in (6.6.285) must have the "twist", and it is not enough to ask that $\xi \to \omega(\xi)$ is a diffeomorphism.

**Lemma 6.6.162.** *For all $n \geq 0$, one has*

$$|R_{ljj'}^{\sigma,\sigma'}(n)| \leq C\gamma\langle\ell\rangle^{-\tau+1}. \tag{6.6.300}$$

*Proof.* Let us define the map $\psi : \mathcal{O}_0 \to \mathbb{C}$

$$\begin{aligned}\psi(\xi) &:= \mathrm{i}\omega_{n+1}(\xi)\cdot l + \mu_{\sigma,j}^{(n)}(\xi) - \mu_{\sigma',k}^{(n)}(\xi)\\ &\overset{(6.5.249),(6.5.250)}{=} \mathrm{i}\Big(\omega^{-1}\cdot l + \sigma j^2 - \sigma' k^2 + \omega^{(0)}(\xi)\cdot l + c_0(\xi)(\sigma j^2 - \sigma' k^2) + \sigma r_0^j - \sigma' r_0^k\Big) \qquad (6.6.301)\\ &+ \mathrm{i}\Big((\omega_{n+1} - \widetilde{\omega})\cdot l + (c_{n+1} - c_0)(\sigma j^2 - \sigma k^2) + (\sigma(r_j^{(n)} - r_0^j) - \sigma'(r_k^{(n)} - r_0^k)\Big),\end{aligned}$$

where $c_0$ and $r_0^j$ are defined in (6.5.218) and (6.4.182). In other words the terms that are linear in $\xi$ are given by $\widetilde{\omega} = \omega^{-1} + \omega^{(0)}(\xi)$, defined in (6.2.16), and $j^2 + c_0 j^2 + r_0^j = \Omega_j^{\rm int}$ defined in (6.6.275). In order to get (6.6.301) we need a lower bound on the lipschitz semi- norm $|\psi|^{lip}$ (as done in [31]). First of all assume that $\omega^{-1} \cdot l + \sigma j^2 - \sigma' k^2 = 0$. Then the (6.6.301) becomes

$$\psi := {\rm i} A l \cdot \xi + {\rm i}\Big((c_{n+1} - c_0)(\sigma j^2 - \sigma k^2) + (\sigma(r_j^{(n)} - r_0^j) - \sigma'(r_k^{(n)} - r_0^k)\Big), \tag{6.6.302}$$

where by formula (6.6.286) we have set

$$A = -(1/4)(\mathtt{a}_1 \mathbb{1} + (\mathtt{a}_3 - \mathtt{a}_2 - \mathtt{b}_2 - \mathtt{b}_3)\mathtt{V}^2 + (-\mathtt{a}_4 + \mathtt{b}_4)\mathtt{V}^4 - \mathtt{a}_6 \mathtt{V}^6).$$

Hence we have for $\xi_1 \neq \xi_2$

$$\frac{|\psi(\xi_1) - \psi(\xi_2)|}{|\xi_1 - \xi_2|} \geq \frac{c}{2d}|l| - C\delta|l| \geq \frac{c'}{d}|l|, \tag{6.6.303}$$

for a suitable pure constant $c' > 0$. To obtain (6.6.303) we used the invertibility of the matrix $A$, equation (6.6.291) and (6.5.250) to estimate the Lipschitz semi-norm of the constants $(r_j^{(n)} - r_0^j)(\xi)$ and $(c_{n+1} - c_0)(\xi)$. This implies that

$$|R_{ljj'}^{\sigma,\sigma'}(n)| \leq C\gamma|\sigma j^2 - \sigma' k^2|\langle l\rangle^{-\tau}\frac{1}{|\psi|^{lip}} \overset{(6.6.303)}{\leq} \widetilde{C}\gamma\langle l\rangle^{-\tau+1}, \tag{6.6.304}$$

that implies (6.6.300). Let us now consider the case $\omega^{-1} \cdot l + \sigma j^2 - \sigma' k^2 := Z \neq 0$. We first prove the following Lemma.

**Lemma 6.6.163.** *Assume that*

$$|l|, |\sigma j^2 - \sigma' k^2| \leq \frac{Z}{\sqrt{\xi}}, \tag{6.6.305}$$

*then* $|\psi| \geq 1/4$.

*Proof.* Since we have $\omega^{-1} \cdot l + \sigma j^2 - \sigma' k^2 \neq 0$ we obtain

$$\begin{aligned} |\psi| &\geq |\omega^{-1} \cdot l + \sigma j^2 - \sigma' k^2| - \Big(|l||\omega_{n+1} - \omega^{-1}| + |\sigma j^2 - \sigma' k^2||c_{n+1} - 1| + |r_j^{(n)}| + |r_k^{(n)}|\Big) \\ &\geq \frac{1}{2} - C_1\frac{\xi}{\sqrt{\xi}} + C_2\xi \geq \frac{1}{4} \end{aligned} \tag{6.6.306}$$

□

Lemma 6.6.163 we have that if (6.6.305) hold, then there is no small divisor, and hence $R_{ljj'}^{\sigma,\sigma'}(n) = \emptyset$. in the last case we rewrite (6.6.301) as

$$\psi := Z + \mathcal{M}^T l \cdot \xi + c_0(\xi)(Z - \omega^{-1} \cdot l) + O(\xi\delta|l|) = Z + A\xi \cdot l + c_0(\xi)Z + O(\xi\delta) \tag{6.6.307}$$

hence one has

$$\frac{|\psi(\xi_1)-\psi(\xi_2)|}{|\xi_1-\xi_2|} \overset{(6.5.218)}{\geq} c|l| - ZC \tag{6.6.308}$$

for some suitable constant $c, C > 0$. Now we use that $|l| \geq Z/\sqrt{\xi}$ to conclude, for $\xi$ small, that one has

$$|\psi|^{lip} \geq c'|l|. \tag{6.6.309}$$

Reasoning as in (6.6.304), we have that (6.6.309) implies the (6.6.300). □

**Conclusions** Proposition 6.6.160 concludes the proof of Theorem 1.2.5. Concerning Theorem 1.2.6 in which the nonlinearity `f` is merely differentiable, one repeats word by word the arguments of Sections 6.1 and 6.2. At this point the vector field in (6.2.14) is defined in the domain (6.2.10) with $s_0 = a_0 \equiv 0$. In this way the norm $\|\cdot\|_{H^p(\mathbb{T}^d_s\times\mathbb{T}_a)}$ is nothing but the Sobolev norm $\|\cdot\|_{H^p(\mathbb{T}^d\times\mathbb{T})} \sim \|\cdot\|_{0,0,p}$ see Remark 6.2.126 and the definition in (3.2.34). Hence one can repeat the proof of Theorem 1.2.5 since the abstract Theorem 3.2.39 holds in Sobolev regularity.

# A. General Tame and Lipschitz estimates

Here we want to illustrate some standard estimates for composition of functions and changes of variables that we use in the paper. We start with classical embedding, algebra, interpolation and tame estimate in Sobolev spaces $H^s := H^s(\mathbb{T}^d, \mathbb{C})$ and $W^{s,\infty} := W^{s,\infty}$, $d \geq 1$.

**Lemma A.164.** *Let $s_0 > d/2$. Then*

*(i)* **Embedding.** $||u||_{L^\infty} \leq C(s_0)||u||_{s_0}$, $\forall\, u \in H^{s_0}$.

*(ii)* **Algebra.** $||uv||_{s_0} \leq C(s_0)||u||_{s_0}||v||_{s_0}$, $\forall\, u, v \in H^{s_0}$.

*(iii)* **Interpolation.** *For $0 \leq s_1 \leq s \leq s_2$, $s = \lambda s_1 + (1-\lambda)s_2$,*

$$||u||_s \leq ||u||_{s_1}^{\lambda}||u||_{s_2}^{1-\lambda}, \quad \forall\, u \in H^{s_2}. \tag{A.1}$$

*Let $a, b \geq 0$ and $p, q > 0$. For all $u \in H^{a+p+q}$ and $v \in H^{b+p+q}$ one has*

$$||u||_{a+p}||v||_{b+q} \leq ||u||_{a+p+q}||v||_b + ||u||_a||v||_{b+p+q}. \tag{A.2}$$

*Similarly, for the $|u|_s^\infty := \sum_{|\alpha|\leq s}||D^\alpha u||_{L^\infty}$ norm, one has*

$$|u|_s^\infty \leq C(s_1, s_2)(|u|_{s_1}^\infty)^\lambda(|u|_{s_2}^\infty)^{1-\lambda}, \quad \forall\, u \in W^{s_2,\infty}, \tag{A.3}$$

*and $\forall\, u \in W^{a+p+q,\infty}$, $v \in W^{b+p+q,\infty}$,*

$$|u|_{a+p}^\infty|v|_{b+q}^\infty \leq C(a,b,p,q)(|u|_{a+p+q}^\infty|v|_b^\infty + |u|_a^\infty|v|_{b+p+q}^\infty). \tag{A.4}$$

*(iv)* **Asymmetric tame product.** *For $s \geq s_0$ one has*

$$||uv||_s \leq C(s_0)||u||_s||v||_{s_0} + C(s)||u||_{s_0}||v||_s, \quad \forall\, u, v \in H^s. \tag{A.5}$$

*(v)* **Asymmetric tame product in $W^{s,\infty}$.** *For $s \geq 0$, $s \in \mathbb{N}$ one has*

$$|uv|_s^\infty \leq \frac{3}{2}||u||_{L^\infty}|v|_s^\infty + C(s)|u|_s^\infty||v||_{L^\infty}, \forall\, u, v \in W^{s,\infty}. \tag{A.6}$$

*(vi)* **Mixed norms asymmetric tame product.** *For $s \geq 0$, $s \in \mathbb{N}$ one has*

$$||uv||_s \leq \frac{3}{2}||u||_{L^\infty}||v||_s + C(s)|u|_{s,\infty}||v||_0, \forall\, u \in W^{s,\infty}, v \in H^s. \tag{A.7}$$

*If $u := u(\lambda)$ and $v := v(\lambda)$ depend in a Lipschitz way on $\lambda \in \Lambda \subset \mathbb{R}$, all the previous statements hold if one replace the norms $||\cdot||_s$, $|\cdot|_s^\infty$ with $||\cdot||_{s,\gamma}$, $|\cdot|_{s,\gamma}^\infty$.*

Now we recall classical tame estimates for composition of functions.

**Lemma A.165. Composition of functions** *Let $f : \mathbb{T}^d \times B_1 \to \mathbb{C}$, where $B_1 := \{y \in \mathbb{R}^m : |y| < 1\}$. it induces the composition operator on $H^s$*

$$\widetilde{f}(u)(x) := f(x, u(x), Du(x), \ldots, D^p u(x)) \tag{A.8}$$

*where $D^k$ denotes the partial derivatives $\partial_x^\alpha u(x)$ of order $|\alpha| = k$.*

*Assume $f \in C^r(\mathbb{T}^d \times B_1)$. Then*

*(i) For all $u \in H^{r+p}$ such that $|u|_{p,\infty} < 1$, the composition operator (A.8) is well defined and*

$$||\widetilde{f}(u)||_r \leq C||f||_{C^r}(||u||_{r+p} + 1), \tag{A.9}$$

*where the constant $C$ depends on $r, p, d$. If $f \in C^{r+2}$, then, for all $|u|_s^\infty, |h|_p^\infty < 1/2$, one has*

$$||\widetilde{f}(u+h) - \widetilde{f}(u)||_r \leq C||f||_{C^{r+1}}(||h||_{r+p} + |h|_p^\infty ||u||_{r+p}), \tag{A.10}$$

$$||\widetilde{f}(u+h) - \widetilde{f}(u) - \widetilde{f}'(u)[h]||_r \leq C||f||_{C^{r+2}}|h|_p^\infty(||h||_{r+p} + |h|_p^\infty||u||_{r+p}).$$

*(ii) the previous statement also hold replacing $||\cdot||_r$ with the norm $|\cdot|_\infty$.*

*Proof.* For the proof see [3] and [48]. □

**Lemma A.166. Lipschitz estimate on parameters** *Let $d \in \mathbb{N}$, $d/2 < s_0 \leq s$, $p \geq 0$, $\gamma > 0$. Let $F : \Lambda \times H^s \to \mathbb{C}$, for $\Lambda \subset \mathbb{R}$, be a $C^1-$map in $u$ satisfying the tame estimates: $\forall\, ||u||_{s_0+p} \leq 1$, $h \in H^{s+p}$,*

$$||F(\lambda_1, u) - F(\lambda_2, u)||_s \leq C(s)|\lambda_1 - \lambda_2|(1 + ||u||_{s+p}), \quad \lambda_1, \lambda_2 \in \Lambda \tag{A.11a}$$

$$\sup_{\lambda \in \Lambda} ||F(\lambda, u)||_s \leq C(s)(1 + ||u||_{s+p}), \tag{A.11b}$$

$$\sup_{\lambda \in \Lambda} ||\partial_u F(\lambda, u)[h]||_s \leq C(s)(||h||_{s+p} + ||u||_{s+p}||h||_{s_0+p}). \tag{A.11c}$$

*Let $u(\lambda)$ be a Lipschitz family of functions with $||u||_{s_0+p,\gamma} \leq 1$. Then one has*

$$||F(\cdot, u)||_{s,\gamma} \leq C(s)(1 + ||u||_{s+p,\gamma}). \tag{A.12}$$

*The same statement holds when the norms $||\cdot||_s$ are replaced by $|\cdot|_s^\infty$.*

*Proof.* We first note that, by (A.11b), one has $sup_\lambda ||F(\lambda, u(\lambda))||_s \leq C(s)(1+||u||_{s+p,\gamma})$. Then, denoting $h = u(\lambda_2) - u(\lambda_1)$, we have

$$\begin{aligned} ||F(\lambda_2,u(\lambda_2)) - F(\lambda_1, u(\lambda_1))||_s &\leq ||F(\lambda_2, u(\lambda_2)) - F(\lambda_1, u(\lambda_2))||_s \\ &+ ||F(\lambda_1, u(\lambda_2)) - F(\lambda_1, u(\lambda_1))||_s \\ &\leq |\lambda_2 - \lambda_1| C(1 + ||u(\lambda_2)||_{s+p}) \\ &+ \int_0^1 ||\partial_u F(u(\lambda_1) + t(u(\lambda_2) - u(\lambda_1)))[h]||_s dt \\ &\overset{(A.11c)}{\leq} |\lambda_2 - \lambda_1| C(1 + ||u(\lambda_2)||_{s+p}) + C(s)\|h\|_{s+p} \\ &+ C(s)||h||_{s_0+p} \int_0^1 ((1-t)||u(\lambda_1)||_{s+p} + t||u(\lambda_2)||_{s+p})\, dt \end{aligned} \tag{A.13}$$

so that

$$\begin{aligned} \gamma \sup_{\substack{\lambda_1,\lambda_2\in\Lambda \\ \lambda_1\neq\lambda_2}} \frac{||F(u(\lambda_1,\lambda_1)) - F(\lambda_2, u(\lambda_2))||_s}{|\lambda_1 - \lambda_2|} &\leq C\gamma(1 + \sup_{\lambda_2\in\Lambda} ||u(\lambda_2)||_{s+p}) \\ &+ C(s)\left[||u||_{s+p,\gamma} + ||u||_{s_0+p,\gamma} \frac{1}{2} \sup_{\lambda_1,\lambda_2} (||u(\lambda_1)||_{s+p,\gamma} + ||u(\lambda_2)||_{s+p,\gamma})\right] \\ &\leq C(s)\left[||u||^2_{s+p,\gamma} + ||u||_{s_0+p,\gamma}||u||_{s+p,\gamma}\right] + C(s)(1 + ||u||_{s+p,\gamma}), \end{aligned}$$

since $||u||_{s_0+p,\gamma} \leq 1$, then the lemma follows. □

In the following we will show some estimates on changes of variables. The lemma is classical, one can see for instance [3].

**Lemma A.167. (Change of variable)** *Let $p : \mathbb{R}^d \to \mathbb{R}^d$ be a $2\pi-$periodic function in $W^{s,\infty}$, $s \geq 1$, with $|p|_1^\infty \leq 1/2$. Let $f(x) = x + p(x)$. Then one has (i) $f$ is invertible, its inverse is $f^{-1}(y) = g(y) = y + q(y)$ where $q$ is $2\pi-$periodic, $q \in W^{s,\infty}(\mathbb{T}^d; \mathbb{R}^d)$ and $|q|_s^\infty \leq C|p|_s^\infty$. More precisely,*

$$|q|_{L^\infty} = |p|_{L^\infty},\ |dq|_{L^\infty} \leq 2|dp|_{L^\infty},\ |dq|_{s-1}^\infty \leq C|dp|_{s-1}^\infty, \tag{A.14}$$

*where the constant $C$ depends on $d, s$.*

*Moreover, assume that $p = p_\lambda$ depends in a Lipschitz way by a parameter $\lambda \in \Lambda \subset \mathbb{R}$, an suppose, as above, that $|d_x p_\lambda|_{L^\infty} \leq 1/2$ for all $\lambda$. Then $q = q_\lambda$ is also Lipschitz in $\lambda$, and*

$$|q|_{s,\gamma}^\infty \leq C\left(|p|_{s,\gamma}^\infty + \left[\sup_{\lambda\in\Lambda} |p_\lambda|_{s+1}^\infty\right] |p|_{L^\infty,\gamma}\right) \leq C|p|_{s+1,\gamma}^\infty, \tag{A.15}$$

*the constant $C$ depends on $d, s$ (it is independent on $\gamma$).*

*(ii) If $u \in H^s(\mathbb{T}^d;\mathbb{C})$, then $u \circ f(x) = u(x+p(x)) \in H^s$, and, with the same $C$ as in (i) one has*

$$||u \circ f||_s \leq C(||u||_s + |dp|_{s-1}^{\infty}||u||_1), \tag{A.16a}$$

$$||u \circ f - u||_s \leq C(|p|_{L^\infty}||u||_{s+1} + |p|_s^{\infty}||u||_2), \tag{A.16b}$$

$$||u \circ f||_{s,\gamma} \leq C(|u|_{s+1,\gamma} + |p|_{s,\gamma}^{\infty}||u||_{2,\gamma}). \tag{A.16c}$$

*The (A.16a), (A.16b) and (A.16c) hold also for $u \circ g$.*

*(iii) Part (ii) also holds with $||\cdot||_s$ replaced by $|\cdot|_s^{\infty}$, and $||\cdot||_{s,\gamma}$ replaced by $|\cdot|_{s,\gamma}^{\infty}$, namely*

$$|u \circ f|_s^{\infty} \leq C(|u|_s^{\infty} + |dp|_{s-1}^{\infty}|u|_1^{\infty}), \tag{A.17a}$$

$$|u \circ f|_{s,\gamma}^{\infty} \leq C(|u|_{s+1,\gamma}^{\infty} + |dp|_{s-1,\gamma}^{\infty}|u|_{2,\gamma}^{\infty}). \tag{A.17b}$$

**Lemma A.168. (Composition).** *Assume that for any $||u||_{s_0+\mu_i,\gamma} \leq 1$ the operator $\mathcal{Q}_i(u)$ satisfies*

$$||\mathcal{Q}_i h||_{s,\gamma} \leq C(s)(||h||_{s+\tau_i,\gamma} + ||u||_{s+\mu_i,\gamma}||h||_{s_0+\tau_i\gamma}), \quad i = 1, 2. \tag{A.18}$$

*Let $\tau := \max\{\tau_1,\tau_2\}$, and $\mu := \max\{\mu_1,\mu_2\}$. Then, for any*

$$||u||_{s_0+\tau+\mu,\gamma} \leq 1, \tag{A.19}$$

*one has that the composition operator $\mathcal{Q} := \mathcal{Q}_1 \circ \mathcal{Q}_2$ satisfies*

$$||\mathcal{Q}h||_{s,\gamma} \leq C(s)(||h||_{s+\tau_1+\tau_2,\gamma} + ||u||_{s+\tau+\mu,\gamma}||h||_{s_0+\tau_1+\tau_2,\gamma}). \tag{A.20}$$

*Proof.* It is sufficient to apply the estimates (A.18) to $\mathcal{Q}_1$ first, then to $\mathcal{Q}_2$ and using the condition (A.19). □

## A.1 Smooth functions and vector fields on the torus

Here we provide some technical results.

The following one is a general result about smooth maps on the torus. First of all, for any $p \geq 0$ and $b \geq 0$ we denote as usual

$$H^p(\mathbb{T}_s^b;\mathbb{C}) := \big\{u = \sum_{l\in\mathbb{Z}^b} u_l e^{\mathrm{i}l\cdot\theta} : \|u\|_{s,p}^2 := \sum_{l\in\mathbb{Z}^b}\langle l\rangle^{2p}|u_l|^2 e^{2s|l|} < \infty\big\}, \tag{A.21}$$

the space of functions which are analytic on the strip $\mathbb{T}_s^b$, Sobolev on its boundary, and have Fourier coefficients $u_l$. By Cauchy formula for analytic complex functions we have that this $u$ is uniquely determined by the values that assume on the edge of the domain i.e $z = x \pm i\sigma s$ where $\sigma \in \{=1, -1\}^b$. We can define a natural norm using the Sobolev norm of the function on the boundary

$$|u|^2_{s,p} := \sum_{\sigma \in \{=1,-1\}^b} \int_{\mathbb{T}^b} \langle \nabla \rangle^{2p} |u(x + i\sigma s)|^2 \tag{A.22}$$

Using the Fourier basis it reads

$$|u|^2_{s,p} := \sum_{\sigma \in \{=1,-1\}^b} \sum_{l \in \mathbb{Z}^b} \langle l \rangle^{2p} |u_l|^2 e^{-2s\sigma \cdot l} .$$

**Lemma A.169.** *The norm $|\cdot|_{s,p}$ and*

$$\|u\|^2_{s,p} := \sum_{l \in \mathbb{Z}^b} \langle l \rangle^{2p} |u_l|^2 e^{2s|l|} \tag{A.23}$$

*are equivalent.*

*Proof.* Let us define the set $A := \{\ell \in \mathbb{Z}^b : \ell_i \geq 0,\ i = 1, \ldots, b\}$ and $B := \{\ell \in \mathbb{Z}^b : \ell_i \leq 0,\ i = 1, \ldots, b\}$. Consider $p = 0$. For $p > 0$ one can follows the same reasoning. One has

$$|u|^2_{s,p} = \sum_{\ell \in A} |u_\ell|^2 e^{2s|\ell|} + \sum_{\ell \in A^c} |u_\ell|^2 e^{2s\ell} + \sum_{\ell \in B} |u_\ell|^2 e^{2s|\ell|} + \sum_{\ell \in B^c} |u_\ell|^2 e^{-2s\ell} \leq 2|A^c| \sum_{\ell \in A} |u_\ell|^2 e^{2s|\ell|} \leq C(b)||u||^2_{s,p}.$$

The opposite inequality is obtained in the same way. □

**Lemma A.170.** *Take $\ell_{a,p}$ of Section 3.1 as $\ell_{a,p} = H^p(\mathbb{T}_a; \mathbb{C})$ and take a map $u : \mathbb{T}^d_s \to \ell_{a,p}$ Define*

$$u_p(\theta) := \|u\|_{a,p}.$$

*Then the norm*

$$A(u) = \|u_p\|_{s,\mathfrak{p}_0} + \|u_{\mathfrak{p}_0}\|_{s,p}, \tag{A.24}$$

*is equivalent to the norm*

$$B(u) = \left( \sum_{l \in \mathbb{Z}^d, j \in \mathbb{Z}} \max\{1, |l|, |j|\}^{2(p+\mathfrak{p}_0)} e^{2s|l|} e^{2a|j|} |u_{l,j}|^2 \right)^{\frac{1}{2}}, \quad p \geq 0. \tag{A.25}$$

*Proof.* Note that the norm in (A.24) is nothing but the norm in (3.2.37) where $\ell_{a,p}$ is the Sobolev space $H^p(\mathbb{T}_a, \mathbb{C})$ The norm in (A.25) is nothing but the norm defined in (A.21) for $b = d + 1$ and different strip of analyticity for the variables $\theta$ and $x$. Assuming that the index $p \in \mathbb{N}$ we first prove that $A(u)$ is equivalent with

$$C(u) := \sum_{\substack{p_1 + p_2 = p + \mathfrak{p}_0 \\ p_1, p_2 \geq \mathfrak{p}_0}} \left( \sum_{l \in \mathbb{Z}^d, j \in \mathbb{Z}} |u_{l,j}|^2 \langle j \rangle^{2p_1} \langle l \rangle^{2p_2} e^{2s|l|} e^{2a|j|} \right)^{\frac{1}{2}}, \quad p \geq 2\mathfrak{p}_0 \tag{A.26}$$

Explicitly one has that

$$A(u) := \left(\sum_{l\in\mathbb{Z}^d, j\in\mathbb{Z}} |u_{l,j}|^2 \langle j\rangle^{2p}\langle l\rangle^{2\mathfrak{p}_0} e^{2s|l|}e^{2a|j|}\right)^{\frac{1}{2}} + \left(\sum_{l\in\mathbb{Z}^d, j\in\mathbb{Z}} |u_{l,j}|^2 \langle j\rangle^{2\mathfrak{p}_0}\langle l\rangle^{2p} e^{2s|l|}e^{2a|j|}\right)^{\frac{1}{2}}.$$

Clearly one inequality is obvious. The other comes form Young's Inequality and

$$\langle j\rangle^{2p_1}\langle l\rangle^{2p_2} = \langle j\rangle^{2\mathfrak{p}_0}\langle l\rangle^{2\mathfrak{p}_0}\langle j\rangle^{2p_1-2\mathfrak{p}_0}\langle l\rangle^{2p_2-2\mathfrak{p}_0} \le \langle j\rangle^{2\mathfrak{p}_0}\langle l\rangle^{2\mathfrak{p}_0}\left(\frac{2(p_1-\mathfrak{p}_0)}{2(p-\mathfrak{p}_0)}\langle j\rangle^{2(p-\mathfrak{p}_0)} + \frac{2(p_2-\mathfrak{p}_0)}{2(p-\mathfrak{p}_0)}\langle l\rangle^{2(p-\mathfrak{p}_0)}\right)$$

Hence $C(u) \le C(p)A(u)$ with $C(p)$ a constant depending on $p$. Clearly $B(u)$ and $A(u)$ are equivalent, hence we have the Lemma follows. □

The following Lemma resume some important properties in Sobolev spaces $H^s := H^s(\mathbb{T}^b;\mathbb{C})$ with *norm*

$$\|u\|_s^2 := \sum_{l\in\mathbb{Z}^b} \langle l\rangle^{2p}|u_l|^2.$$

The same results of Lemma A.164 holds also for our analytic norm in (A.21).

We now introduce the space

$$W^{p,\infty}(\mathbb{T}^b_\zeta) := \left\{\beta : \mathbb{T}^b_\zeta \to \mathbb{T}^b_\zeta \ : \ |\beta|_{p,\zeta,\infty} := \sum_{k=0}^{p} \|d^k\beta\|_{L^\infty(\mathbb{T}^n_s)} < \infty\right\}, \tag{A.27}$$

and note that one has $H^{\zeta,p+\mathfrak{p}_0}(\mathbb{T}^b_\zeta) \subset W^{p,\infty}(\mathbb{T}^b_\zeta)$.

**Lemma A.171 (Diffeomorphism).** *Let $\beta \in W^{p,\infty}(\mathbb{T}^b_\zeta)$ for some $p, \zeta \ge 0$ such that*

$$\|\beta\|_{\zeta,\mathfrak{p}_0} \le \frac{\delta}{2C_1}, \quad \|\beta\|_{\zeta,\mathfrak{p}_0} \le \frac{1}{2C_2}, \quad 0 < \delta < \frac{\zeta}{2}, \ \ C_1, C_2 > 0, \tag{A.28}$$

*and let us consider $\Phi : \mathbb{T}^b_\zeta \to \mathbb{T}^b_{2\zeta}$ of the form*

$$x \mapsto x + \beta(x) = \Phi(x). \tag{A.29}$$

*Then the following is true.*

*(i) There exists $\Psi : \mathbb{T}^b_{\zeta-\delta} \to \mathbb{T}^b_\zeta$ of the form $\Psi(y) = y + \widetilde{\beta}(y)$ with $\widetilde{\beta} \in W^{p,\infty}(\mathbb{T}^b_{\zeta-\delta})$ satisfying*

$$\|\widetilde{\beta}\|_{\zeta-\delta,\mathfrak{p}_0} \le \frac{\delta}{2}, \quad \|\widetilde{\beta}\|_{\zeta-\delta,p} \le 2\|\beta\|_{\zeta,p}, \tag{A.30}$$

*such that for all $x \in \mathbb{T}^b_{\zeta-2\delta}$ one has $\Psi \circ \Phi(x) = x$.*

*(ii) For all $u \in H^{\zeta,p}(\mathbb{T}^b_\zeta)$, the composition $(u \circ \Phi)(x) = u(x + \beta(x))$ satisfies*

$$\|u \circ \Phi\|_{\zeta-\delta,p} \le C(\|u\|_{\zeta,p} + |d\beta|_{p-1,\zeta,\infty}\|u\|_{\zeta,\mathfrak{p}_0}). \tag{A.31}$$

*Proof.* For $\zeta = 0$ the result is proved in A.167 thus in the following we assume $\zeta > 0$.
*(i)* First of all recall that, if $\mathfrak{p}_0 \geq b/2$ then $\|u\|_{L^\infty} \leq \|u\|_{\zeta,\mathfrak{p}_0}$. We look for $\widetilde{\beta}$ such that

$$\widetilde{\beta}(y) = -\beta(y + \widetilde{\beta}(y)). \tag{A.32}$$

The idea is to rewrite the problem as a fixed point equation. We define the operator $\mathcal{G} : H^{\zeta,p} \to H^{\zeta,p}$ as $\mathcal{G}(\widetilde{\beta}) = -\beta(y+\widetilde{\beta})$. First of all we need to show that $\mathcal{G}$ maps the ball $B_{\delta/2} := \{\|u\|_{\zeta-\delta,p} < \delta/2\}$ into itself.

One has

$$\|\mathcal{G}(\widetilde{\beta})\|_{\zeta-\delta,\mathfrak{p}_0} = \left\| \sum_{n\geq 0} \frac{1}{n!} (\partial^n \beta)\widetilde{\beta}^n \right\|_{\zeta-\delta,\mathfrak{p}_0} \leq \sum_{n\geq 0} \frac{1}{n!} \|\beta\|_{\zeta-\delta,\mathfrak{p}_0+n} \|\widetilde{\beta}\|^n_{\zeta-\delta,\mathfrak{p}_0}\,, \tag{A.33}$$

where $\partial\beta$ denotes the derivative of $\beta$ w.r.t. its argument. Note that for any $u \in H^{\zeta+\delta,s}$ and $\tau > 0$ one has

$$\|u\|_{\zeta,s+\tau} \leq \left(\frac{\tau}{e}\right)^\tau \frac{1}{\delta^\tau} \|u\|_{\zeta+\delta,s}\,; \tag{A.34}$$

indeed

$$\|u\|^2_{\zeta,p+\tau} = \sum_{l\in\mathbb{Z}^b} \langle l\rangle^{2(p+\tau)} e^{2\zeta|l|} |u_l|^2 \leq \sum_{l\in\mathbb{Z}^b} \langle l\rangle^{2p} |l|^{2\tau} e^{-2\delta(|l|)} e^{2(\zeta+\delta)|l|} |b_l|^2\,,$$

and the function $f(x) := x^{2\tau} e^{-2\delta x}$ reach its maximum at $x = \tau/\delta$ and $f(\tau/\delta) = (\tau/\delta e)^{2\tau}$, so that (A.34) follows. Then using the (A.34) and the fact that $n! = (1/\sqrt{2\pi n})(n/e)^n(1 + O(1/n))$ as $n \to \infty$, we obtain

$$\begin{aligned} \|\mathcal{G}(\widetilde{\beta})\|_{\zeta-\delta,\mathfrak{p}_0} &\leq \sum_{n\geq 0} \frac{1}{n!} \left(\frac{n}{e}\right)^n \frac{1}{\delta^n} \|\beta\|_{\zeta,\mathfrak{p}_0} \|\widetilde{\beta}\|^n_{\zeta-\delta,\mathfrak{p}_0} \leq \|\beta\|_{\zeta,\mathfrak{p}_0} \sum_{n\geq 0} C \left( \frac{\|\widetilde{\beta}\|_{\zeta-\delta,\mathfrak{p}_0}}{\delta} \right)^n \\ &\leq 2C\|\beta\|_{\zeta,\mathfrak{p}_0} \stackrel{(A.28)}{\leq} \frac{\delta}{2}\,. \end{aligned} \tag{A.35}$$

Finally we show that $\mathcal{G}$ is a contraction. One has

$$\begin{aligned} \|\mathcal{G}(\widetilde{\beta}_1) - \mathcal{G}(\widetilde{\beta}_2)\|_{\zeta-\delta,p} &= \left\| \sum_{n\geq 1} \frac{1}{n!} (\partial^n \beta)\widetilde{\beta}_1^n - \sum_{n\geq 1} \frac{1}{n!} (\partial^n \beta)\widetilde{\beta}_2^n \right\|_{\zeta-\delta,p} \\ &= \left\| \sum_{n\geq 1} \frac{1}{n!} (\partial^n \beta)(\widetilde{\beta}_1 - \widetilde{\beta}_2) \left( \sum_{k=0}^{n-1} \widetilde{\beta}_1^k \widetilde{\beta}_2^{n-1-k} \right) \right\|_{\zeta-\delta,p} \\ &\leq \|\widetilde{\beta}_1 - \widetilde{\beta}_2\|_{\zeta-\delta,p} \sum_{n\geq 1} \frac{1}{n!} \left(\frac{n}{e}\right)^n \frac{1}{\delta^n} \|\beta\|_{\zeta,p\|} \sum_{k=0}^{n-1} \|\widetilde{\beta}_1\|^k_{\zeta-\delta,p} \|\widetilde{\beta}_2\|^{n-1-k}_{\zeta-\delta,s} \\ &\leq \|\widetilde{\beta}_1 - \widetilde{\beta}_2\|_{\zeta-\delta,p} C_2 \|\beta\|_{\zeta-\delta,p} \stackrel{(A.28)}{\leq} \frac{1}{2} \|\widetilde{\beta}_1 - \widetilde{\beta}_2\|_{\zeta-\delta,p}\,. \end{aligned} \tag{A.36}$$

Then we deduce that there exists a unique fixed point in $B_{\delta/2}$, hence a solution of the equation (A.31).
*(ii)* One can follow almost word by word the proof of Lemma 11.4 in [?] using the norm (A.22) instead of (A.23) and the interpolation properties of the $W^{p,\infty}(\mathbb{T}^b_\zeta)$-norms. ■

**Remark A.172.** *Note that by Lemma A.169, one has*

$$\|V^{(\theta)}(\theta)\cdot\partial_\theta\|_{s,a,p} \approx \frac{1}{s_0}\max_{1\le i\le d}\sum_{\sigma\in\{\pm1\}^d}\|V^{(\theta_i)}(\mathrm{Re}(\theta)+\mathrm{i}\sigma s)\|_{H^p}\,,$$

$$\|V^{(y)}(\theta)\cdot\partial_y\|_{s,a,p} \approx \frac{1}{r_0^2}\sum_{i=1}^{d}\sum_{\sigma\in\{\pm1\}^d}\|V^{(y_i)}(\mathrm{Re}(\theta)+\mathrm{i}\sigma s)\|_{H^p}\,,$$

$$\|V^{(w)}(\theta)\partial_w\|_{s,a,p} \approx \frac{1}{r_0}\sum_{\sigma\in\{\pm1\}^d}\sum_{p_1+p_2=p}\|\|V^{(w)}(\mathrm{Re}(\theta)+\mathrm{i}\sigma s)\|_{a,p_1}\|_{H^{p_2}}\,,$$

*In particular this means that for all $s\ge 0$, $a\ge 0$ and $p\ge\overline{\mathfrak{p}}>n/2$ one has the standard algebra, interpolation and tame properties; see for instance [12, 4, 31, 14] just to mention a few.*

From Lemma A.171 above we deduce the following result; from now on we drop the labels $s,a$ in the scales of constants since all the results below do not depend on them, i.e. we write $C_p(F)=C_{s,a,p}(F)$.

**Lemma A.173.** *Given a tame vector field $F\in\mathcal{V}_{a,p}(s,r)$ with scale of constants $C_p(F)$ of the form* (3.2.39) *and given a map $\Phi(\theta)=\theta+\beta(\theta)$ as in* (A.29) *with $b=d$ and $\zeta=s$, then the composition $F\circ\Phi$ is a tame vector field with constant*

$$C_p(F\circ\Phi)\le C_p(F)+C_{\mathfrak{p}_0}(F)\|\beta\|_{s,p+\nu+3} \tag{A.37}$$

*Proof.* For simplicity we drop the indexes $\lambda$ and $\mathcal{O}$ in the tameness constants. By Lemma A.171 one has that if $\|\beta\|_{s,\mathfrak{p}_0}$ is sufficiently small, then the vector field $F\circ\Phi$ is defined on $D_{a,p}(s-\rho s_0, r-\rho r_0)$. Lemma A.171 guarantess that for a function $u(\theta)\in\mathbb{C}$ the estimate (A.31) holds. Hence also the components $F^{(v)}(\theta,y,w)$ for $v=\theta,y$ satisfy the same bounds (recall that for the norm (3.2.34) $y,w$ are simply parameters). Let us study the composition of $F^{(w)}(\theta+\beta(\theta),y,w)$.

We define the function $u_{p_1}:\mathbb{T}^d_s\to\mathbb{C}$ as $u_{p_1}(\theta)=\|F^{(w)}\|_{a,p_1}$. Hence one has that

$$\|F^{(w)}\|_{s,a,p}=\frac{1}{r_0}(\|u_{\mathfrak{p}_0}\|_{s,p}+\|u_p\|_{s,\mathfrak{p}_0}).$$

Consider the composition with the diffeomorphism $\Phi$, then one has

$$\|F^{(w)}\circ\Phi\|_{s,a,p}\le\frac{1}{r_0}\left(\|u_{\mathfrak{p}_0}\|_{s,p}+|\beta|_{p,\xi,\infty}\|u_{\mathfrak{p}_0}\|_{s,\mathfrak{p}_0}+C\|u_p\|_{s,\mathfrak{p}_0}\right)\le\frac{C}{r_0}(\|F^{(w)}\|_{s,a,p}+|\beta|_{p,\xi,\infty}\|F^{(w)}\|_{s,a,\mathfrak{p}_0})$$

By the estimates on the norm we get the estimates on the tameness constants. □

# B. Properties of tame vector fields

We collects some properties of vector fields that are "tame" in the sense of Definition 3.2.26. These results are used in Section 3.1.18 in order to prove Theorem 3.2.39. First we set up some notations.

**Notations.** Given a map $f : \mathbb{T}^d_s \to \mathbb{C}^d$ $f(\theta) \cdot \partial_\theta :\ \ g \mapsto dg[f] = \sum_i f_i(\theta)\partial_{\theta_i} g$. The same holds for $F(\theta)y \cdot \partial_y$ when $F$ maps $\mathbb{T}^d_s$ into $\mathcal{L}(\mathbb{C}^d, \mathbb{C}^d)$ ( the $d \times d$ complex matrices).

Consider now a map $M$ from $\mathbb{T}^d_s$ into $\overbrace{\ell^*_{a,q} \times \ldots \times \ell^*_{a,q}}^{d}$ so $M(\theta) = \{M_i(\theta)\}_{i=1}^d$ with $M_i \in \ell^*_{a,q}$. The symbol $(M(\theta)[w])$ is a linear map in $\mathcal{L}(\ell_{a,q}, \mathbb{C}^d)$ and so $(M(\theta) \cdot w) \cdot \partial_y$ is well defined as above.

Finally given a map $W$ from $\mathbb{T}^d_s$ into $\ell_{a,q}$, the vector field $W(\theta) \cdot \partial_w$ acts on functions $g : \ell_{a,q} \to \ell_{a,q}$ as $g \mapsto dg[W] = \nabla g \cdot W$. Then given a map $L$ from $\mathbb{T}^d_s$ into $\mathcal{L}(\ell_{a,q}, \ell_{a,q})$ the vector field $Lw \cdot \partial_w$ is defined in the same way.

It is important to note that our operators depend on the angles $\theta$ and it is more convenient describe them in terms of the Fourier coefficients. More precisely we define

$$
\begin{aligned}
H^p(\mathbb{T}^d_s; \ell_{a,q}) &:= \{u = \{u_l\}_{l \in \mathbb{Z}^d} : u_l \in \ell_{a,q} \ \ \|u\|_{s,a,p} < \infty\}\,, \\
H^p(\mathbb{T}^d_s; \ell^*_{a,q}) &:= \left\{ f = \{f_l[\cdot]\}_{l \in \mathbb{Z}^d} : f_l[\cdot] \in \ell^*_{a,q} \ \ \sum_{l \in \mathbb{Z}^d} |f_l[\cdot]|^2 \langle l \rangle^{2p} e^{2s|l|} < \infty \right\},
\end{aligned}
\tag{B.1}
$$

The action of a linear operator $M$ on the space $H^q(\mathbb{T}^d_s; \ell_{a,p})$ is described, in Fourier coefficients, by

$$
Mu = M\{u_l\}_{l \in \mathbb{Z}^d} := \left\{ \sum_{l' \in \mathbb{Z}^d} A_l^{l'} u_{l'} \right\}_{l \in \mathbb{Z}^d}
\tag{B.2}
$$

**Definition B.174.** *Given a map $M$ as in* (B.2) *we define an operator $M^\dagger$ on the space $H^p(\mathbb{T}^d_s; \ell^*_{a,q})$ as follows: for any $w \in \ell_{a,q}$ one has*

$$
M^\dagger f[w] := f[Mw] = \left\{ \sum_{l' \in \mathbb{Z}^d} f_{l'}[M_l^{l'} w] \right\}_{l \in \mathbb{Z}^d}
\tag{B.3}
$$

**Remark B.175.** *By Definition B.174 follows that if an operator $M$ is left-invertible, i.e. there exists $L$ such that*

$$LMu = u \in H^p(\mathbb{T}^d_s; \ell_{a,q}),$$

*then the operator $L^\dagger$ is a right-inverse for $M^\dagger$, i.e.*

$$M^\dagger L^\dagger f[w] = f[w] \in H^p(\mathbb{T}^d_s; \ell^*_{a,q})$$

**Remark B.176.** *Definition B.174 is quite natural. Indeed, if the matrix $\{M_l^{l'}\}_{l,l'\in\mathbb{Z}^d}$ is the representation of an operator the is a function of $\theta$ one has that*

$$M^\dagger(\theta) f[w] = f[M(\theta)w] = -\mathrm{ad}(Mw\cdot\partial_w)(f[w]\cdot\partial_y) = [Mw\cdot\partial_w, f[w]\cdot\partial_y].$$

*Clearly Definition B.5 is more general.*

**Proof of Lemma 3.2.31 (Conjugation).** By (3.2.42) one has that the vector field $G$ is defined in $D_{a,\mathfrak{p}_0}(s-2\rho s_0, r-2\rho r_0)$. Then, given a change of coordinates $\Gamma$ which maps $D_{a,\mathfrak{p}_0}(s-3\rho s_0, r-3\rho r_0)$ into $D_{a,\mathfrak{p}_0}(s-2\rho s_0, r-2\rho r_0)$ we can consider the composition of $G$ with $\Gamma$. Let us check the property $(F1)$ for the vector field $G$. In the following to simplify the notation we will drop the indices $a, s, \lambda$ in the norms since they are essentially fixed. One has

$$\begin{aligned}
\|G(\Gamma)\|_p &\le \|F(\Phi^{-1}(\Gamma))\|_p + \|df(\Phi^{-1}(\Gamma))[F(\Phi^{-1}(\Gamma))]\|_p \\
&\overset{(F2)}{\le} (1+C_{\mathfrak{p}_0+1}(f))\|F(\Phi^{-1}(\Gamma))\|_p + \left(C_{p+1}(f) + C_{\mathfrak{p}_0+1}(f)\|\Phi^{-1}(\Gamma)\|\right)\|F(\Phi^{-1}(\Gamma))\|_{\mathfrak{p}_0} \\
&\overset{(F1)}{\le} (1+C_{\mathfrak{p}_0+1}(f))\left[C_p(F) + C_{\mathfrak{p}_0}(F)\|\Phi^{-1}(\Gamma)\|_{p+\nu}\right] \\
&+ \left(C_{p+1}(f) + C_{\mathfrak{p}_0+1}(f)\|\Phi^{-1}(\Gamma)\|_p\right)\left[C_{\mathfrak{p}_0}(F) + C_{\mathfrak{p}_0}(F)\|\Phi^{-1}(\Gamma)\|_{\mathfrak{p}_0+\nu}\right],
\end{aligned} \tag{B.4}$$

and moreover

$$\begin{aligned}
&\|\Phi^{-1}(\Gamma)\|_p \le C_p(f) + (1+C_{\mathfrak{p}_0}(f))\|\Gamma\|_p, \\
&\|\Phi^{-1}(\Gamma)\|_{\mathfrak{p}_0+\nu} \le 1 + 2C_{\mathfrak{p}_0+\nu}(f).
\end{aligned} \tag{B.5}$$

Hence by (B.4) and (B.5) we obtain

$$\begin{aligned}
\|G(\Gamma)\|_p \le\, & C_p(F)(1+C_{\mathfrak{p}_0+\nu+1}(f)) + 5C_{\mathfrak{p}_0}(F)(1+C_{\mathfrak{p}_0+1}(f))C_{p+\nu+1}(f) \\
&+ \|\Gamma\|_{p+\nu}\left[C_{\mathfrak{p}_0}(F)(1+3C_{\mathfrak{p}_0+\nu+1}(f))^2\right],
\end{aligned} \tag{B.6}$$

that is the $(F1)$. The other properties are obtained with similar calculations using also the fact that the vector field $f$ is linear in the variables $y, w$. Hence $G$ is tame with scale of constants in (3.2.49). $\square$

**Lemma B.177.** *All $f$ as in Definition 3.2.34 are tame with tameness constant*

$$C_{\vec{v},p}(f) = C_{d,q}|f|_{\vec{v},p}, \tag{B.7}$$

*where $|\cdot|_{\vec{v},p}$ is defined in Definition 3.2.34.*

*Proof.* Consider a linear vector field $f$ (see (3.2.34)) and a map $\Phi = \mathbb{1} + g$ as in Definition (3.2.26). For simplicity we drop the indeces $\vec{v}, \vec{v}_1, \vec{v}_2$. Without loss of generality we can also assume that $g^{(\theta)}$ depends only on $\theta$ . Let us check the $(F1)$ for $f$. One has that

$$\begin{aligned}
&(f\circ\Phi)^{(\theta)} := h^{(\theta,0)}(\theta), \qquad (f\circ\Phi)^{(w)} := h^{(w,0)}(\theta),\\
&(f\circ\Phi)^{(y)} := h^{(y,0)}(\theta) + h^{(y,y)}(\theta)\Phi^{(y)}(\theta,y,w) + h^{(y,w)}(\theta)\Phi^{(w)}(\theta,y,w)
\end{aligned}$$

where

$$h^{(v,v')}(\theta) := f^{(v,v')}(\theta + g^{(\theta,0)}(\theta)), \quad v, v' = \theta, y, w.$$

We first give bounds on the norm of $f\circ\Phi$ in terms of the norms of $h$ and $\Phi$. Then we need a Lemma to estimate the norms of the composition with diffeomorphisms of the torus. For the $\theta$ and $w$ components it is trivial that the tameness constant is the norm of the function. Concerning the $y-$components we will give an explicit estimates only for the non trivial terms. One has

$$\begin{aligned}
\|h^{(y,y)}\Phi^{(y)}\|^2_{s,a,p} &\le C(d)\frac{1}{r_0^4}\sum_{\ell\in\mathbb{Z}^d}\sum_{i=1}^{d}\sum_{k=1}^{d}|(h^{(y_i,y_k)}g^{(y_k)})(\ell)|^2 e^{2s|\ell|}\langle\ell\rangle^{2p}\\
&= C(d)\frac{1}{r_0^4}\sum_{i=1}^{d}\sum_{k=1}^{d}\|h^{(y_i,y_k)}(\theta)\Phi^{(y_k)}(\theta)\|^2_{s,p}\\
&\overset{(A.5)}{\le} C(d)\frac{1}{r_0^4}\sum_{i=1}^{d}\sum_{k=1}^{d}(\|h^{(y_i,y_k)}\|_{s,p}\|\Phi^{(y_k)}\|_{s,\mathfrak{p}_0} + \|h^{(y_i,y_k)}\|_{s,\mathfrak{p}_0}\|\Phi^{(y_k)}\|_{s,p})^2,
\end{aligned} \tag{B.8}$$

hence one obtains

$$\|h^{(y,y)}\Phi^{(y)}\|_{s,a,p} \le K\left(r_0^2\|h^{(y,y)}\|_{s,a,p}\|\Phi^{(y)}\|_{s,a,\mathfrak{p}_0} + r_0^2\|h^{(y,y)}\|_{s,a,\mathfrak{p}_0}\|\Phi^{(y)}\|_{s,a,p}\right). \tag{B.9}$$

Finally one has

$$
\begin{aligned}
\|h^{(y,w)}\Phi^{(w)}\|_{s,a,p}^2 &\le C(d)\frac{1}{r_0^4}\sum_{i=1}^d \|h^{(y_i,w)}\Phi^{(w)}\|_{s,p}^2 = C(d)\frac{1}{r_0^4}\sum_{i=1}^d\sum_{l\in\mathbb{Z}^d}\langle l\rangle^{2p}e^{2s|l|}|(h^{(y_i,w)}\Phi^{(w)})(l)|^2 \\
&\le C(d)\frac{1}{r_0^4}\sum_{i=1}^d\sum_{l\in\mathbb{Z}^d}\langle l\rangle^{2p}e^{2s|l|}(\sum_{k\in\mathbb{Z}^d}|h^{(y_i,w)}(l-k)\Phi^{(w)}(k)|)^2 \\
&\le \frac{C(d,\mathfrak{p}_0)}{r_0^4}\sum_{i=1}^d\sum_{l,k\in\mathbb{Z}^d}\langle l-k\rangle^{2p}e^{2s|l-k|}\langle k\rangle^{2\mathfrak{p}_0}e^{2s|k|}|h^{(y_i,w)}(l-k)\Phi^{(w)}(k)|^2 \\
&+ \frac{C(d,\mathfrak{p}_0)}{r_0^4}\sum_{i=1}^d\sum_{l,k\in\mathbb{Z}^d}\langle l-k\rangle^{2\mathfrak{p}_0}e^{2s|l-k|}\langle k\rangle^{2p}e^{2s|k|}|h^{(y_i,w)}(l-k)\Phi^{(w)}(k)|^2 \\
&\overset{(3.2.29)}{\le} \frac{C(d,\mathfrak{p}_0)}{r_0^4}\sum_{i=1}^d\sum_{l,k\in\mathbb{Z}^d}\langle l-k\rangle^{2p}e^{2s|l-k|}\langle k\rangle^{2\mathfrak{p}_0}e^{2s|k|}\|h^{(y_i,w)}(l-k)\|_{a,p}^2\|\Phi^{(w)}(k)\|_{a,\mathfrak{p}_0}^2 \\
&+ \frac{C(d,\mathfrak{p}_0)}{r_0^4}\sum_{i=1}^d\sum_{l,k\in\mathbb{Z}^d}\langle l-k\rangle^{2p}e^{2s|l-k|}\langle k\rangle^{2\mathfrak{p}_0}e^{2s|k|}\|h^{(y_i,w)}(l-k)\|_{a,\mathfrak{p}_0}^2\|\Phi^{(w)}(k)\|_{a,p}^2,
\end{aligned}
\tag{B.10}
$$

where we used the fact that $\mathfrak{p}_0 > d/2$. By (B.10) follows that

$$
\|h^{(y,w)}\Phi^{(w)}\|_{s,a,p} \le C(d,\mathfrak{p}_0)\left(r_0\|h^{(y,w)}\|_{s,a,p}\|\Phi^{(w)}\|_{s,a,\mathfrak{p}_0} + r_0\|h^{(y,w)}\|_{s,a,\mathfrak{p}_0}\|\Phi^{(w)}\|_{s,a,p},\right). \tag{B.11}
$$

Now set $\Phi^{(v)} := v + g^{(v)}(\theta, y, w)$ for $v = \theta, y, w$. Then, by the discussion above, we have

$$
\begin{aligned}
\|(f\circ\Phi)^{(y)}\|_{s,a,p} &\le \|h^{(y,0)}\|_{s,a,p} + \|h^{(y,w)}\|_{s,a,p}\|\Phi^{(w)}\|_{s,a,\mathfrak{p}_0} + \|h^{(y,w)}\|_{s,a,\mathfrak{p}_0}\|\Phi^{(w)}\|_{s,a,p} \\
&\le \|h^{(y,y)}\|_{s,a,p}\|\Phi^{(y)}\|_{s,a,\mathfrak{p}_0} + \|h^{(y,y)}\|_{s,a,\mathfrak{p}_0}\|\Phi^{(y)}\|_{s,a,p}
\end{aligned}
\tag{B.12}
$$

Now each component of the vector fields $h$ is a function of $\theta$ composed with a diffeomorphism of the torus given by $\theta \mapsto \theta + g^{(\theta,0)}(\theta)$. Hence we obtain, by using $(ii)$ of Lemma A.171,

$$
\|f\circ\Phi\|_{s,a,p} \le C(|f|_{s,a,p} + |f|_{s,a,\mathfrak{p}_0}\|\Phi\|_{s,a,p+\nu}), \tag{B.13}
$$

with a constant depending only on $d, \mathfrak{p}_0$. □

**Lemma B.178.** *Fix $\mathfrak{p}_1$. Given any vector field $g \in \mathcal{B}$, $p \ge \mathfrak{p}_1$ with $|g|_{\vec{v},\mathfrak{p}_1} \le \rho$ then for $0 \le t \le 1$ there exists $f_t \in \mathcal{B}$ such that the time$-t$ map of the flow of $g$ is of the form $\mathbb{1} + f_t$ moreover we have $|f_t|_{\vec{v},p} \le 2|g|_{\vec{v},p}$.*

**Lemma B.179.** *For all $s, a \ge 0$, $\lambda \in (0,1)$, consider a vector field $f \in \mathcal{B}$ such that*

$$
f : \mathcal{O} \times D_{a',p}(r') \times \mathbb{T}_s^d \to V_{a,p} \tag{B.14}
$$

*and*

$$|f|_{\vec{v},\mathfrak{p}_1} \leq \frac{\rho}{K}, \tag{B.15}$$

*for some $\rho > 0$, $\mu > 0$, and $K = K(\mathfrak{p}_0, d) \geq 1$. If $\rho$ is small enough, then for all $\xi \in \mathcal{O}$ the following holds.*

*(i) The map $\Phi := \mathbb{1} + f$ is such that*

$$\Phi : \mathcal{O} \times D_{a,p}(r) \times \mathbb{T}^d_s \longrightarrow D_{a,p}(r + \rho r_0) \times \mathbb{T}^d_s. \tag{B.16}$$

*(ii) There exists a vector field $g \in \mathcal{B}$ such that*

- $|g|_{\vec{v},p} \leq 2|f|_{\vec{v},p}$, *the map $\Psi := \mathbb{1} + g$ is such that*

$$\Psi : \mathcal{O} \times D_{a,p}(r - \rho r_0) \times \mathbb{T}^d_s \to D_{a,p}(r) \times \mathbb{T}^d_s \tag{B.17}$$

- *for all $(\theta, y, w) \in D_{a,\mathfrak{p}_0+\mu}(r - 2\rho r_0, s)$ one has*

$$\Psi \circ \Phi(\theta, y, w) = (\theta, y, w). \tag{B.18}$$

*Proof. (i)* We want to bound the components of $\Phi = \mathbb{1} + f$. First of all we see that for $\theta \in \mathbb{T}^d_s$ one has

$$|\Phi^{(\theta)}|_\infty \leq s + |f^{(\theta)}|_\infty \leq s + \|f^{(\theta)} \cdot \partial_\theta\|_{s,a,\mathfrak{p}_0} \overset{(B.15)}{\leq} s + \rho^2 s_0\,, \tag{B.19}$$

where we used the standard Sobolev embedding Theorem. The bound on $\|\Phi^{(w)}\|_{a,\mathfrak{p}_0} \leq r + \rho^2 r_0$ follows directly by hypothesis (B.15). In order to obtain the estimates on the $y-$components, since the components of $f$ are in 1-to-1 correspondence with those of the vector field $f$ we need to check that

$$\begin{gathered} |f^{(y,0)}(\theta)|_1 \leq K\|f^{(y,0)}(\theta) \cdot \partial_y\|_{s,a,\mathfrak{p}_0}, \qquad |f^{(y,y)}(\theta)y|_1 \leq K\|f^{(y,y)}(\theta)y \cdot \partial_y\|_{s,a,\mathfrak{p}_0}, \\ |f^{(y,w)}(\theta)w|_1 \leq K\|f^{(y,w)}(\theta)w \cdot \partial_y\|_{s,a,\mathfrak{p}_0}, \end{gathered} \tag{B.20}$$

for some constant $K > 0$ depending on $\mathfrak{p}_0$. Since for an $d$-dimensional vector $\mathtt{v}$ one has $|\mathtt{v}|_1 \leq d|\mathtt{v}|_\infty$ we get

$$\begin{aligned} |f^{(y,w)}(\theta)w|_1 &\leq d \max_{v=y_1,\dots,y_d} |f^{(v,w)}(\theta)w| \leq d \max_{v=y_1,\dots,y_d} \sup_{\theta\in\mathbb{T}^d_s} |f^{(y,w)}(\theta)w| = d \max_{v=y_1,\dots,y_d} \|f^{(v,w)}(\theta)w\|_\infty \\ &\leq K(n,\mathfrak{p}_0)\|f^{(y,w)}(\theta)w\|_{s,\mathfrak{p}_0} \leq K(d,\mathfrak{p}_0)\|f^{(y,w)}(\theta)w \cdot \partial_y\|_{s,a,\mathfrak{p}_0}\,. \end{aligned} \tag{B.21}$$

The other bounds in (B.20) follow in the same way. The extension of the bounds for the Lipschitz norm is standard; see for instance [31]. Thus, choosing properly $K = K(d, \mathfrak{p}_0)$ we obtain $|\Phi^{(y)}|_1 \leq (r + \rho r_0)^2$ so that (B.16) follows.

*(ii)* On the first $d$ components of the map $(\theta_+, y_+, w_+) = \Phi(\theta, y, w)$ we have $\theta_+ = \theta + f^{(\theta)}(\theta)$. If $\rho$ is small enough, we can apply Lemma A.171 in order to define a map $\widetilde{f}^{(\theta)}(\theta_+) \in W^{p,\infty}(\mathbb{T}^n_{s-\rho s_0})$ with $\|\widetilde{f}^{(\theta)}\|_{s-\rho s_0,p} \leq 2\|f^{(\theta)}\|_{s,p}$ and such that for $\theta_+ \in \mathbb{T}^n_{s-\rho s_0}$ one has $\widetilde{f}^{(\theta)}(\theta_+) + f^{(\theta)}(\theta_+ + \widetilde{f}^{(\theta)}(\theta_+)) = 0$. On the other hand, for $\theta \in \mathbb{T}_{s-2\rho s_0}$, one has $f^{(\theta)}(\theta) + \widetilde{f}^{(\theta)}(\theta + f^{(\theta)}(\theta)) = 0$, hence we set

$$\Psi^{(\theta)}(\theta_+) := \theta_+ + \widetilde{f}^{(\theta)}(\theta_+), \qquad \theta_+ \in \mathbb{T}^n_{s-\rho s_0}. \tag{B.22}$$

Define the map $\mathcal{T} : (\theta, y, w) \mapsto (\theta + f^{(\theta)}(\theta), y, w)$ and the vector field $G := \Pi_{\mathcal{A}} F - \Pi_{\mathcal{A}^{(\theta,0)}} F$ Denoting by $\hat{f}^{(v,k)}$ for $v = y, w$ and $k = y, w, 0$ the components of $G \circ \mathcal{T}$, by Lemma A.173 the vector field

$$\begin{aligned} G \circ \mathcal{T} &:= (\hat{f}^{(y,0)}(\theta_+) + \hat{f}^{(y,y)}(\theta_+)y + \hat{f}^{(y,w)}(\theta_+)w) \cdot \partial_y + \hat{f}^{(w,0)}(\theta_+)\partial_w \\ &= (f^{(y,0)} \circ \Psi^{(\theta)}(\theta_+) + f^{(y,y)} \circ \Psi^{(\theta)}(\theta_+)y + f^{(y,w)} \circ \Psi^{(\theta)}(\theta_+)w) \cdot \partial_y + f^{(w,0)} \circ \Psi^{(\theta)}(\theta_+)\partial_w \end{aligned} \tag{B.23}$$

is tame with scale of constants $C_p(G)$ as in (A.37) with $\beta \rightsquigarrow f^{(\theta)}$.

The third component of $\Phi$ is given by $w_+ := w + f^{(w,0)}(\theta)$ so that by (B.15) we deduce $w_+ \in B_{r+\rho r_0,a,\mathfrak{p}_0+\mu}$. On the latter component the map is a translation, hence we can write

$$w = w_+ + \widetilde{f}^{(w,0)}(\theta_+) := w_+ - f^{(w,0)}(\Psi^{(\theta)}(\theta_+)) =: \Psi^{(w)}(\theta_+, w_+), \tag{B.24}$$

so that by Lemma A.173 one has

$$\|\widetilde{f}^{(w,0)}\partial_w\|_{s-\rho s_0,a,p} \leq C_p(f) + C_{\mathfrak{p}_0}(f)|\widetilde{f}^{(\theta)}|_{p,s,\infty}, \tag{B.25}$$

hence by the smallness condition (B.15) we have that $w_+ \in B_{r-\rho r_0,a,\mathfrak{p}_0+\mu}$ implies $w \in B_{r,a,\mathfrak{p}_0+\mu}$.

The $y-$components are more delicate. Let us start by studying the invertibility of the finite-dimensional matrix $\mathbb{1} + f^{(y,y)}\left(\Psi^{(\theta)}(\theta_+)\right)$. Setting formally

$$\mathbb{1} + \widetilde{f}^{(y,y)}(\theta_+) := (\mathbb{1} + f^{(y,y)} \circ \Psi^{(\theta)})^{-1} = (\mathbb{1} + \hat{f}^{(y,y)})^{-1} = \sum_{n\geq 0}(-\hat{f}^{(y,y)}(\theta_+))^n, \tag{B.26}$$

we have

$$\begin{aligned} \|\widetilde{f}^{(y,y)}(\theta_+)y \cdot \partial_y\|_{s-\rho s_0,a,p} &\leq \sum_{n\geq 1} \|\hat{f}^{(y,y)}(\theta_+)^n y \cdot \partial_y\|_{s-\rho s_0,a,p} \\ &\overset{(F2)}{\leq} \sum_{n\geq 1} C^{n-1}_{G,\mathfrak{p}_0}\|\hat{f}^{(y,y)}(\theta_+)y \cdot \partial_y\|_{s-\rho s_0,a,p}, \end{aligned} \tag{B.27}$$

hence, by (A.37), (B.15) and (B.27) we obtain that $\widetilde{f}^{(y,y)}(\theta_+)y \cdot \partial_y$ is a well defined tame vector field with constant $C_p(\widetilde{f}^{(y,y)}) \leq 2C_p(G)$.

We now define the $y$–components of the map $\Psi$ as

$$y = \Psi^{(y)}(\theta_+, y_+, w_+) := y_+ + \widetilde{f}^{(y,0)}(\theta_+) + \widetilde{f}^{(y,y)}(\theta_+)y_+ + \widetilde{f}^{(y,w)}(\theta_+)w_+, \tag{B.28}$$

with $\widetilde{f}^{(y,y)}$ as above and

$$\widetilde{f}^{(y,w)}(\theta_+)w_+ := -(\mathbb{1} + \hat{f}^{(y,y)}(\theta_+))(f^{(y,w)} \circ \Psi^{(\theta)})(\theta_+)w_+, \tag{B.29a}$$

$$\widetilde{f}^{(y,0)}(\theta_+) := -(\mathbb{1} + \widetilde{f}^{(y,y)}(\theta_+))\Big((f^{(y,0)} \circ \Psi^{(\theta)})(\theta_+) + (f^{(y,w)} \circ \Psi^{(\theta)})(\theta_+)(f^{(w,0)} \circ \Psi^{(\theta)})(\theta_+)\Big). \tag{B.29b}$$

We will show that each summand can be associated with a tame vector fields with scale of constants such that $C_{\mathfrak{p}_0}(\widetilde{f}^{(y,k)})$ is small with $\rho$ for $k = 0, y, w$. This will implies the assertion, reasoning as in the proof of (B.16) i.e. via bounds of the form (B.20): in order to simplify the notation we drop the Lipschitz parameter $\lambda$. The bound on the term $\widetilde{f}^{(y,y)}(\theta_+)y_+ \cdot \partial_y$ follows from the discussion after (B.27). Let us consider (B.29a) and note that it is the sum of two terms, i.e.

$$(\text{B.29a}) = -(f^{(y,w)} \circ \Psi^{(\theta)})(\theta_+)w_+ - \hat{f}^{(y,y)}(\theta_+)(f^{(y,w)} \circ \Psi^{(\theta)})(\theta_+)w_+$$

The bound for $(f^{(y,w)} \circ \Psi^{(\theta)})(\theta_+)$ is a direct consequence of Lemma A.173. The second summand can be bounded by

$$\begin{aligned}\|\hat{f}^{(y,y)}(\theta_+)(f^{(y,w)} \circ \Psi^{(\theta)})(\theta_+)w_+\partial_y\|_{s-\rho s_0,a,p} &\overset{(F2)}{\le} C_p(G)\|(f^{(y,w)} \circ \Psi^{(\theta)})(\theta_+)w_+\|_{s-\rho s_0,a,\mathfrak{p}_0} \\ &\quad + C_{\mathfrak{p}_0}(G)\|(f^{(y,w)} \circ \Psi^{(\theta)})(\theta_+)w_+\|_{s-\rho s_0,a,p} \\ &\le C_p(G)C_{\mathfrak{p}_0}(G)\|w_+\|_{a,\mathfrak{p}_0} + C_{\mathfrak{p}_0}(G)^2\|w_+\|_{a,p}.\end{aligned} \tag{B.30}$$

The bound (B.29b) follows in the same way.

Collecting together the bounds (B.25), (B.19), (B.27), (B.30) we obtain that the map $\Psi$ with components defined in (B.22), (B.24) and (B.28) is such that $\Psi : D_{a,\mathfrak{p}_0+\mu}(r - \rho r_0, s - \rho s_0) \to D_{a,\mathfrak{p}_0+\mu}(r, s)$. Finally from the discussion above it follows that that the vector field

$$g := \widetilde{f}^{(\theta,0)}(\theta_+) \cdot \partial_\theta + (\widetilde{f}^{(y,0)}(\theta_+) + \widetilde{f}^{(y,y)}(\theta_+)y_+ + \widetilde{f}^{(y,w)}(\theta_+)w_+) \cdot \partial_y + \widetilde{f}^{(w,0)}(\theta_+)\partial_w$$

is tame with scale of constants $C_p(g) \le C_p(f) + C_{\mathfrak{p}_0}(f)\|f\|_{s,a,p}$ hence the result follows. □

**Lemma B.180.** *Let $g \in \mathcal{B}$ that satisfies the Hypothesis of Lemma B.179 with $\rho$ sufficiently small and consider $\Phi^t$ the flow at time $t \in [0,1]$ of the vector field $g$. Set $\Phi := \Phi^1$. Then, for all $p \ge \mathfrak{p}_0$ for all tame vector field $F : \mathcal{O} \times D_{a,p+\nu}(r + \rho r_0) \times \mathbb{T}^d_{s+\rho s_0} \to V_{a,p}$ one has*

$$\begin{aligned}\Phi_* F - F &= \int_0^1 \Phi^t_*[g, F]dt, \\ \Phi_* F - [g, F] - F &= \int_0^1 \int_0^t \Phi^s_*[g, [g, F]]dsdt.\end{aligned} \tag{B.31}$$

*Proof.* We start from the first of (B.31). Consider $\Phi^t$ the flow at time $t \in [0,1]$. It is know that it is a 1-parameter group of transformations. We want to evaluate its derivative in the parameter. Hence one has

$$\frac{d}{dt}\left(\Phi^t_* F\right)|_{t=t_0} = \frac{d}{ds}\left(\Phi^{t_0}_* \Phi^s_* F\right)|_{s=0} = \Phi^{t_0}_* \frac{d}{ds}\left(\Phi^s_* F\right)|_{s=0} = \Phi^{t_0}_*[g,F] \tag{B.32}$$

by the definition of Lie derivative. □

Now we have the following Lemma.

**Lemma B.181.** *Given any $F \in \mathcal{V}_{\vec{v},p}$ that satisfy properties $(F1)$ and $(F2)$ in Definition 3.2.26 with scale of constant $C_{\vec{v},p}(F)$, then the projection $\Pi_{\mathcal{R}^\perp} F$ is a tame vector field with the same scale of constant.*

*Proof.* The projection on the subspaces $\mathcal{N}$ and $\mathcal{A}$ can be defined in terms of the derivatives of $F$ (see formula (3.2.52)). Clearly the sum of tame vector fields is a tame vector field. Let study the single projections on the subspaces $\mathcal{N}$, $\mathcal{A}^{(v,0)}$, for $v = \theta, y, w$ and $\mathcal{A}^{(y,v)}$ with $v = y, w$ (see formulæ(3.2.53)). The tame properties for the fields in $\mathcal{A}^{(v,0)}$ are derived form $(F1)$ on the field $F$ since $\Pi_{\mathcal{A}^{(v,0)}} F = F^{(v)}(\theta, 0, 0)$. The vector fields

$$\Pi_{\mathcal{A}^{(y,v)}} F = \partial_v F^{(y)}(\theta,0,0)[v] \cdot \partial y, \quad v = y, w,$$

can be controlled by using property $(F1)$ on the field $F$. Same hold for the projection on $\mathcal{N}$. □

**Lemma B.182 (Commutator).** *Consider tame vector fields $F \in \mathcal{V}_{\vec{v},p}$ and $g \in \mathcal{B}$. Then the commutator $[F,g]$ satisfies properties $(F1)$ and $(F2)$ of Definition 3.2.26, for $p \le q - \nu - 3$, with constant*

$$C_{\vec{v},p}([F,g]) \le C_{\vec{v},p+1}(F) C_{\vec{v},\mathfrak{p}_0+\nu+1}(g) + C_{\vec{v},\mathfrak{p}_0+1}(F) C_{\vec{v},p+\nu+1}(g). \tag{B.33}$$

*Proof.* Let us check the properties $(F1)$ for the vector field $[g,f]$. Consider a map $\Phi := \mathbb{1} + h$ with $h$ that satisfies the hypothesis of Lemma B.179. We need to estimate the norm of

$$[g,f](\Phi) := dg(\Phi)[f(\Phi)] - df(\Phi)[g(\Phi)]. \tag{B.34}$$

In the following to simplify the notation we will drop the indices $a, s, \lambda$ in the norms since they are essentially fixed. First of all we have

$$\begin{aligned}\|dg(\Phi)[f(\Phi)]\|_p &\le (|g|_{p+1} + |g|_{\mathfrak{p}_0+1}\|\Phi\|_p)\|f(\Phi)\|_{\mathfrak{p}_0} + |g|_{\mathfrak{p}_0+1}\|f(\Phi)\|_p \\ &\le (|g|_{p+1} + |g|_{\mathfrak{p}_0+1}\|\Phi\|_p)(|f|_{\mathfrak{p}_0} + |f|_{\mathfrak{p}_0}\|\Phi\|_{\mathfrak{p}_0+\nu}) + |g|_{\mathfrak{p}_0+1}(|f|_p + |f|_{\mathfrak{p}_0}\|\Phi\|_{p+\nu}) \\ &\le 2(|g|_{p+1}|f|_{\mathfrak{p}_0} + |g|_{\mathfrak{p}_0+1}|f|_p) + 2|g|_{\mathfrak{p}_0+1}|f|_{\mathfrak{p}_0}\|\Phi\|_{p+\nu}.\end{aligned} \tag{B.35}$$

Moreover

$$\begin{aligned}df(\Phi)[g(\Phi)] &\le (|f|_{p+1} + |f|_{\mathfrak{p}_0+1}\|\Phi\|_{p+\nu})\|g(\Phi)\|_{\mathfrak{p}_0+\nu} + |f|_{\mathfrak{p}_0+1}\|g(\Phi)\|_{p+\nu} \\ &\le 2(|f|_{p+1}|g|_{\mathfrak{p}_0+\nu+1} + |g|_{\mathfrak{p}_0+\nu+1}|f|_{p+1}) + 2|g|_{\mathfrak{p}_0+\nu+1}|f|_{\mathfrak{p}_0+1}\|\Phi\|_{p+\nu}.\end{aligned} \tag{B.36}$$

The $(F2)$ and $(F3)$ follows in the same way. $\square$

Note that in Lemma B.182 one cannot in general conclude that the commutator between is a *tame* vector field in the sense of Definition 3.2.26 Indeed some problems appear in proving properties $(F3)$ and $(F4)$ on the commutator between two fields. It turn out one need a control on the fourth order derivative of the field $F$. This is **not** a problem if $F \in \mathcal{A}$ because in this $[F, g] \in \mathcal{R}^{\perp}$, hence one can use Lemma B.181 to conclude that $[F, g]$ is tame.

**Lemma B.183 (Remainder).** *Let $g \in \mathcal{B}$ that satisfies the Hypothesis of Lemma B.179 with $\rho$ sufficiently small and consider $\Phi^t$ the flow at time $t \in [0,1]$ of the vector field $g$. Then, for all $p \geq \mathfrak{p}_0$ for all tame vector field $F \in \mathcal{V}_{a,p}(s+\rho s_0, r+\rho r_0)$, one has that the vector field*

$$G := \Pi_{\mathcal{R}^{\perp}} \int_0^1 \int_0^t \Phi_*^s [g,[g,F]] ds dt \tag{B.37}$$

*is a tame vector field for $\mathfrak{p}_0 \leq p \leq q - \nu - 4$ and*

$$C_{\vec{v},p}(G) \lessdot C_{\vec{v},p+2}(F) C^2_{\vec{v},\mathfrak{p}_0+\nu+1}(g) + C_{\vec{v},\mathfrak{p}_0+2}(R) C_{\vec{v},\mathfrak{p}_0+\nu+1}(g) C_{\vec{v},p+\nu+1}(g) \tag{B.38}$$

*Proof.* We consider the map $\Phi^t := \mathbb{1} + f_t$ the flow at time $t$ of the vector field $g$. Moreover by Lemma B.178 one has that the norm of $f_t$ is controlled by the norm of $g$. The first step is to prove that the commutator $[g,[g,F]]$ satisfies properties $(F1)$ and $F(2)$. Consider a map $\Gamma$ as in Definition 3.2.26, one has that

$$\begin{aligned} &[g,[g,F]](\Gamma) := dg(\Gamma)[dg(\Gamma)[R(\Gamma)]] - 2dg(\Gamma)[dR(\Gamma)[g(\Gamma)]] - d^2g(\Gamma)[R(\Gamma), g(\Gamma)] \\ &- d^2R(\Gamma)[g(\Gamma), g(\Gamma)] + dR(\Gamma)[dg(\Gamma)[g(\Gamma)]]. \end{aligned} \tag{B.39}$$

Hence by using properties $(F1), (F2), (F3)$ on $R$ and $g$ we have that the double commutator satisfies $(F1)$ with

$$C_{\vec{v},p}([g,[g,R]]) \lessdot C_{\vec{v},p+2}(F) C^2_{\vec{v},\mathfrak{p}_0+\nu+1}(g) + C_{\vec{v},\mathfrak{p}_0+2}(R) C_{\vec{v},\mathfrak{p}_0+\nu+1}(g) C_{\vec{v},p+\nu+1}(g). \tag{B.40}$$

The property $(F2)$ for $[g,[g,F]]$ can be checked in the same way by using in addiction property $F(4)$ on the field $R$. This is the most important point where we need to work with $C^3$ tame vector field. Now by definition we have that

$$\Phi_*^s [g,[g,F]] := [g,[g,F]]((\Phi^t)^{-1}) + df((\Phi^t)^{-1})[[g,[g,F]]((\Phi^t)^{-1})], \tag{B.41}$$

hence properties $(F1)$ and $(F2)$ are still satisfied with the constant

$$C_{\vec{v},p}(\Phi_*^s S) \leq (1 + 5|f_t|_{\vec{v},\mathfrak{p}_0+\nu+1}) C_{\vec{v},p}(s) + 5C_{\vec{v},\mathfrak{p}_0}(s)(1+5|f_t|_{\vec{v},\mathfrak{p}_0+\nu+1})|f_t|_{\vec{v},p+\nu+1}, \tag{B.42}$$

where $S := [g,[g,F]]$ (see the proof of Lemma 3.2.31 for further details). To finish the proof of the Lemma it is sufficient to use the smallness condition on the $|g|_{\vec{v},\mathfrak{p}_1}$ and use Corollary B.181 in order to conclude that the projection on $\mathcal{R}^{\perp}$ is "tame" (i.e. satisfies also properties $(F3)$ and $(F4)$) with constant given in (B.38). $\square$

In Lemma remainder we use strongly the $(F4)$ on a non linear vector field $F$. Essentially we need to perform three derivatives just to be able to control the remainder of the second order in $g$ of the push-forward of a vector field.

The last Lemma we need is quite technical. In our application we will always deal with vector field with a particular form, i.e. a constant coefficients vector field $N_0$ plus something in $\mathcal{V}_{a,p}(s,r)$ that is "perturbative" with respect to the size of $N_0$. We will explain later what we mean with the term "perturbative".

**Lemma B.184 (Normal form).** *Let $g \in \mathcal{B}$ that satisfies the Hypothesis of Lemma B.179 with $\rho$ sufficiently small and assume also that $g \in E^{(K)}$ for some $K > 0$ large (see Definition 3.2.32 for the definition of the $E^{(K)}$). Let $N_0 := N_0(\xi) \in \mathcal{N}^{\mathcal{O}}$ be a vector field of the form*

$$N_0 := \omega_0 \cdot \partial_\theta + \Omega_0 w \cdot \partial_w \tag{B.43}$$

*with constant coefficients. Then one has*

$$\Pi_{\mathcal{A}^{\perp}}[g, N_0] = 0, \tag{B.44}$$

*and*

$$\Pi_K^{\perp}\Pi_{\mathcal{A}}[g, N_0] = 0, \tag{B.45}$$

*where $\Pi_K^{\perp} := \mathbb{1} - \Pi_K$.*

*Proof.* First of all note that since $N_0$ has constant coefficients it preserves the subspaces $E^K$. Hence formula (B.45) follows since $g \in E^{(K)}$ so that the projection $\Pi_K^{\perp}$ is zero. Let us check the (B.44). Since $g \in \mathcal{B}^{\mathcal{O}}_{a,p+3}(s)$ and $N_0 \in \mathcal{N}$ then $\Pi_{\mathcal{R}}[g, N_0] = 0$. Then by explicit calculation we have that

$$[g, N_0]^{(\theta)} := (\omega \cdot \partial_\theta g^{(\theta,0)}(\theta)) \cdot \partial_\theta, \quad [g, N_0]^{(w)} := (\omega \cdot \partial_\theta g^{(w,0)}(\theta) - \Omega_0 g^{(w,0)}) \cdot \partial_w. \tag{B.46}$$

Clearly the $w-$component does not have a linear term in $w$, moreover the $\theta-$component has zero average in $\theta$, hence also the projection on the subspace $\mathcal{N}$ is zero. $\square$

**Lemma B.185.** *Consider a map $\Phi = \mathbb{1} + f$ and a tame field $F \in \mathcal{N}$ as in Lemma 3.2.31. Assume in addiction that $f \in \mathcal{B}$. Then, for all $p \geq \mathfrak{p}_0$ one has that*

$$\Pi_{\mathcal{N}^{\perp}} C_{\vec{v}_1,p}(\Phi_* F) \leq C_{\vec{v},p}(F) C_{\vec{v}_2,\mathfrak{p}_1+1}(f) + C_{\vec{v},\mathfrak{p}_1}(F) C_{\vec{v}_2,p+1}(f)(1 + 2C_{\vec{v}_2,\mathfrak{p}_1+1}(f)) \tag{B.47}$$

*where* $\vec{v} := (\lambda, \mathcal{O}, s, a)$, $\vec{v}_1 := (\lambda, \mathcal{O}, s - 2\rho s_0, a - 2\delta a_0)$ *and* $\vec{v}_2 := (\lambda, \mathcal{O}, s - \rho s_0, a - \delta a_0)$.

*Proof.* We have that the field $F$ has the form $F = F^{(\theta,0)}(\theta) \cdot \partial_\theta + F^{(w,w)}(\theta) w \cdot \partial_w$. If we consider the map

$$\theta_+ = \theta, \quad y_+ = y + f^{(y,0)}(\theta) + f^{(y,y)}(\theta) y + f^{(y,w)}(\theta) w, \qquad w_+ = w + f^{(w,0)}(\theta)$$

one can write explicitly the conjugated field. Hence one has

$$\Pi_{\mathcal{N}^\perp}\Big(\Phi_* F\Big) = \Pi_{\mathcal{N}^\perp}\Big(F \circ \Phi^{-1} + d_u f(\Phi^{-1})[F \circ \Phi^{-1}]\Big) = -F^{(w,w,)}(\theta) f^{(w,0)}(\theta) + \Pi_{\mathcal{N}^\perp}\Big(d_u f(\Phi^{-1})[F \circ \Phi^{-1}]\Big). \tag{B.48}$$

All terms in (B.48) can be estimated using the tameness of the fields, i.e. the properties $(Fi)$ i=1,2, in Definition 3.2.26. □

# C. Proofs

## C.1 Proof of Lemmata 4.5.84 and 4.5.83

*Proof of Lemma 4.5.84.* Define the function $\psi : \Lambda \to \mathbb{C}$,

$$\begin{aligned}\psi(\lambda) &:= i\lambda\bar{\omega}\cdot\ell + \Omega_{\sigma,j}(\lambda) - \Omega_{\sigma',j'}(\lambda) \\ &\overset{(4.3.89)}{=} i\lambda\bar{\omega}\cdot\ell - im(\lambda)(\sigma j^2 - \sigma' j'^2) + r^{\infty}_{\sigma,j}(\lambda) - r^{\infty}_{\sigma',j'}(\lambda),\end{aligned} \tag{C.1}$$

where with abuse of notation we set $r^{\infty}_{\sigma,0} \equiv 0$. Note that, by $(N1)_n$ of Theorem 3.1.18, we have $||\mathbf{u}_n||_{\mathfrak{s}_0+\mu_2,\gamma} \leq 1$ on $G_n$. Then (4.3.90) holds and we have

$$\begin{aligned}|\Omega_{\sigma,j} - \Omega_{\sigma',j'}|^{\text{lip}} &\leq |m|^{\text{lip}}|\sigma j^2 - \sigma' j'^2| + |r^{\infty}_{\sigma,j}|^{\text{lip}} + |r^{\infty}_{\sigma',j'}|^{\text{lip}} \\ &\leq C\varepsilon\gamma^{-1}|\sigma j^2 - \sigma' j'^2| \overset{(4.5.232)}{\leq} C\varepsilon\gamma^{-1}|\bar{\omega}\cdot\ell|.\end{aligned} \tag{C.2}$$

We can estimate, for any $\lambda_1, \lambda_2 \in \Lambda$,

$$\frac{|\psi(\lambda_1) - \psi(\lambda_2)|}{|\lambda_1 - \lambda_2|} \overset{(4.5.232),(C.2)}{\geq} \left(\frac{1}{8} - C\varepsilon\gamma^{-1}\right)|\bar{\omega}\cdot\ell| \geq \frac{|\sigma j^2 - \sigma' j'^2|}{9}, \tag{C.3}$$

if $\varepsilon\gamma^{-1}$ is small enough. Then, using standard measure estimates on sub-levels of Lipschitz functions, we conclude

$$|R^{\sigma,\sigma'}_{\ell jj'}| \leq 4\gamma_n|\sigma j^2 - \sigma' j'^2|\langle\ell\rangle^{-\tau}\frac{9}{|\sigma j^2 - \sigma' j'^2|} \leq C\gamma\langle\ell\rangle^{-\tau}. \tag{C.4}$$

□

*Proof of Lemma 4.5.83.* We first prove the (4.5.232); note that if $(\sigma, j) = (\sigma', j')$ then it is trivially true. If $R^{\sigma,\sigma'}_{\ell jj'}(\mathbf{u}_n) \neq \emptyset$, then, by definition (4.5.230), there exists a $\lambda \in \Lambda$ such that

$$|\Omega_{\sigma,j}(\mathbf{u}_n) - \Omega_{\sigma',j'}(\mathbf{u}_n)| < 2\gamma_n|\sigma j^2 - \sigma' j'^2|\langle\ell\rangle^{-\tau} + 2|\bar{\omega}\cdot\ell|. \tag{C.5}$$

On the other hand, for $\varepsilon$ small and since $(\sigma, j) \neq (\sigma', j')$,

$$|\Omega_{\sigma,j}(\mathbf{u}_n) - \Omega_{\sigma',j'}(\mathbf{u}_n)| \overset{(4.2.35a),(4.3.90)}{\geq} \frac{1}{2}|\sigma j^2 - \sigma' j'^2| - C\varepsilon \geq \frac{1}{3}|\sigma j^2 - \sigma' j'^2|. \tag{C.6}$$

By the (C.5), (C.6) and $\gamma_n \leq 2\gamma$ follows

$$2|\bar{\omega}\cdot\ell| \geq \left(\frac{1}{3} - \frac{4\gamma}{\langle\ell\rangle^{\tau}}\right)|\sigma j^2 - \sigma' j'^2| \geq \frac{1}{4}|\sigma j^2 - \sigma' j'^2|, \tag{C.7}$$

since $\gamma \leq \gamma_0$, by choosing $\gamma_0$ small enough. It is sufficient $\gamma_0 < 1/48$. Then, the (4.5.232) hold.
In order to prove the (4.5.231) we need to understand the variation of the eigenvalues $\Omega_{\sigma,j}(\mathbf{u})$ with respect to the function $\mathbf{u}$. If we assume that

$$|(\Omega_{\sigma,j} - \Omega_{\sigma',j'})(\mathbf{u}_n) - (\Omega_{\sigma,j} - \Omega_{\sigma',j'})(\mathbf{u}_{n-1})| \leq C\varepsilon|\sigma j^2 - \sigma' j'^2| N_n^{-\alpha}, \tag{C.8}$$

then, for $j \neq j'$, $|\ell| \leq N_n$, and $\lambda \in G_n$, we have

$$\begin{aligned} |i\lambda\bar{\omega}\cdot\ell + \Omega_{\sigma,j}(\mathbf{u}_n) \quad &- \quad \Omega_{\sigma',j'}(\mathbf{u}_n)| \overset{(C.8)}{\geq} 2\gamma_{n-1}|\sigma j^2 - \sigma' j'^2|\langle\ell\rangle^{-\tau} \\ &- \quad C\varepsilon|\sigma j^2 - \sigma' j'^2| N_n^{-\alpha} \geq 2\gamma_n|\sigma j^2 - \sigma' j'^2|\langle\ell\rangle^{-\tau}, \end{aligned} \tag{C.9}$$

because $C\varepsilon\gamma^{-1}N_n^{\tau-\alpha}2^{n+1} \leq 1$ if $\varepsilon\gamma^{-1}$ small enough. We complete the proof by verifying (C.8).

By Lemma 4.3.71, using the $(\mathbf{S4})_{n+1}$ with $\gamma = \gamma_{n-1}$ and $\gamma - \rho = \gamma_n$, and with $\mathbf{u}_1 = \mathbf{u}_{n-1}$, $\mathbf{u}_2 = \mathbf{u}_n$, we have

$$\Lambda_{n+1}^{\gamma_{n-1}}(\mathbf{u}_{n-1}) \subseteq \Lambda_{n+1}^{\gamma_n}(\mathbf{u}_n), \tag{C.10}$$

since, for $\varepsilon\gamma^{-1}$ small enough,

$$\varepsilon C N_n^{\tau} \sup_{\lambda\in G_n} ||\mathbf{u}_n - \mathbf{u}_{n-1}||_{\mathfrak{s}_0+\mu} \overset{(3.1.7)}{\leq} \varepsilon^2\gamma^{-1}CC_{\star}N_n^{\tau-\kappa_3} \leq \gamma_{n-1} - \gamma_n =: \rho = \gamma 2^{-n}.$$

where $\kappa_3$ is defined in (3.1.7) with $\nu = 2$, $\mu$ defined in (4.4.206) with $\eta = \eta_1 + \beta$, $\mu > \tau$ (see Lemmata 4.5.79, 5.4.113 and (4.3.117), (4.2.31)). We also note that,

$$\begin{aligned} G_n \quad &\overset{\text{Def.5.4.211,(4.1.23)}}{\subseteq} \quad \Lambda_{\infty}^{2\gamma_{n-1}}(\mathbf{u}_{n-1}) \overset{(4.3.187)}{\subseteq} \cap_{\nu\geq 0}\Lambda_{\nu}^{\gamma_{n-1}}(\mathbf{u}_{n-1}) \\ &\subseteq \quad \Lambda_{n+1}^{\gamma_{n-1}}(\mathbf{u}_{n-1}) \overset{(C.10)}{\subseteq} \Lambda_{n+1}^{\gamma_n}(\mathbf{u}_n). \end{aligned} \tag{C.11}$$

This means that $\lambda \in G_n \subset \Lambda_{n+1}^{\gamma_{n-1}}(\mathbf{u}_{n-1}) \cap \Lambda_{n+1}^{\gamma_n}(\mathbf{u}_n)$, and hence, we can apply the $(\mathbf{S3})_{\nu}$, with $\nu = n+1$, in Lemma 4.3.71 to get

$$\begin{aligned} |r_{\sigma,j}^{\infty}(\mathbf{u}_n) - r_{\sigma,j}^{\infty}(\mathbf{u}_{n-1})| &\leq |r_{\sigma,j}^{n+1}(\mathbf{u}_n) - r_{\sigma,j}^{n+1}(\mathbf{u}_{n-1})| \\ &+ |r_{\sigma,j}^{\infty}(\mathbf{u}_n) - r_{\sigma,j}^{n+1}(\mathbf{u}_n)| + |r_{\sigma,j}^{\infty}(\mathbf{u}_{n-1}) - r_{\sigma,j}^{n+1}(\mathbf{u}_{n-1})| \\ &\overset{(4.3.185),(4.2.36a),(4.3.129b)}{\leq} \varepsilon C||\mathbf{u}_n - \mathbf{u}_{n-1}||_{\mathfrak{s}_0+\eta_2} \\ &+ \varepsilon\left(1 + ||\mathbf{u}_{n-1}||_{\mathfrak{s}_0+\eta_1+\beta} + ||\mathbf{u}_n||_{\mathfrak{s}_0+\eta_1+\beta}\right) N_n^{-\alpha} \\ &\overset{(3.1.7)}{\leq} C\varepsilon^2\gamma^{-1}N_n^{-\mu_3} + \varepsilon\left(1 + ||\mathbf{u}_{n-1}||_{\mathfrak{s}_0+\eta_1+\beta} + ||\mathbf{u}_n||_{\mathfrak{s}_0+\eta_1+\beta}\right) N_n^{-\alpha}. \end{aligned}$$

Now, first of all $\mu_3 > \alpha$ by (3.1.7), (4.3.117), moreover $\eta_1 + \beta < \eta_5$ then by $(\mathbf{S1})_n$, $(\mathbf{S1})_{n-1}$, one has $||\mathbf{u}_{n-1}||_{\mathfrak{s}_0+\eta_5} + ||\mathbf{u}_n||_{\mathfrak{s}_0+\eta_5} \leq 2$, we obtain

$$|r^{\infty}_{\sigma,j}(\mathbf{u}_n) - r^{\infty}_{\sigma,j}(\mathbf{u}_{n-1})| \leq \varepsilon N_n^{-\alpha}. \tag{C.12}$$

Then, by (4.2.35b) and (C.12), one has that the (C.8) hold and the proof of Lemma (4.5.83) is complete.

□

## C.2 Proof of Lemma 2.3.15

We first show that $T$ is symplectic. Consider $W = (w^{(1)}, w^{(2)}), V = (v^{(1)}, v^{(2)}) \in H^s(\mathbb{T}^{d+1};\mathbb{R}) \times H^s(\mathbb{T}^{d+1};\mathbb{R})$ and set $w = w^{(1)} + iw^{(2)}$, $v = v^{(1)} + iv^{(2)}$, then one has

$$\widetilde{\Omega}(TW, TV) := \int_{\mathbb{T}} \begin{pmatrix} \frac{i}{\sqrt{2}} w \\ \frac{1}{\sqrt{2}} \bar{w} \end{pmatrix} \cdot J \begin{pmatrix} \frac{i}{\sqrt{2}} v \\ \frac{1}{\sqrt{2}} \bar{v} \end{pmatrix} dx = \int_{\mathbb{T}} WJV dx =: \widetilde{\Omega}(W, V). \tag{C.13}$$

To show the (2.3.35) is sufficient to apply the definition of $T_1$. First of all consider the linearized operator in some $z = (z^{(1)}, z^{(2)})$

$$D_z\mathcal{F}(\omega t, x, z) = D_\omega + \varepsilon D_z g(\omega t, x, z) = D_\omega + \varepsilon \partial_{z_0} g + \varepsilon \partial_{z_1} g \partial_x + \varepsilon \partial_{z_2} g \partial_{xx} \tag{C.14}$$

where $D_\omega$ and $g$ are defined in (2.1.11) and (2.1.12) and

$$\partial_{z_i} g := (a^{(i)}_{jk})_{j,k=1,2} := (\partial_{z_i^{(j)}} g_k)_{j,k=1,2}. \tag{C.15}$$

All the coefficients $a^{(i)}_{jk}$ are evaluated in $(z^{(1)}, z^{(2)}, z^{(1)}_x, z^{(2)}_x, z^{(1)}_{xx}, z^{(2)}_{xx})$. By using the definitions (C.14), (C.15) and recalling that $g = (g_1, g_2) = (-f_1, f_2)$ and $\mathtt{f} = f_1 + if_2$, one can check with an explicit computation that

$$\mathcal{L}(z) = T_1^{-1} T d_z \mathcal{F}(\omega t, x, z) T^{-1} T_1$$

has the desired form. ■

## C.3 Proof of Lemma 3.2.43

We write $F = N_0 + G$ and define

$$\mathfrak{L} := \mathfrak{D} + \mathfrak{R} := \Pi_K \Pi_{\mathcal{A}}([N, \cdot]) + \Pi_K \Pi_{\mathcal{A}}([R, \cdot])$$

with $R := \Pi_{\mathcal{R}} G$ and $N = \Pi_{\mathcal{N}} F$. Let us first notice that by construction the *Melnikov conditions* (see Definition 3.2.42) allow us to invert approximately the operator $\mathfrak{D}$ and that the approximate solution of the homological equation is in $\hat{\mathcal{B}}$. Let us show that that the invertibility of $\mathfrak{D}$ implies the invertibility of $\mathfrak{L}$, by showing that $\mathfrak{R} := \Pi_K \Pi_{\mathcal{A}}([R, \cdot])$ is upper triangular with zero on the diagonal and hence nilpotent.

**Lemma C.186 (commutator).** *Given $\Psi \in \mathcal{A}$ and a vector field $R \in \mathcal{R}$. Let us write $\Psi = (\Psi^{(y,0)} + \Psi^{(y,y)} y + \Psi^{(y,w)} w) \cdot \partial_y + \Psi^{(w,0)} \partial_w$, then one has*

$$\begin{aligned}
&\Pi_{\mathcal{A}}[R, \Psi^{(y,y)} y \partial_y] = 0,\\
&\Pi_{\mathcal{A}}[R, \Psi^{(y,w)} w \partial_y] = -\Psi^{(y,w)}(\theta) R^{(w,y)}(\theta) y \cdot \partial_y\\
&\Pi_{\mathcal{A}}[R, \Psi^{(w,0)} \partial_w] = (\partial_{wy} R^{(y)}(\theta,0,0) \Psi^{(w,0)} y + \partial_{ww} R^{(y)}(\theta,0,0) \Psi^{(w,0)} w) \cdot \partial_y\,,\\
&\Pi_{\mathcal{A}}[R, \Psi^{(y,0)} \partial_y] = -(\partial_\theta \Psi^{(y,0)}(\theta) R^{(\theta,y)}(\theta) y + \partial_\theta \Psi^{(y,0)}(\theta) R^{(\theta,w)} w\\
&\qquad - \partial_{yy} R^{(y)}(\theta,0,0) \Psi^{(y,0)} y - \partial_{yw} R^{(y)}(\theta,0,0) \Psi^{(y,0)}(\theta) w) \cdot \partial_y + R^{(w,y)}(\theta) \Psi^{(y,0)}(\theta) \cdot \partial_w.
\end{aligned} \tag{C.16}$$

*Moreover*

$$\Pi_{\mathcal{N}}[R, \Psi] = [R, \Psi]^{(\theta)}(\theta,0,0) \cdot \partial_\theta + G(\theta) w \partial_w \tag{C.17}$$

*with*

$$G(\theta) = R^{(w,y)}(\theta) \Psi^{(y,w)} + \partial_{yw} R^{(w)}(\theta,0,0) \Psi^{(y,0)} + \partial_{ww} R^{(w)}(\theta,0,0) \Psi^{(w,0)} - \partial_\theta \Psi^{(w,0)} R^{(\theta,w)}(\theta)$$

*Proof.* It is an explicit computation. ∎

A vector field in $\mathcal{A}$ is determined by its five components $\Psi_i$. Hence the operator $\mathfrak{R} : \Psi \mapsto \Pi_{\mathcal{A}}[R, \Psi]$ can be written in the form

$$\begin{pmatrix} (\Pi_{\mathcal{A}}[R,\Psi])^{(y,y)} \\ (\Pi_{\mathcal{A}}[R,\Psi])^{(y,w)} \\ (\Pi_{\mathcal{A}}[R,\Psi])^{(w,0)} \\ (\Pi_{\mathcal{A}}[R,\Psi])^{(y,0)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & -R^{(w,y)} & \partial_{wy} R^{(y)} & \partial_{yy} R^{(y)} - R^{(\theta,y)} \partial_\theta \\ 0 & 0 & 0 & \partial_{ww} R^{(y)} & R^{(w,y)} \\ 0 & 0 & 0 & 0 & R^{(w,y)} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \Psi^{(y,y)} \\ \Psi^{(y,w)} \\ \Psi^{(w,0)} \\ \Psi^{(y,0)} \end{pmatrix} \tag{C.18}$$

Recall that $\mathfrak{D}$ is block-diagonal w.r.t. the above decomposition; by definition $\mathcal{W}$ is block-diagonal as well and provides an approximate inverse for $\mathfrak{D}$. Let us set $\mathfrak{U} := \mathfrak{R}\mathcal{W}$, we note that $\mathfrak{U}$ is upper triangular so $\mathfrak{U}^4 = 0$ Now set $\mathfrak{B} = \mathfrak{D}\mathcal{W} - \mathbb{1}$, by definition we have

$$(\mathfrak{D} + \mathfrak{R})\mathcal{W}(1 + \mathfrak{U})^{-1} = (\mathbb{1} + \mathfrak{U} + \mathfrak{B})(1 + \mathfrak{U})^{-1} = \mathbb{1} + \mathfrak{B}(1 + \mathfrak{U})^{-1}\,, \quad (1 + \mathfrak{U})^{-1} = \sum_{j=0}^{4} (-1)^j \mathfrak{U}^j \tag{C.19}$$

Now set $X = \Pi_K \Pi_{\mathcal{A}} F$ we claim that

$$g = \mathcal{W} \sum_{j=0}^{4} (-1)^j \mathfrak{U}^j X, \tag{C.20}$$

satisfies the bounds (3.2.59), (3.2.60). By hypothesis $X \in \mathcal{E}$, by the properties of $\mathcal{E}$ and definition of $\mathcal{W}$ we have that $\mathcal{W}\mathcal{U}^j X \in \hat{\mathcal{B}}$ so that $g \in \hat{\mathcal{B}}$. We estimate the five terms of the sum separately. Obviously for $j = 0$ the estimate (3.2.79) holds. Moreover for $k = 1$ one has

$$\begin{aligned}
|\mathcal{W}\mathfrak{R}(\mathcal{W}X)|_{\vec{v}_1,p} &\overset{(3.2.79)}{\leq} \gamma^{-1}K^{\eta}\Big[|\mathfrak{R}\mathcal{W}X|_{\vec{v},p} + K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,p}(G)|\mathfrak{R}\mathcal{W}X|_{\vec{v},\mathfrak{p}_1}\Big] \\
&\overset{(B.33)}{\leq} \gamma^{-1}K^{\eta+\nu+1}\Big[C_{\vec{v}_1,p}(R)C_{\vec{v}_1,\mathfrak{p}_1}(\mathcal{W}X) + C_{\vec{v}_1,\mathfrak{p}_1}(R)C_{\vec{v}_1,p}(\mathcal{W}X) \\
&\qquad + 2K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,p}(G)C_{\vec{v}_1,\mathfrak{p}_1}(R)C_{\vec{v}_1,\mathfrak{p}_1}(\mathcal{W}X)\Big] \\
&\overset{(3.2.79)}{\leq} \gamma^{-1}K^{\eta+\nu+1}K^{\eta}\Big[\gamma^{-1}C_{\vec{v}_1,p}(R)(1+\gamma^{-1}C_{\vec{v}_1,\mathfrak{p}_1}(G))|X|_{\vec{v}_1,\mathfrak{p}_1} \\
&\qquad + \gamma^{-1}C_{\vec{v}_1,p}(R)(|X|_{\vec{v}_1,p} + K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,p}(G)|X|_{\vec{v}_1,\mathfrak{p}_1}) \\
&\qquad + \gamma^{-1}C_{\vec{v}_1,p}(G)K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,\mathfrak{p}_1}(R)(1+\gamma^{-1}C_{\vec{v}_1,\mathfrak{p}_1}(G))|X|_{\vec{v}_1,\mathfrak{p}_1}\Big] \\
&\leq \gamma^{-1}K^{2\eta+\nu+4}\Big[|X|_{\vec{v}_1,p} + K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,p}(G)|X|_{\vec{v}_1,\mathfrak{p}_1}\Big]
\end{aligned} \tag{C.21}$$

where in the last inequality we have used the fact that $\gamma^{-1}C_{\vec{v}_1,\mathfrak{p}_1}(R)$ is controlled by $\gamma^{-1}C_{\vec{v}_1,\mathfrak{p}_1}(G) \leq K$. The term with $k = 2$ can be estimated by

$$\begin{aligned}
|\mathcal{W}\mathfrak{R}\mathcal{W}\mathfrak{R}(\mathcal{W}X)|_{\vec{v}_1,p} &\leq \gamma^{-1}K^{\eta}\Big[|\mathfrak{R}\mathcal{W}\mathfrak{R}(\mathcal{W}X)|_{\vec{v},p} + K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,p}(G)|\mathfrak{R}\mathcal{W}\mathfrak{R}(\mathcal{W}X)|_{\vec{v},\mathfrak{p}_1}\Big] \\
&\overset{(B.33)}{\leq} \gamma^{-1}K^{\eta+\nu+1}\Big[C_{\vec{v}_1,p}(R)C_{\vec{v}_1,\mathfrak{p}_1}(\mathcal{W}\mathfrak{R}(\mathcal{W}X)) + C_{\vec{v}_1,\mathfrak{p}_1}(R)C_{\vec{v}_1,p}(\mathcal{W}\mathfrak{R}(\mathcal{W}X)) \\
&\qquad + 2K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,p}(G)C_{\vec{v}_1,\mathfrak{p}_1}(R)C_{\vec{v}_1,\mathfrak{p}_1}(\mathcal{W}\mathfrak{R}(\mathcal{W}X))\Big] \\
&\overset{(C.21)}{\leq} \gamma^{-1}K^{\eta+\nu+1}\Big[\gamma^{-1}C_{\vec{v}_1,p}(R)K^{2\eta+\nu+4}(1+\gamma^{-1}C_{\vec{v}_1,\mathfrak{p}_1}(G))|X|_{\vec{v}_1,\mathfrak{p}_1} \\
&\qquad + 2K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,p}(G)C_{\vec{v}_1,\mathfrak{p}_1}(R)\gamma^{-1}K^{2\eta+\nu+4}(1+\gamma^{-1}C_{\vec{v}_1,\mathfrak{p}_1}(G))|X|_{\vec{v}_1,\mathfrak{p}_1} \\
&\qquad + \gamma^{-1}C_{\vec{v}_1,\mathfrak{p}_1}(R)K^{2\eta+\nu+4}(|X|_{\vec{v}_1,p} + K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,p}(G)|X|_{\vec{v}_1,\mathfrak{p}_1})\Big] \\
&\leq \gamma^{-1}K^{3\eta+2\nu+8}\Big[|X|_{\vec{v}_1,p} + K^{\alpha(p-\mathfrak{p}_1)}\gamma^{-1}C_{\vec{v}_1,p}(G)|X|_{\vec{v}_1,\mathfrak{p}_1}\Big].
\end{aligned} \tag{C.22}$$

By following the same reasoning for $k = 3$ and $k = 4$ we have that the whole sum satisfies (3.2.79) with $\mu := 5\eta + 4\nu + 20$. Let us check (3.2.60). By following the same reasoning used to estimate $g$ in (C.20) we have that

$$|(\mathbb{1} + \mathfrak{U})^{-1}X|_{\vec{v}_1,p} \leq K^{4\eta+4\nu+20}\Big(|X|_{\vec{v}_1,p} + \gamma^{-1}C_{\vec{v}_1,p}(G)K^{\alpha(p-\mathfrak{p}_1)}|X|_{\vec{v}_1,\mathfrak{p}_1}\Big) \tag{C.23}$$

By equation (C.19) one has $[(\mathfrak{D}+\mathfrak{R})g - X] = \mathfrak{B}(\mathbb{1}+\mathfrak{U})^{-1}X$. Moreover one has

$$\begin{aligned}
|\mathfrak{B}(\mathbb{1}+\mathfrak{U})^{-1}X|_{\vec{v}_1,p} &\overset{(3.2.80)}{\leq} \gamma^{-1}K^{\eta}\mathfrak{r}(|(\mathbb{1}+\mathfrak{U})^{-1}X|_{\vec{v}_1,p}C_{\vec{v}_1,\mathfrak{p}_1}(G) + C_{\vec{v}_1,p}(G)K^{\alpha(p-\mathfrak{p}_1)}|(\mathbb{1}+\mathfrak{U})^{-1}X|_{\vec{v}_1,\mathfrak{p}_1}) \\
&\overset{(C.23)}{\leq} \gamma^{-1}\mathfrak{r}K^{5\eta+4\nu+20}\Big((|X|_{\vec{v}_1,p} + \gamma^{-1}C_{\vec{v}_1,p}(G)K^{\alpha(p-\mathfrak{p}_1)}|X|_{\vec{v}_1,\mathfrak{p}_1})C_{\vec{v}_1,\mathfrak{p}_1}(G) \\
&\quad + C_{\vec{v}_1,p}(G)K^{\alpha(p-\mathfrak{p}_1)}(1+\gamma^{-1}C_{\vec{v}_1,\mathfrak{p}_1})|X|_{\vec{v}_1,\mathfrak{p}_1}\Big) \\
&\leq \gamma^{-1}\mathfrak{r}K^{5\eta+4\nu+21}\Big(|X|_{\vec{v}_1,p}C_{\vec{v}_1,\mathfrak{p}_1}(G) + C_{\vec{v}_1,p}(G)K^{\alpha(p-\mathfrak{p}_1)}|X|_{\vec{v}_1,\mathfrak{p}_1}\Big),
\end{aligned} \tag{C.24}$$

which implies (3.2.60).

# Bibliography


[1] V. I. Arnold, *Small denominators and problems of stability of motion in classical celestial mechanics*, Russian Math. Surveys **18** (1963), 85–193.

[2] P. Baldi, *Periodic solutions of forced Kirchhoff equations*, Ann. Scuola Norm. Sup. Pisa, Cl. Sci. **8** (2009), no. 5, 117–141.

[3] ______, *Periodic solutions of fully nonlinear autonomous equations of Benjamin-Ono type*, Ann. I. H. Poincaré (C) Anal. Non Linéaire **30** (2013), 33–77.

[4] P. Baldi, M. Berti, and R. Montalto, *KAM for autonomous quasilinear perturbations of KdV*, arXiv:1404-3125, 2014.

[5] ______, *KAM for quasi-linear and fully nonlinear forced KdV*, Math. Ann. **359** (2014), 471–536.

[6] D. Bambusi and B. Grébert, *Birkhoff normal form for partial differential equations with tame modulus*, Duke Math. J. **135 n. 3** (2006), 507–567.

[7] M. Berti, L. Biasco, and M. Procesi, *KAM theory for the Hamiltonian derivative wave equation*, Annales Scientifiques de l'ENS **46** (2013), no. 2, 299–371.

[8] M. Berti, P. Biasco, and M. Procesi, *KAM theory for reversible derivative wave equations*, Archive for Rational Mechanics and Analysis **212** (2014), no. 3, 905–955.

[9] M. Berti and P. Bolle, *A Nash-Moser approach to KAM theory*, to appear on Field Institute Communications.

[10] ______, *Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential*, Nonlinearity **25** (2012), 2579–2613.

[11] M. Berti, P. Bolle, and M. Procesi, *An abstract Nash-Moser theorem with parameters and applications to PDEs*, Annales de l'Institut Henri Poincare (C) Non Linear Analysis **27** (2010), no. 1, 377 – 399.

[12] M. Berti and Ph. Bolle, *Quasi-periodic solutions for Schrödinger equations with Sobolev regularity of NLS on $T^d$ with a multiplicative potential*, to appear on Eur. Jour. Math.

[13] ———, *Quasiperiodic solutions with sobolev regularity for NLS on $\mathbb{T}^d$ and a multiplicative potential.*, J. European Math. Society **15** (2013), no. 1, 229–286.

[14] M. Berti, L. Corsi, and M. Procesi, *An abstract Nash-Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds*, Comm. Math. Phys. **334** (2015), no. 3, 1413–1454.

[15] J. Bourgain, *On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE*, Internat. Math. Res. Notices (1996), no. 6, 277–304.

[16] ———, *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations*, Ann. of Math. (2) **148** (1998), no. 2, 363–439. MR MR1668547 (2000b:37087)

[17] ———, *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies, vol. 158, Princeton University Press, Princeton, NJ, 2005. MR MR2100420 (2005j:35184)

[18] ———, *On invariant tori of full dimension for the 1d periodic nls*, J. Funct. Anal. **229** (2005).

[19] Jean Bourgain, *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices (1994), no. 11, 475ff., approx. 21 pp. (electronic). MR 1316975 (96f:58170)

[20] Ch.-Q. Cheng, *Birkhoff-kolmogorov-arnold-moser tori in convex hamiltonian systems*, Comm. Math. Phys, **177** (1996), no. 3, 529–559.

[21] L. Chierchia and P. Perfetti, *Second order hamiltonian equations on f tsp infty and almost-periodic solutions*, J. Differential Equations **116** (1995).

[22] L. Chierchia and J. You, *Kam tori for 1d nonlinear wave equations with periodic boundary conditions*, Comm. Math. Phys. **211** (2000), 497–525.

[23] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, *Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation*, Invent. Math. **181** (2010), no. 1, 39–113. MR 2651381 (2011f:35320)

[24] A. Constantin and D. Lannes, *The hydrodynamical relevance of the camassa-holm and degasperis-procesi equations*, Arch. Ration. Mech. Anal. **192** (2009), 165–186.

[25] L. Corsi, R. Feola, and G. Gentile, *Lower-dimensional invariant tori for perturbations of a class of non-convex hamiltonian functions*, Journal of Statistical Physics **1** (2013), 156–189.

[26] L. Corsi and G. Gentile, *Resonant tori of arbitrary codimension for quasi-periodically forced systems*, (preprint 2014), arxiv:1403.2872.

[27] Walter Craig and C. Eugene Wayne, *Newton's method and periodic solutions of nonlinear wave equations*, Comm. Pure Appl. Math. **46** (1993), no. 11, 1409–1498.

[28] J. M. Delort, *A quasi-linear birkhoff normal forms method. application to the quasi-linear klein-gordon equation on* $\mathbb{S}^1$, Astérisque **341** (2012).

[29] L. Hakan Eliasson and Sergei B. Kuksin, *KAM for the nonlinear Schrödinger equation*, Ann. of Math. (2) **172** (2010), no. 1, 371–435. MR 2680422 (2011g:37203)

[30] R. Feola, *KAM for a quasi-linear forced Hamiltonian NLS*, In preparation, 2015.

[31] R. Feola and M. Procesi, *Quasi-periodic solutions for fully nonlinear forced reversible schrödinger equations*, Journal of Differential Equations (2014).

[32] Jiansheng Geng and Jiangong You, *A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces*, Comm. Math. Phys. **262** (2006), no. 2, 343–372. MR MR2200264 (2007b:37184)

[33] G. Gentile, *Degenerate lower-dimensional tori under the bryuno condition*, Ergodic Theory Dynam. Systems **27** (2007), 427–457.

[34] G. Gentile, V. Mastropietro, and M. Procesi, *Periodic solutions for completely resonant nonlinear wave equations with Dirichlet boundary conditions*, Comm. Math. Phys. **256** (2005), no. 2, 437–490. MR MR2160800 (2006m:35250)

[35] S. M. Graff, *On the conservation of hyperbolic invariant tori for hamiltonian systems*, Journal of Differential Equations **15** (1974), 1–69.

[36] B. Grébert and T. Kappeler, *Symmetries of the nonlinear schrödinger equation*, Bull. Soc. math. France **130** (2002), 603–618.

[37] G. Iooss, P.I. Plotnikov, and J.F. Toland, *Standing waves on an infinitely deep perfect fluid under gravity*, Arch. Ration. Mech. Anal. **177** (2005), no. 3, 367–478.

[38] T. Kappeler, P. Lohrmann, P. Topalov, and Zung. N.T., *Birkhoff coordinates for the focusing nls equation*, Comm. Math. Phys, **285** (2009).

[39] T. Kappeler and J. Pöschel, *KAM and KdV*, vol. 45, Springer-Verlag, Berlin, 2003.

[40] A. Kolmogorov, *On conservation of conditionally periodic motions for a small change in hamil- ton's function*, Dokl. Akad. Nauk SSSR **98** (1954), 527–530.

[41] S. Kuksin, *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional Anal. i Prilozhen. **21** (1987), no. 3, 22–37.

[42] ______, *A KAM theorem for equations of the Korteweg-de Vries type*, Rev. Math. Phys. **10** (1998), no. 3, 1–64.

[43] ______, *Analysis of hamiltonian pdes*, Oxford University Press,, 2000.

[44] S. Kuksin and G. Perelman, *Vey theorem in infinite dimensions and its application to kdv*, Discrete Contin. Dyn. Syst. **27** (2010), 1–24.

[45] Sergej Kuksin and Jürgen Pöschel, *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Ann. of Math. **143** (1996), no. 1, 149–179. MR MR1370761 (96j:58147)

[46] J. Liu and X. Yuan, *A KAM Theorem for Hamiltonian partial differential equations with unbounded perturbations*, Comm. Math. Phys, **307** (2011), 629–673.

[47] H.P. McKean and K.L. Vaninsky, *Action-angle variables for the cubic schrödinger equation*, CPAM **50** (1997), 489–562.

[48] J. Moser, *Rapidly convergent iteration method and non-linear partial differential equations - i*, Ann. Sc. Norm. Sup. Pisa **20** (1966), no. 2, 265–315.

[49] ______, *Convergent series expansions for quasi-periodic motions*, Math. Ann. **169** (1967), 136–176.

[50] J. Pöschel, *A KAM-theorem for some nonlinear partial differential equations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **23** (1996), no. 1, 119–148. MR MR1401420 (97g:58146)

[51] ______, *On the construction of almost periodic solutions for a nonlinear Schrödinger equation*, Ergodic Theory Dynam. Systems **22** (2002), no. 5, 1537–1549. MR MR1934149 (2005a:37142)

[52] C. Procesi and M. Procesi, *A normal form for the Schrödinger equation with analytic non-linearities*, Communications in Mathematical Physics **312** (2012), no. 2, 501–557.

[53] M. Procesi and C. Procesi, *A KAM algorithm for the non–linear Schrödinger equation*, Advances in Math. **272** (2015), 399–470.

[54] M. Procesi, C. Procesi, and V. Nguyen Bich, *The energy graph of the non–linear Schrödinger equation*, Rend. Lincei Mat. Appl. **24** (2013), 1–73.

[55] Michela Procesi, *Quasi-periodic solutions for completely resonant non-linear wave equations in 1D and 2D*, Discrete Contin. Dyn. Syst. **13** (2005), no. 3, 541–552. MR MR2152330 (2006b:35231)

[56] G. Staffilani, *On the growth of high sobolev norms of solutions for kdv and schrödinger equations.*, Duke Math. J. **86** (1997), 109–142.

[57] A. S. Švarc, *The genus of a fibered space*, Trudy Moskov. Mat. Obšč. **10** (1961), 217–272. MR MR0154284 (27 #4233)

[58] C. Eugene Wayne, *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. **127** (1990), no. 3, 479–528. MR MR1040892 (91b:58236)

[59] X. Xu and M. Procesi, *Quasi-Töplitz Functions in KAM Theorem*, SIAM J. of Math. Anal. **45** (2013), no. 4, 2148 – 2181.

[60] V. E. Zakharov (ed.), *What is integrability?*, Springer, 1991.

[61] J. Zhang, M. Gao, and X. Yuan, *Kam tori for reversible partial differential equations*, Nonlinearity **24** (2011), 1189–1228.