Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 312 (2022) 276–316

Journal of Differential Equations

www.elsevier.com/locate/jde

Quadratic lifespan and growth of Sobolev norms for derivative Schrödinger equations on generic tori ☆

Roberto Feola [a], Riccardo Montalto [b,*]

[a] Dipartimento di Matematica e Fisica, Università degli studi ROMA TRE, Largo San Leonardo Murialdo 1, 00146, Roma, Italy

[b] Dipartimento di Matematica, Università degli Studi di Milano, Via Saldini 50, I-20133, Italy

Received 12 April 2021; accepted 21 December 2021
Available online 4 January 2022

Abstract

We consider a family of Schrödinger equations with unbounded Hamiltonian quadratic nonlinearities on a generic tori of dimension $d \geq 1$. We study the behavior of high Sobolev norms H^s, $s \gg 1$, of solutions with initial conditions in H^s whose H^ρ-Sobolev norm, $1 \ll \rho \ll s$, is smaller than $\varepsilon \ll 1$. We provide a control of the H^s-norm over a time interval of order $O(\varepsilon^{-2})$.

Due to the lack of conserved quantities controlling high Sobolev norms, the key ingredient of the proof is the construction of a modified energy equivalent to the "low norm" H^ρ (when ρ is sufficiently high) over a nontrivial time interval $O(\varepsilon^{-2})$. This is achieved by means of normal form techniques for quasi-linear equations involving para-differential calculus. The main difficulty is to control the possible loss of derivatives due to the small divisors arising form three waves interactions. By performing "tame" energy estimates we obtain upper bounds for higher Sobolev norms H^s.

MSC: 35Q55; 37K10; 35S50

Keywords: Derivative Schrödinger equations; Para-differential calculus; Energy estimates; Normal form theory

☆ *Acknowledgments.* Riccardo Montalto is supported by INDAM-GNFM. The authors warmly thank Dario Bambusi for many useful discussions and comments.

* Corresponding author.
E-mail addresses: roberto.feola@uniroma3.it (R. Feola), riccardo.montalto@unimi.it (R. Montalto).

https://doi.org/10.1016/j.jde.2021.12.018
0022-0396/

Contents

1. Introduction

We consider the derivative Schrödinger equation (DNLS)

$$\partial_t u = \mathrm{i}(\Delta_g u - mu - Q(u,\overline{u}))\,, \quad u = u(t,x)\,,\ \ x \in \mathbb{T}^d := (\mathbb{R}/2\pi\mathbb{Z})^d\,, \quad d \geq 1\,, \tag{1.1}$$

where $m > 0$ is the mass, the operator Δ_g is defined by linearity as

$$\Delta_g e^{\mathrm{i}j\cdot x} = -\|j\|_g^2 e^{\mathrm{i}j\cdot x}\,, \quad \|j\|_g^2 := Gj\cdot j\,, \quad j \in \mathbb{Z}^d\,, \tag{1.2}$$

with $G = (g_{ij})_{i,j=1,\dots,d}$ a strictly positive definite, symmetric, matrix, i.e. $G\xi\cdot\xi \geq c_0|\xi|^2$ for any $\xi \in \mathbb{Z}^d \setminus \{0\}$, and $c_0 > 0$. The nonlinearity $Q(u,\overline{u})$ has the form

$$Q(u,\overline{u}) := (\partial_{\overline{u}} f)(u,\nabla u) - \sum_{i=1}^{d} \partial_{x_i}(\partial_{\overline{u_{x_i}}} f)(u,\overline{u})\,, \tag{1.3}$$

where we denoted $\partial_u := (\partial_{\mathrm{Re}(u)} - \mathrm{i}\partial_{\mathrm{Im}(u)})/2$ and $\partial_{\overline{u}} := (\partial_{\mathrm{Re}(u)} + \mathrm{i}\partial_{\mathrm{Im}(u)})/2$ the Wirtinger derivatives and where $f(y_0, y_1, \dots, y_d) \in C^\infty(\mathbb{C}^{d+1};\mathbb{R})$ (in the *real* sense, i.e. f is C^∞ as function of $\mathrm{Re}(y_i)$, $\mathrm{Im}(y_i)$) is a homogeneous polynomial of degree 3 satisfying

$$\partial_{y_i}\partial_{\overline{y_j}} f = \partial_{\overline{y_i}}\partial_{\overline{y_j}} f \equiv 0\,, \quad \forall\, i,j = 1,\dots,d\,, \quad \forall\,(y_0, y_1, \dots, y_d) \in \mathbb{C}^{d+1}\,. \tag{1.4}$$

Thanks to (1.3) one can note that the equation is Hamiltonian, namely (1.1) can be written as

$$\partial_t u = -\mathrm{i}\nabla_{\overline{u}} H(u,\overline{u})\,, \qquad H(u,\overline{u}) := \int_{\mathbb{T}^d} (\Lambda u)\cdot\overline{u}\,dx + \int_{\mathbb{T}^d} f(u,\overline{u})\,dx\,,$$

where[1] $\Lambda := \Lambda(D) := -\Delta_g + m$ is the Fourier multiplier with symbol

$$\Lambda(\xi) := \|\xi\|_g^2 + m := G\xi \cdot \xi + m\,, \quad \xi \in \mathbb{Z}^d\,. \tag{1.5}$$

The main result of the paper is the following.

Theorem 1.1. *For almost every $m \in (0, +\infty)$ the following holds. There exists $\rho \gg 0$ large enough and $\varepsilon \equiv \varepsilon(\rho) \ll 1$ small enough such that for any initial datum $u_0 \in H^\rho(\mathbb{T}^d; \mathbb{C})$, $\|u_0\|_{H^\rho} \le \varepsilon$, there exists a unique solution $u \in C^0\big([-T_\rho, T_\rho], H^\rho(\mathbb{T}^d; \mathbb{C})\big)$ of (1.1) with $u(0, \cdot) = u_0$, with*

$$\|u(t)\|_{H^\rho} \lesssim_\rho \varepsilon\,, \quad \forall t \in [-T_\rho, T_\rho]\,, \quad T_\rho := \mathtt{c}(\rho)\varepsilon^{-2} \tag{1.6}$$

for some $\mathtt{c}(\rho) \le 1$. Moreover, assume in addition that $u_0 \in H^s(\mathbb{T}^d; \mathbb{C})$, $s \ge \rho$ (without any smallness assumption of $\|u_0\|_{H^s}$). Then $u \in C^0\big([-T_\rho, T_\rho], H^s(\mathbb{T}^d; \mathbb{C})\big)$ of (1.1) which remains bounded on $[-T_\rho, T_\rho]$, namely

$$\|u(t)\|_{H^s} \lesssim_s \|u_0\|_{H^s}\,, \quad \forall t \in [-T_\rho, T_\rho]\,. \tag{1.7}$$

Some comments on the result above are in order.

First notice that equation (1.1) can be seen as a nonlinear Schrödinger equation posed on a torus $\mathbb{T}^d_\Gamma := \mathbb{R}^d/\Gamma$ with arbitrary periodicity lattice Γ. In view of the assumptions in (1.3)-(1.4), we have that $Q(u, \overline{u})$ is a Hamiltonian nonlinearity containing *at most* one spatial derivative of the unknown $u(t, x)$. Hence, as far as we know, Theorem 1.1 provides the first long existence results of solutions for nonlinear Schrödinger equations with derivatives and on a manifold different from the *square* torus.

We also could consider nonlinearities of order m with $m < 2$ but we preferred to write the paper for nonlinearities depending on ∇u, since they are more physical (it is basically the case of magnetic potentials).

The bound (1.6) shows indeed that solutions evolving from sufficiently regular initial data of size $\varepsilon \ll 1$ remain small over a time interval of size $O(\varepsilon^{-2})$. This lifespan which is strictly larger than the time of existence provided by local theory which is of order $O(\varepsilon^{-1})$. In addition to this our result provide a control, on the same time interval, of the growth of high Sobolev norms of solutions of (1.1). The bound (1.7) actually shows that the H^s-Sobolev norm of a solution remains *bounded* by only requiring a smallness conditions on a *low* norm of the initial datum. The second part of the Theorem is a consequence of sharp tame *à priori* estimates on the solutions. Unfortunately the equation has no conserved quantities which control for every time the Sobolev norms H^ρ with $\rho > 1$. Hence we are only able to obtain the bound (1.7) over a long (but finite) time interval, by constructing a modified energy for the H^ρ-norm with normal forms techniques. The result we obtained is in the same spirit of [20] by Delort-Masmoudi.

We require the Hamiltonian assumption on the nonlinearity in order to guarantee the well-posedness of the Cauchy problem associated to (1.1) at least for short time. Actually this hypothesis could be weakened. For more details, we refer for instance to the introduction of [24].

[1] $\nabla_u := (\nabla_{\mathrm{Re}(u)} - \mathrm{i}\nabla_{\mathrm{Im}(u)})/2$ and $\nabla_{\overline{u}} := (\nabla_{\mathrm{Re}(u)} + \mathrm{i}\nabla_{\mathrm{Im}(u)})/2$, ∇ denotes the L^2-gradient.

We also remark that the mass parameter $m > 0$ in (1.5) will be used to provide suitable lower bounds on *three wave interactions*.

Some related literature. We now present some known results on the long time existence and stability for derivative Schrödinger equations.

Local well-posedness. Many authors considered equations of the type (1.1) (even without the assumption (1.4)) in the *Euclidean case* (i.e. when $x \in \mathbb{R}^d$). The first existence result is due to Poppenberg in [36] for a special model in one dimension, later extended by Colin [15] to any dimension. A more general class of quasilinear Schrödinger equation is studied in the pioneering work of Kenig-Ponce-Vega [31]. We also mention a recent paper [32] by Marzuola-Metcalfe-Tataru (see also references therein) which optimize the result in [31] in terms of the regularity of the initial data. The situation drastically changes when the equation is posed on a compact manifold. Indeed, Christ in [16] provides examples of Schrödinger equations with derivatives which are ill-posed on the circle $\mathbb{S}^1$ and well posed on $\mathbb{R}$. We mention that a local existence has been obtained on the circle by Baldi-Haus-Montalto in [1] and by Feola-Iandoli in [24] with different techniques. In [25] the authors extend the latter results to any *squared d-dimensional* tori. Our assumption in (1.3) on the nonlinearity guarantees that the local well-posedness for (1.1) can be obtained in the same spirit of [25].

Global well-posedness. All the aforementioned results regard the *local in time* well-posedness for quasilinear Schrödinger equations. The global well-posedness has been established on $\mathbb{R}^2$ and $\mathbb{R}^3$ by de Bouard-Hayashi-Saut [13] in dimension two and three for small data on a model quasilinear Schrödinger. In [13], dispersive properties of the flow are exploited in order to obtain a control of the Sobolev norms for long time. We also mention the paper [34] by Murphy-Pusateri about the almost global existence for a non-gauge-invariant cubic nonlinear Schrödinger equation on $\mathbb{R}$.

Long time regularity and normal forms. On tori (or more in general on compact manifolds) there are no dispersive effects that could help in controlling the behavior of the solutions for long times. In order to extend the lifespan of solution we use the powerful tool of normal form theory. This approach has been successfully and widely used in the past starting form the study of semi-linear PDEs. Without trying to be exhaustive we quote Bourgain [14], Bambusi [2] and Bambusi-Grébert [4] where the authors considered the Klein-Gordon equation on the circle. They proved almost global existence in the sense that, for any $N \geq 1$ and any initial datum in $H^s(\mathbb{T})$ of size $\varepsilon \ll 1$ with $s \gg 1$ large enough, the solution exist and its H^s-Sobolev norm remains small over a time interval of size $O(\varepsilon^{-N})$. Similar results have been obtained for semilinear equations also in higher space dimension. We refer, for instance, to [3] by Bambusi-Delort-Grébert-Szeftel which considered PDEs on Zoll manifolds (see also [23,21]). Normal form theory for quasi-linear equations has been constructed more recently. We quote Delort [18,19] for the Klein-Gordon on $\mathbb{S}^d$ and Berti-Delort [7] for the gravity capillary water waves on $\mathbb{T}$. For equations like (1.1) we mention [26,27] where it is exploited the fact that (following the ideas of [7]) quasi-linear Schrödinger equations may be reduced to constant coefficients through a *para-composition* generated by a diffeomorphism of the circle.

Normal forms on irrational tori. All the papers mentioned above have in common that the spectrum of the linearized problem at zero has "good separation" properties. This fact depends on the geometry of the eigenvalues of the Laplace-Beltrami operator. On irrational tori, for instance, differences of eigenvalues can accumulate to zero. In this case, one typically gets very weak lower bounds on "small divisors" arising from n-waves interactions (see Appendix A). The same problem occurs for the Klein Gordon equation posed on $\mathbb{T}^d$, $d \geq 2$. In dealing with this problem is out of reach (at the moment), but nevertheless one can obtain partial results. We refer to Delort [17], Fang and Zhang [22], Zhang [38] for the Klein-Gordon, Imekraz in [29] for the Beam equation on $\mathbb{T}^2$ and Feola-Grébert-Iandoli [28]. In this last case a special class of quasi linear Klein Gordon equation is considered. We finally quote the remarkable work on multidimensional periodic water wave by Ionescu-Pusateri [30].

Growth of Sobolev norms for PDEs on tori. For linear Schrödinger equations with time dependent potentials on tori, there are several results providing an upper bound t^ε for the high Sobolev norms of the solutions. On rational tori $\mathbb{T}^d$, we mention the results of Bourgain [9], [10] and Delort [12]. These results have been extended on the irrational torus by Berti and Maspero in [8]. In these aforementioned results, the potential is bounded and the proof basically relies on the so called *Bourgain Lemma*. For Schrödinger equations on irrational tori with unbounded potentials (of order strictly smaller than 2), the upper bound t^ε on the growth of Sobolev norms has been proved in [6]. The proof relies on a Pseudo-differential normal form and on a careful analysis of the resonant vector field, by showing that the flow generated by it is uniformly bounded in time.

For nonlinear Schrödinger equations on tori, by completely different methods, Bourgain [11] proved an upper bound t^s for the H^s-norm of the solutions of nonlinear Schrödinger equations on $\mathbb{T}^2$. This result has been also generalized on more general manifolds in [35] (see also references therein).

Plan of the paper and scheme of the proof. In the remaining part of the introduction, we briefly explain the strategy of our proof.

In order to prove a time of existence of size $O(\varepsilon^{-2})$, we need to perform one step of normal form, in order to remove the quadratic terms. Hence in Section 2.1 we consider symbols which are sums of symbols linear in $u, \overline{u}$ plus symbols which are quadratic in $(u, \overline{u})$. Similarly we define classes of smoothing operators. Since one is able to impose only very weak lower bounds on the three wave interactions (cf. Section A), the normal form procedure requires to use paradifferential calculus. In Section 4, we construct a change of variables $u = \Phi(u)[w]$ ($u(t,x)$ is a smooth solution of (1.1) defined on a time interval $[-T, T]$) which transforms the equation (1.1) into another one which has the form

$$\partial_t w + \mathrm{i}(-\Delta_g + m)w + \mathrm{i}\mathrm{Op}^{bw}(z(u; x, \xi))w + \mathcal{R}(u)w = 0$$

where $\mathcal{R}(u)$ is a smoothing remainder, i.e. $\|\mathcal{R}(u)w\|_{H^{s+N}} \lesssim_{s,N} \|u\|_{H^\rho}\|w\|_{H^s}$ for $N \gg 0$, $\rho \gg N$, $s \gg N$ and the *normal form symbol* $z(U; x, \xi)$ is real and it has the property that its Fourier transform $\widehat{z}(U; k, \xi)$ is non zero if

$$|(\xi; k)| \le \langle \xi \rangle^\delta |k|^{-\tau} \text{ and } |k| \le \langle \xi \rangle^\epsilon$$

cf. Definition 4.3. This normal form step is essentially a *nonlinear analogue* of the method developed in [5], [6] at a linear level.

At this point, in Section 5, we perform a Poincaré-Birkhoff normal form step in order to remove the quadratic terms from the smoothing remainder $\mathcal{R}(u)w$. The loss of derivatives in the estimates of the three wave interactions is then compensated by the fact that the remainder is smoothing. In [6], it is proved that the flow associated to normal form symbols is well defined on H^s and uniformly bounded in time. This fact allows in Section 6 to perform an energy estimate which shows that $\|u(t)\|_{H^\rho} \lesssim_\rho \varepsilon$ for $t \in [-T_\rho, T_\rho]$ with $T_\rho = O(\varepsilon^{-2})$, for some $\rho \gg 1$ large enough, provided the initial datum $\|u_0\|_{H^\rho} \le \varepsilon$ is small enough. If in addition, the initial datum

$u_0 \in H^s$, with $s > \rho$ (but with no smallness assumption on $\|u_0\|_{H^s}$), a bootstrap argument shows that $\|u(t)\|_{H^s} \lesssim_s \|u_0\|_{H^s}$ for any $t \in [-T_\rho, T_\rho]$, implying that there is no growth of *high Sobolev norms* over the time interval $[-T_\rho, T_\rho]$.

We finally remark that if one considers the equation (1.1) with the standard Laplacian, there are no small divisors since if the mass is not an integer, the three wave interactions are bounded from below by a constant. On the other hand, generically (meaning for a generic choice of the matrix G in (1.2)), the three wave interactions accumulate to zero and one is able to prove only very weak non-resonance conditions, see Appendix A for more details.

2. Functional setting

We denote by $H^s(\mathbb{T}^d;\mathbb{C})$ (respectively $H^s(\mathbb{T}^d;\mathbb{C}^2)$) the usual Sobolev space of functions $\mathbb{T}^d \ni x \mapsto u(x) \in \mathbb{C}$ (resp. $\mathbb{C}^2$). We expand a function $u(x)$, $x \in \mathbb{T}^d$, in Fourier series as

$$u(x) = \frac{1}{(2\pi)^{d/2}} \sum_{n\in\mathbb{Z}^d} \widehat{u}(n) e^{\mathrm{i} n\cdot x}\,, \qquad \widehat{u}(n) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{T}^d} u(x) e^{-\mathrm{i} n\cdot x}\, dx\,.$$

We endow $H^s(\mathbb{T}^s;\mathbb{C})$ with the norm

$$\|u\|_s^2 := \|u\|_{H^s}^2 := (\langle D\rangle^s u, \langle D\rangle^s u)_{L^2}\,, \qquad \langle D\rangle e^{\mathrm{i} j\cdot x} = \langle j\rangle e^{\mathrm{i} j\cdot x}\,, \quad \forall\, j \in \mathbb{Z}^d\,,$$

where $\langle j\rangle := \sqrt{|j|^2+1}$ and $(\cdot,\cdot)_{L^2}$ denotes the standard complex L^2-scalar product

$$(u,v)_{L^2} := \int_{\mathbb{T}^d} u\cdot \overline{v}\, dx\,, \qquad \forall\, u,v \in L^2(\mathbb{T}^d;\mathbb{C})\,. \tag{2.1}$$

For $U = (u_1,u_2) \in H^s(\mathbb{T}^d;\mathbb{C}^2)$ we just set $\|U\|_s = \|u_1\|_s + \|u_2\|_s$.

Notation. We shall use the notation $A \lesssim B$ to denote $A \le CB$ where C is a positive constant depending on parameters fixed once for all, for instance d and s. We will emphasize by writing $\lesssim_q$ when the constant C depends on some other parameter q. To shorten the notation we shall write $H^s = H^s(\mathbb{T}^d;\mathbb{C})$.

2.1. *Classes of symbols and operators*

In this section we introduce symbols and operators we shall use along the paper. We follow the notation of [25] but with symbols introduced in [5].

Given a symbol $a(x,\xi)$ of order m, and fixing $\delta \in (0,1)$ (very close to one) we define for any $s \in \mathbb{N}$, the norm $|a|_{m,s}$ as

$$|a|_{m,s} := \sup_{|\alpha_1|+|\alpha_2|\le s}\ \sup_{(x,\xi)\in\mathbb{T}^d\times\mathbb{R}^d} |\partial_x^{\alpha_1}\partial_\xi^{\alpha_2} a(x,\xi)\langle\xi\rangle^{-m+\delta|\alpha_2|}|\,, \tag{2.2}$$

and we define $\mathcal{N}_s^m$ the space of the C^s functions $(x,\xi) \mapsto a(x,\xi)$ such that $|a|_{m,s} < \infty$. If the symbol $a(x,\xi)$ is independent of x, namely it is a Fourier multiplier $a(\xi)$, then the norm is given by

$$|a|_{m,s} := \sup_{|\alpha|\le s} \sup_{\xi\in\mathbb{R}^d} |\partial_\xi^\alpha a(\xi)\langle\xi\rangle^{-m+\delta|\alpha|}| .$$

The following elementary lemma holds.

Lemma 2.1. *Let $m \in \mathbb{R}$, $N, s \in \mathbb{N}$, $a \in \mathcal{N}^m_{s+N}$. Then for any $k \in \mathbb{Z}^d$,*

$$|\widehat{a}(k,\cdot)|_{m,s} \lesssim_N \langle k\rangle^{-N} |a|_{m,s+N} .$$

Proof. By a simple integration by parts, one has

$$\begin{aligned} k_i^N \widehat{a}(k,\xi) &= -\frac{1}{(-\mathrm{i})^N}\frac{1}{(2\pi)^{d/2}}\int_{\mathbb{T}^d} \widehat{a}(x,\xi)\partial_{x_i}^N(e^{-\mathrm{i}k\cdot x})\,dx \\ &= \frac{(-1)^{N+1}}{(-\mathrm{i})^N}\frac{1}{(2\pi)^{d/2}}\int_{\mathbb{T}^d} \partial_{x_i}^N a(x,\xi)e^{-\mathrm{i}k\cdot x}\,dx . \end{aligned}$$

Hence

$$\langle k\rangle^N |\widehat{a}(k,\cdot)|_{m,s} \lesssim_N \max_{i=1,\dots,d}|\partial_{x_i}^N a|_{m,s} \lesssim_N |a|_{m,s+N}$$

and the claimed statement has been proved. □

The Bony-Weyl quantization. Let $0 < \epsilon < 1/2$ and consider a smooth function $\eta : \mathbb{R} \to [0,1]$

$$\eta(\xi) = \begin{cases} 1 & \text{if} |\xi| \le 5/4 \\ 0 & \text{if} |\xi| \ge 8/5 \end{cases} \qquad \text{and define} \qquad \eta_\epsilon(\xi) := \eta(|\xi|/\epsilon) .$$

For a symbol $a(x,\xi)$ in $\mathcal{N}^m_s$ we define its (Weyl) quantization as

$$Op^{bw}(a)h := \frac{1}{(2\pi)^d}\sum_{j\in\mathbb{Z}^d} e^{\mathrm{i}j\cdot x}\sum_{k\in\mathbb{Z}^d}\eta_\epsilon\Big(\frac{|j-k|}{\langle j+k\rangle}\Big)\widehat{a}\Big(j-k,\frac{j+k}{2}\Big)\widehat{h}(k) \tag{2.3}$$

where $\widehat{a}(\eta,\xi)$ denotes the Fourier transform of $a(x,\xi)$ in the variable $x \in \mathbb{T}^d$.

Remark 2.2. Notice that the symbol $\Lambda(\xi)$ in (1.5) belongs to $\mathcal{N}^2_s$ for any $s \in \mathbb{R}$, with $|\Lambda|_{2,s} \lesssim_s 1$. Moreover (recall (1.2)) we have that $-\Delta_g + m = Op^{bw}(\Lambda(\xi))$.

The following results follow by standard paradifferential calculus.

Lemma 2.3. ***(Action of Sobolev spaces).*** *Let $m \in \mathbb{R}$, $s_0 > d/2$. Then for any $s \ge 0$, the linear map*

$$\mathcal{N}^m_{s_0} \to \mathcal{L}(H^{s+m}, H^s), \quad a \mapsto Op^{bw}(a)$$

is continuous, namely

$$\|\mathrm{Op}^{bw}(a)\|_{\mathcal{L}(H^{s+m},H^s)} \lesssim_s |a|_{m,s_0}\,.$$

In the paper we shall deal with symbols in $\mathcal{N}_s^m$ depending nonlinearly on a function $u \in H^s(\mathbb{T}^d;\mathbb{C})$. Let us now introduce the spaces

$$\mathbf{H}^s := \Big(H^s(\mathbb{T}^d;\mathbb{C}) \times H^s(\mathbb{T}^d;\mathbb{C})\Big) \cap \mathcal{U}\,, \qquad \mathcal{U} := \{(u^+,u^-) \in \mathbb{L}^2(\mathbb{T}^d;\mathbb{C}^2) \,:\, \overline{u^+} = u^-\}\,. \tag{2.4}$$

We denote by $B_s(r)$ the ball

$$B_s(r) := \Big\{U = (u,\overline{u}) \in \mathbf{H}^s \,:\, \|U\|_s \le r\Big\}\,.$$

Definition 2.4. (Non-homogeneous symbols). Let $m \in \mathbb{R}$, $p \in \mathbb{N}$. We say that a map $U = (u,\overline{u}) \mapsto a(U;x,\xi)$ belongs to the class Γ_p^m if there exists $s_0 > 0$ such that for any $s \ge s_0$, there exists $r = r(s) \in (0,1)$, $\sigma_s \gg s$ such that the map

$$B_{\sigma_s}(r) \to \mathcal{N}_s^m\,, \quad U \mapsto a(U;x,\xi)\,,$$

is $\mathcal{C}^\infty$-smooth and vanishes at $U = 0$ of order p.

Remark 2.5. (Estimates on non-homogeneous symbols). Clearly by the latter definition, one has the following estimates.

$$|a(U;\cdot)|_{m,s} \lesssim_s \|U\|_{\sigma_s}^p$$

If $n \le p$, $H_1,\dots,H_n \in H^{\sigma_s}$,

$$|d^n a(U;\cdot)[H_1,\dots,H_n]|_{m,s} \lesssim \|U\|_{\sigma_s}^{p-k}\|H_1\|_{\sigma_s}\dots\|H_n\|_{\sigma_s}\,.$$

If $n > p$, then

$$|d^n a(U;\cdot)[H_1,\dots,H_n]|_{m,s} \lesssim \|H_1\|_{\sigma_s}\dots\|H_n\|_{\sigma_s}\,.$$

Definition 2.6. (Linear symbols in $(u,\overline{u})$). Let $m \in \mathbb{R}$. We say that a linear map $U = (u,\overline{u}) \mapsto a(U;x,\xi)$ belongs to the class O_1^m if it is in the class Γ_1^m and the symbol $a(U;x,\xi)$ is linear w.r.t. U, namely it has the form

$$a(U;x,\xi) = \sum_{k\in\mathbb{Z}^d,\sigma\in\{\pm\}} m_\sigma(k,\xi)\widehat{u}^\sigma(k)e^{\sigma \mathrm{i}k\cdot x}$$

where, for any $k \in \mathbb{Z}^d$, we denoted

$$\widehat{u}^\sigma(k) = \widehat{u}(k)\,, \text{ if } \sigma = +\,, \qquad \widehat{u}^\sigma(k) = \overline{\widehat{u}(k)}\,, \text{ if } \sigma = -\,.$$

Remark 2.7. Notice that one has the inclusion $O_1^m \subseteq \Gamma_1^m$.

Definition 2.8. (Symbols). Given $m \in \mathbb{R}$, we say that a symbol $a \in \Sigma_1^m$ if $a = a_l + a_q$ with $a_l \in O_1^m$ and $a_q \in \Gamma_2^m$.

Definition 2.9. (Classes of para-differential operators). (i) We say that a linear operator A is in the class $\mathcal{OB}_\Gamma(m, p)$ if there exists $a \in \Gamma_p^m$ such that $A = \mathrm{Op}^{bw}(a)$.

(ii) We say that a linear operator A is in the class $\mathcal{OB}_O(m)$, if there exists $a \in O_1^m$ such that $A = \mathrm{Op}^{bw}(a)$.

(iii) We say that a linear operator A is in the class $\mathcal{OB}_\Sigma(m)$, if there exists $a \in \Sigma_1^m$ such that $A = \mathrm{Op}^{bw}(a)$.

We now start by defining the classes of smoothing operators that we use in our procedure.

Definition 2.10. (Non-homogeneous smoothing operators). Let $N \in \mathbb{N}$. We say that a map $(U, w) \mapsto \mathcal{R}(U)[w]$ belongs to the class $\mathcal{S}_2(N)$ if there exists $\rho \equiv \rho_N > N$ such that for any $s \geq \rho$ the map

$$B_\rho(r) \to \mathcal{B}(H^s, H^{s+N}), \quad U \mapsto \mathcal{R}(U)$$

is continuous and satisfies the *tame* estimate

$$\|\mathcal{R}(U)\|_{\mathcal{L}(H^s, H^{s+N})} \lesssim_{s,N,\rho} \|U\|_\rho^2 , \quad \forall s \geq \rho . \tag{2.5}$$

Definition 2.11. (Smoothing operators depending linearly on $(u, \overline{u})$). Let $N \in \mathbb{N}$. We say that a **bilinear** map $(u, w) \mapsto \mathcal{R}(u)[w]$ belongs to the class $\mathcal{OS}_1(N)$ if it is of the form

$$\mathcal{R}(u)[w] = \sum_{\xi, k \in \mathbb{Z}^d} r(k, \xi) \widehat{u}(k - \xi) \widehat{w}(\xi) e^{\mathrm{i} x \cdot k} , \tag{2.6}$$

and there exists $\rho \equiv \rho_N > N$ such that the **linear** map $H^\rho \to \mathcal{B}(H^s, H^{s+N})$, $u \mapsto \mathcal{R}(u)$ satisfies the tame estimate

$$\|\mathcal{R}(u)\|_{\mathcal{L}(H^s, H^{s+N})} \lesssim_{s,\rho,N} \|u\|_\rho , \quad \forall s \geq \rho . \tag{2.7}$$

With a slight abuse of terminology we use the same notation for the class of operators of the form $(U, w) \mapsto \mathcal{R}(U)[w] = \mathcal{R}_+(u)[w] + \mathcal{R}_-(\overline{u})[w]$ where $\mathcal{R}_+, \mathcal{R}_- \in \mathcal{OS}_1(N)$.

Definition 2.12. (Smoothing operators). We say that $\mathcal{R}$ is in $\mathcal{S}(N)$ if $\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2$ with $\mathcal{R} \in \mathcal{OS}_1(N)$ and $\mathcal{R}_2 \in \mathcal{S}_2(N)$.

Definition 2.13 ***(Matrix valued symbols and operators).*** (i) Consider a matrix valued symbol

$$A := A(U; x, \xi) := \begin{pmatrix} a(U; x, \xi) & b(U; x, \xi) \\ \overline{b(U; x, -\xi)} & \overline{a(U; x, -\xi)} \end{pmatrix}$$

We say that $A \in \Gamma_p^m$, resp. O_1^m, resp Σ_m^1 if its entries $a, b \in \Gamma_p^m$, resp. O_1^m, resp. Σ_m^1. We then denote by $\mathrm{Op}^{bw}(A)$ the matrix valued operator

$$\mathrm{Op}^{bw}(A) = \begin{pmatrix} \mathrm{Op}^{bw}\Big(a(U;x,\xi)\Big) & \mathrm{Op}^{bw}\Big(b(U;x,\xi)\Big) \\ \mathrm{Op}^{bw}\Big(\overline{b(U;x,-\xi)}\Big) & \mathrm{Op}^{bw}\Big(\overline{a(U;x,-\xi)}\Big) \end{pmatrix} \tag{2.8}$$

and we use the same notations to denote the classes given in the Definition 2.9.

(ii) Similarly if $\mathcal{R}_1, \mathcal{R}_2 \in \mathcal{O}$ where $\mathcal{O} = \mathcal{S}_2(N), \mathcal{OS}_1(N), \mathcal{S}(N)$ we say that

$$\mathcal{R}(U) = \begin{pmatrix} \mathcal{R}_1(U) & \mathcal{R}_2(U) \\ \overline{\mathcal{R}_2(U)} & \overline{\mathcal{R}_1(U)} \end{pmatrix} \tag{2.9}$$

belongs to the class $\mathcal{O}$. Here the operators $\overline{R_j(U)}$, $j = 1, 2$, are defined as

$$\overline{R_j(U)}[h] := \overline{R(U)[\overline{h}]}, \qquad \forall h \in H^s(\mathbb{T}^d; \mathbb{C}). \tag{2.10}$$

One can easily check that a linear operator $\mathcal{R}$ of the form (2.9) (or (2.8)) is *real-to-real* in the sense that it preserves the spaces $\mathbf{H}^s$ (see (2.4)). On the space $\mathbf{H}^0$ we define the scalar product

$$(U, V)_{\mathbf{H}^0} := \int_{\mathbb{T}} U \cdot \overline{V} dx. \tag{2.11}$$

Given an operator $\mathcal{R}$ of the form (2.9) we denote by $\mathcal{R}^*$ its adjoint with respect to the scalar product (2.11), i.e.

$$(\mathcal{R}U, V)_{\mathbf{H}^0} = (U, \mathcal{R}^* V)_{\mathbf{H}^0}, \quad \forall\, U,\ V \in \mathbf{H}^0.$$

One can check that

$$\mathcal{R}^* := \begin{pmatrix} \mathcal{R}_1^* & \overline{\mathcal{R}_2}^* \\ \mathcal{R}_2^* & \overline{\mathcal{R}_1}^* \end{pmatrix},$$

where $\mathcal{R}_1^*$ and $\mathcal{R}_2^*$ are respectively the adjoints of the operators $\mathcal{R}_1$ and $\mathcal{R}_2$ with respect to the complex scalar product on $L^2(\mathbb{T}; \mathbb{C})$ defined in (2.1).

Definition 2.14. Let $\mathcal{R}$ be an operator as in (2.9). We say that $\mathcal{R}$ is *self-adjoint* if

$$\mathcal{R}_1^* = \mathcal{R}_1, \quad \overline{\mathcal{R}_2} = \mathcal{R}_2^*. \tag{2.12}$$

We say that an operator $\mathcal{M}$ as in (2.9) is *Hamiltonian* is $-\mathrm{i}E\mathcal{M}$ is self-adjoint.

Consider now a symbol $a = a(x,\xi) \in \Gamma_p^m$ (resp. O_1^m, resp. Σ_m^1), and set $A := Op^{bw}(a(x,\xi))$. Using (2.3) and (2.10) one can check that

$$\overline{A} = Op^{bw}(\widetilde{a}(x,\xi)), \qquad \widetilde{a}(x,\xi) = \overline{a(x,-\xi)};$$

$$(\textbf{Ajdoint}) \qquad A^* = Op^{bw}\big(\overline{a(x,\xi)}\big).$$

Therefore, a matrix valued paradifferential operator as in (2.8) is self-adjoint according to Definition 2.14 if and only if (recall (2.12)) one has

$$a(x,\xi)=\overline{a(x,\xi)}\,,\qquad b(x,-\xi)=b(x,\xi)\,. \tag{2.13}$$

Definition 2.15 ***(Symplectic map).*** Let $\mathcal{Q}=\mathcal{Q}(U)$ be a matrix valued operator of the form (2.9) (resp. (2.8)). We say that $\mathcal{Q}$ is *symplectic* if

$$\mathcal{Q}^*(-\mathrm{i}E)\mathcal{Q}=-\mathrm{i}E\,,\qquad E=\left[\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right]. \tag{2.14}$$

2.2. Symbolic calculus

In this section we provide some abstract lemmas on the classes that we defined before that we shall apply in our normal form procedure. We introduce the following differential operator

$$\sigma(D_x,D_\xi,D_y,D_\eta):=D_\xi D_y-D_xD_\eta\,,$$

where $D_x:=\frac{1}{\mathrm{i}}\partial_x$ and D_ξ, D_y, D_η are similarly defined.

Definition 2.16. (Asymptotic expansion of composition symbol). Let $\rho\in\mathbb{N}$, $m_1,m_2\in\mathbb{R}$ and $a\in\Sigma_1^{m_1}$, $b\in\Sigma_1^{m_2}$. We define the symbol

$$(a\#_\rho b)(U;x,\xi):=\sum_{k=0}^{\rho-1}\frac{1}{k!}\left(\frac{\mathrm{i}}{2}\sigma(D_x,D_\xi,D_y,D_\eta)\right)^k\Big[a(x,\xi)b(y,\eta)\Big]_{|x=y,\xi=\eta} \tag{2.15}$$

modulo symbols in $\Sigma_1^{m_1+m_2-\rho\delta}$.

Remark 2.17. Recalling (2.2) we note that the symbol

$$\sigma(D_x,D_\xi,D_y,D_\eta)^k\big[a(x,\xi)b(y,\eta)\big]_{|x=y,\xi=\eta}$$

belongs to $\Sigma_1^{m_1+m_2-\delta k}$. In particular we have the expansion $a\#_\rho b=ab+\frac{1}{2\mathrm{i}}\{a,b\}+\Sigma_1^{m_1+m_2-2\delta}$.

We shall prove the following result on the composition of paradifferential operator.

Proposition 2.18. *Fix $\rho\in\mathbb{N}$, $m_1,m_2\in\mathbb{R}$ and $s_0>d/2$. There is $q=q(\rho)\gg1$ such that, for $a\in\mathcal{N}_{s_0+q}^{m_1}$, $b\in\mathcal{N}_{s_0+q}^{m_2}$, one has that*

$$Op^{bw}(a)\circ Op^{bw}(b)=Op^{bw}(a\#_\rho b)+R(a,b) \tag{2.16}$$

where, for any $s\geq s_0>d/2$, the bilinear and continuous map

$$\mathcal{N}_{s_0+q(\rho)}^{m_1}\times\mathcal{N}_{s_0+q(\rho)}^{m_2}\to\mathcal{L}(H^s,H^{s-m_1-m_2+\rho}),\quad (a,b)\mapsto R(a,b)$$

satisfies

$$\|R(a,b)h\|_{s-m_1-m_2+\rho} \lesssim_s |a|_{m_1,s_0+q(\rho)}|b|_{m_2,s_0+q(\rho)}\|h\|_s\,, \quad \forall h \in H^s\,. \tag{2.17}$$

Proof. In order to prove the lemma above we reason as follows. First of all notice that[2]

$$\mathcal{F}(Op^{bw}(a)\circ Op^{bw}(b)h)(\xi) = \sum_{\eta,\theta\in\mathbb{Z}^d} r_1(\xi,\theta,\zeta)\widehat{a}\big(\xi-\theta,\frac{\xi+\theta}{2}\big)\widehat{b}\big(\theta-\zeta,\frac{\theta+\zeta}{2}\big)\widehat{h}(\zeta)\,, \tag{2.18}$$

where

$$r_1(\xi,\theta,\zeta) := \eta_\epsilon\left(\frac{|\xi-\theta|}{|\xi+\theta|}\right)\eta_\epsilon\left(\frac{|\theta-\zeta|}{|\theta+\zeta|}\right). \tag{2.19}$$

Fix $L\in\mathbb{N}$ with $L\gg\rho$ to be chosen later. By Taylor expanding the symbols we have

$$\begin{aligned}\widehat{a}\Big(\xi-\theta,\frac{\xi+\theta}{2}\Big) &= \sum_{k=0}^{L}\frac{1}{2^k i^k k!}\widehat{(\partial_\xi^k a)}\Big(\xi-\theta,\frac{\xi+\zeta}{2}\Big)[\mathrm{i}(\theta-\zeta)]^k + \\ &\quad + \frac{1}{2^{L+1}\mathrm{i}^{L+1}L!}\int_0^1(1-\tau)^L\widehat{(\partial_\xi^{L+1}a)}\Big(\xi-\theta,\frac{\xi+\zeta}{2}+\tau\frac{\theta-\zeta}{2}\Big)[\mathrm{i}(\theta-\zeta)]^{L+1}d\tau,\end{aligned} \tag{2.20}$$

$$\begin{aligned}\widehat{b}\Big(\theta-\zeta,\frac{\theta+\zeta}{2}\Big) &= \sum_{j=0}^{L}\frac{(-1)^j}{2^j i^j j!}\widehat{(\partial_\xi^j b)}\Big(\theta-\zeta,\frac{\xi+\zeta}{2}\Big)[\mathrm{i}(\xi-\theta)]^j + \\ &\quad + \frac{(-1)^{L+1}}{2^{L+1}\mathrm{i}^{L+1}L!}\int_0^1(1-\tau)^L\widehat{(\partial_\xi^{L+1}b)}\Big(\theta-\zeta,\frac{\xi+\zeta}{2}+\tau\frac{\theta-\xi}{2}\Big)[\mathrm{i}(\xi-\theta)]^{L+1}d\tau\,.\end{aligned} \tag{2.21}$$

Therefore we deduce that

$$\widehat{a}\Big(\xi-\theta,\frac{\xi+\theta}{2}\Big)\widehat{b}\Big(\theta-\zeta,\frac{\theta+\zeta}{2}\Big) = \sum_{\ell=1}^{4} g_\ell(\xi,\theta,\zeta) \tag{2.22}$$

where

$$\begin{aligned}g_1(\xi,\theta,\zeta) := \sum_{p=0}^{L}\frac{1}{2^p\mathrm{i}^p p!}\sum_{k=0}^{p}\binom{p}{k}(-1)^{p-k}\widehat{(\partial_\xi^k\partial_x^{p-k}a)}\Big(\xi-\theta,\frac{\xi+\zeta}{2}\Big)\times \\ \times\widehat{(\partial_\xi^{p-k}\partial_x^k b)}\Big(\theta-\zeta,\frac{\xi+\zeta}{2}\Big),\end{aligned} \tag{2.23}$$

[2] We denote the Fourier transform in $x\in\mathbb{T}^d$ of a function $f(x)$ by $\mathcal{F}(f)(\xi)=\widehat{f}(\xi)$.

$$g_2(\xi,\theta,\zeta) := \sum_{p=L+1}^{2L} \frac{1}{2^p \mathrm{i}^p p!} \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} (\widehat{\partial_\xi^k \partial_x^{p-k} a})\Big(\xi-\theta, \frac{\xi+\zeta}{2}\Big) \times \\ \times (\widehat{\partial_\xi^{p-k} \partial_x^k b})\Big(\theta-\zeta, \frac{\xi+\zeta}{2}\Big), \tag{2.24}$$

$$g_3(\xi,\theta,\zeta) := \frac{\widehat{b}\Big(\theta-\zeta, \frac{\theta+\zeta}{2}\Big)}{2^{\rho+1}\mathrm{i}^{\rho+1}\rho!} \int_0^1 (1-\tau)^\rho (\widehat{\partial_\xi^{\rho+1} a})\Big(\xi-\theta, \frac{\xi+\zeta}{2} + \tau\frac{\theta-\zeta}{2}\Big)[\mathrm{i}(\theta-\zeta)]^{\rho+1} d\tau\,, \tag{2.25}$$

$$g_4(\xi,\theta,\zeta) := \sum_{k=0}^{\rho} \frac{(-1)^{\rho+1}}{2^{\rho+1}\mathrm{i}^{\rho+1}\rho!} \int_0^1 (1-\tau)^\rho (\widehat{\partial_\xi^{\rho+1} b})\Big(\theta-\zeta, \frac{\xi+\zeta}{2} + \tau\frac{\theta-\xi}{2}\Big)[\mathrm{i}(\xi-\theta)]^{\rho+1} d\tau \times \\ \times \frac{1}{2^k i^k k!} (\widehat{\partial_\xi^k a})\Big(\xi-\theta, \frac{\xi+\zeta}{2}\Big)[\mathrm{i}(\theta-\zeta)]^k\,. \tag{2.26}$$

We set

$$Op^{bw}(a) \circ Op^{bw}(b) = \sum_{\ell=1}^{4} R_\ell \tag{2.27}$$

where the operators R_ℓ are defined by

$$\widehat{R_\ell h}(\xi) = \sum_{\zeta,\theta\in\mathbb{Z}^d} r_1(\xi,\theta,\zeta) g_\ell(\xi,\theta,\zeta)\widehat{h}(\zeta)\,, \qquad \ell = 1,\dots,4\,, \tag{2.28}$$

where r_1 is in (2.19) and g_ℓ are in (2.23)-(2.26).

We now study the explicit form of the symbol $(a\#_\rho b)(x,\xi)$ (recall (2.15)). First of all we note that (formally)

$$\frac{1}{p!}\Big[\frac{\mathrm{i}}{2}\sigma(D_x, D_\xi, D_y, D_\eta)\Big]^p = \frac{1}{2^p\mathrm{i}^p p!}(\partial_\xi\partial_y - \partial_x\partial_\eta)^p \\ = \frac{1}{2^p\mathrm{i}^p p!} \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} (\partial_\xi\partial_y)^k (\partial_x\partial_\eta)^{p-k}\,.$$

Then it is easy to note that (using (2.15) and (2.23))

$$\mathcal{F}(a\#_\rho b)(\xi-\zeta, \frac{\xi+\zeta}{2}) = \sum_{\theta\in\mathbb{Z}^d} g_1(\xi,\theta,\zeta)\,.$$

Hence we have that $Op^{bw}(a\#_\rho b)h =: Qh$ has the form

$$\widehat{Qh}(\xi) := \sum_{\zeta\in\mathbb{Z}^d} r_2(\xi,\zeta)\mathcal{F}(a\#_\rho b)(\xi-\zeta, \frac{\xi+\zeta}{2})\widehat{h}(\zeta) = \sum_{\zeta,\theta\in\mathbb{Z}^d} \chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big) g_1(\xi,\theta,\zeta)\widehat{h}(\zeta)\,, \tag{2.29}$$

where

$$r_2(\xi,\zeta) := \eta_\epsilon\Big(\frac{|\xi-\zeta|}{\langle \xi+\zeta\rangle}\Big). \tag{2.30}$$

In conclusion, by (2.27), (2.28) and (2.29), we obtained

$$Op^{bw}(a)\circ Op^{bw}(b) = Op^{bw}(a\#_\rho b) + \mathcal{R} + \sum_{\ell=2}^{4} R_\ell$$

where

$$\begin{aligned}\widehat{(\mathcal{R}h)}(\xi) &:= \mathcal{F}\big((R_1-Q)h\big)(\xi) := \sum_{\zeta,\theta\in\mathbb{Z}^d}\mathcal{R}(\xi,\theta,\zeta)\widehat{h}(\zeta)\\ \mathcal{R}(\xi,\theta,\zeta) &\overset{(2.19)}{:=} \Big[\eta_\epsilon\Big(\frac{|\xi-\theta|}{|\xi+\theta|}\Big)\eta_\epsilon\Big(\frac{|\theta-\zeta|}{|\theta+\zeta|}\Big) - \eta_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big)\Big]g_1(\xi,\theta,\zeta)\end{aligned} \tag{2.31}$$

To obtain the (2.16) it remains to show that the terms $\mathcal{R}$, R_ℓ, $\ell=2,3,4$, satisfy the estimate (2.17).

We start by considering the remainder $\mathcal{R}$ in (2.31). First of all, using the explicit formula (2.23) for the coefficients $g_1(\xi,\theta,\zeta)$ and reasoning as in Lemma 2.1, we deduce that

$$|g_1(\xi,\theta,\zeta)| \lesssim \langle\xi-\theta\rangle^{-p}\langle\theta-\zeta\rangle^{-q}|a|_{m_1,p+L}|b|_{m_2,q+L}\langle\xi+\zeta\rangle^{m_1+m_2}, \tag{2.32}$$

for any $p,q\in\mathbb{N}$. We now study the properties of the cut-off function $(r_1-r_2)(\xi,\theta,\zeta)$ (see (2.19), (2.30)) appearing in (2.31). Let us define the sets

$$\begin{aligned} D &:= \Big\{(\xi,\theta,\zeta)\in\mathbb{Z}^{3d} \,:\, (r_1-r_2)(\xi,\theta,\zeta)=0\Big\},\\ A &:= \Big\{(\xi,\theta,\zeta)\in\mathbb{Z}^{3d} \,:\, \frac{|\xi-\theta|}{\langle\xi+\theta\rangle}\le\frac{5\epsilon}{4},\ \frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\le\frac{5\epsilon}{4},\ \frac{|\theta-\zeta|}{\langle\theta+\zeta\rangle}\le\frac{5\epsilon}{4}\Big\},\\ B &:= \Big\{(\xi,\theta,\zeta)\in\mathbb{Z}^{3d} \,:\, \frac{|\xi-\theta|}{\langle\xi+\theta\rangle}\ge\frac{8\epsilon}{5},\ \frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\ge\frac{8\epsilon}{5},\ \frac{|\theta-\zeta|}{\langle\theta+\zeta\rangle}\ge\frac{8\epsilon}{5}\Big\}.\end{aligned}$$

We note that

$$D\supseteq A\cup B \quad\Rightarrow\quad D^c\subseteq A^c\cap B^c.$$

Let $(\xi,\theta,\zeta)\in D^c$ and assume in particular that $(\xi,\theta,\zeta)\in \mathrm{Supp}(r_1) := \overline{\{(\xi,\theta,\zeta): r_1\neq 0\}}$. Then we can note that

$$|\xi-\zeta| \ll \langle\xi+\zeta\rangle \quad\text{and}\quad \langle\xi\rangle\sim\langle\zeta\rangle. \tag{2.33}$$

Notice also that $(\xi,\theta,\zeta)\in\mathrm{Supp}(r_2)$ implies the (2.33) as well. We need to estimate

$$\|R_0 h\|^2_{s+\rho-m_1-m_2} \lesssim \sum_{\xi\in\mathbb{Z}^d} \Big(\sum^{*}_{\zeta,\theta} |g_1(\xi,\theta,\zeta)||\widehat{h}(\zeta)|\langle\xi\rangle^{s+\rho} \Big)^2 = I + II + III ,$$

where $\sum^*_{\zeta,\theta}$ denotes the sum over indexes satisfying (2.33), the term I denotes the sum on indexes satisfying also $|\xi-\theta| > c\epsilon|\xi|$, II denotes the sum on indexes satisfying also $|\zeta-\theta| > c\epsilon|\zeta|$, for some $0 < c \ll 1$ and III is defined by difference. We estimate the term I. By using (2.33), $|\xi-\theta| > c\epsilon|\xi|$ and (2.32), we get

$$\begin{aligned} I &\lesssim \sum_{\xi\in\mathbb{Z}^d} \Big(\sum^{*}_{\zeta,\theta} |g_1(\xi,\theta,\zeta)||\widehat{h}(\zeta)|\langle\zeta\rangle^{s}\langle\xi-\theta\rangle^{\rho-m_1-m_2} \Big)^2 \\ &\lesssim |a|^2_{m_1,s_0+\rho+L}|b|^2_{m_2,s_0+L}\| |\widehat{h}(\xi)|\langle\xi\rangle^s \star \langle\xi\rangle^{s_0+\rho} \star \langle\xi\rangle^{-s_0} \|^2_{\ell^2(\mathbb{Z}^d)} \\ &\lesssim |a|^2_{m_1,s_0+\rho+L}|b|^2_{m_2,s_0+L}\| |\widehat{h}(\xi)|\langle\xi\rangle^s \|^2_{\ell^2(\mathbb{Z}^d)} \lesssim |a|^2_{m_1,s_0+\rho+L}|b|^2_{m_2,s_0+L}\|h\|^2_s , \end{aligned}$$

where we used $s_0 > d > d/2$.

Reasoning similarly one obtains $II \lesssim \|h\|^2_s |a|^2_{m_1+s_0+L}|b|^2_{m_2,s_0+\rho+L}$. The sum III is restricted to indexes satisfying (2.33) and $|\xi-\theta| \le c\epsilon|\xi|$, $|\zeta-\theta| \le c\epsilon|\zeta|$. For $c \ll 1$ small enough this restriction implies that $(\xi,\theta,\zeta) \in A$, which is a contradiction since $(\xi,\theta,\zeta) \in D^c \subseteq A^c$.

For the remainders R_ℓ, $\ell = 2,3,4$ in (2.28) one can reason similarly using the explicit formulæ (2.24)-(2.26) to show that g_ℓ are symbols of order at least $L+1$ or ρ. Therefore one concludes the proof by choosing L large enough. □

By the Proposition above we deduce the following.

Lemma 2.19. ***(Compositions and commutators).*** *(i) Let $a \in \Sigma^m_1$, $b \in \Sigma^{m'}_1$ and let $N \in \mathbb{N}$. Then the operator $\mathrm{Op}^{bw}(a) \circ \mathrm{Op}^{bw}(b)$ satisfies*

$$\mathrm{Op}^{bw}(a) \circ \mathrm{Op}^{bw}(b) = \mathrm{Op}^{bw}(ab + \frac{1}{2\mathrm{i}}\{a,b\}) + \mathrm{Op}^{bw}(r_{ab}) + \mathcal{R}_{ab}(U) ,$$

where $r_{ab} \in \Gamma^{m+m'-2\delta}_2$ and the map $(U,w) \mapsto \mathcal{R}_{ab}(U)[w]$ belongs to the class $\mathcal{S}_2(N)$. As a consequence, the commutator

$$[\mathrm{Op}^{bw}(a), \mathrm{Op}^{bw}(b)] = \frac{1}{\mathrm{i}}\mathrm{Op}^{bw}(\{a,b\}) + \mathrm{Op}^{bw}(r_{ab} - r_{ba}) + \mathcal{R}_{ab}(U) - \mathcal{R}_{ba}(U) .$$

(ii) Let $a \in \Sigma^m_1$ and $N \in \mathbb{N}$. Then, recalling (1.5), the Poisson bracket $\{\Lambda, a\} \in \Sigma^{m+1}_1$ and

$$\begin{aligned} \mathrm{Op}^{bw}(\Lambda) \circ \mathrm{Op}^{bw}(a) &= \mathrm{Op}^{bw}(\Lambda a) + \frac{1}{2\mathrm{i}}\mathrm{Op}^{bw}(\{\Lambda,a\}) + \mathrm{Op}^{bw}(r_{\Lambda a}) + \mathcal{R}_{\Lambda a}(U) , \\ [\mathrm{Op}^{bw}(\Lambda), \mathrm{Op}^{bw}(a)] &= \frac{1}{\mathrm{i}}\mathrm{Op}^{bw}(\{\Lambda,a\}) + \mathrm{Op}^{bw}(r_{\Lambda a} - r_{a\Lambda}) + \mathcal{R}_{\Lambda a}(U) - \mathcal{R}_{a\Lambda}(U) , \end{aligned}$$

where $r_{\Lambda a} \in \Sigma^{m+1-2\delta}_1$ and $\mathcal{R}_{\Lambda a}(U), \mathcal{R}_{a\Lambda}(U)$ are in $\mathcal{S}(N)$.

Proof. It follows by Proposition 2.18, using formula (2.15). The homogeneity expansions of symbols and remainders can be deduced by the formulæ in the proof of the Proposition. □

We also have the following result about the composition between the smoothing operators introduced in Definition 2.10-2.12.

Lemma 2.20. *Let $N \in \mathbb{N}$, $m \in \mathbb{R}$, $a \in \Sigma_1^m$ and $R, Q \in \mathcal{S}(N)$. Then one has*
(i) $R(U) \circ Q(U)$ and $Q(U) \circ R(U)$ are smoothing operators in $\mathcal{S}_2(N)$.
(ii) $R(U) \circ Op^{bw}(a(U; x, \xi))$, $Op^{bw}(a(U; x, \xi)) \circ R(U)$ are in $\mathcal{S}_2(N - m)$.

Proof. By Definition 2.12 we can write $R = R_1 + R_2$, $Q = Q_1 + Q_2$ for some $R_1, Q_1 \in \mathcal{OS}_1(N)$ and $R_2, Q_2 \in \mathcal{S}_2(N)$ (see Definition 2.10-2.11). Then item (i) follows by using estimates (2.5) and (2.7). Item (ii) follows similarly by using also Lemma 2.3 and Remark 2.5. □

3. Technical lemmata

3.1. Flows and conjugations

In this section we prove some abstract results about the conjugation of paradifferential operators and smoothing remainders under flows.

Consider a real symbol $g \in \Sigma_1^m$ with $m < 1$ and the flow $\Phi_g^\tau(U)$, $\tau \in [-1, 1]$ defined by

$$\begin{cases} \partial_\tau \Phi_g^\tau(U) = \mathrm{i} G(U) \Phi_g^\tau(U) \,, \qquad G(U) := Op^{bw}(g(U; x, \xi)) \,, \\ \Phi^0(U) = \mathbb{1} \,. \end{cases} \tag{3.1}$$

We have the following.

Lemma 3.1. ***(Linear flows).*** *There are $s_0 > d/2$ and $r > 0$ such that, for any $U = \left[\begin{smallmatrix} u \\ \overline{u} \end{smallmatrix}\right]$ with $u \in H^s(\mathbb{T}^d; \mathbb{C}) \cap B_{s_0}(r)$, for any $s > 0$ the problem (3.1) admits a unique solution $\Phi_g^\tau(U)$ satisfying*

$$\begin{aligned} \|\Phi_g^\tau(U) w\|_s \le \|w\|_s (1 + C(s) \|u\|_\rho) \,, \qquad & \forall\, w \in H^s(\mathbb{T}^d; \mathbb{C}) \,, \\ \|(\Phi_g^\tau(U) - \mathrm{Id}) v\|_s \lesssim_s \|u\|_\rho \|v\|_{s+m} \,, \qquad & \forall\, v \in H^{s+m}(\mathbb{T}^d; \mathbb{C}) \,, \end{aligned} \tag{3.2}$$

for some $C(s) > 0$, uniformly in $\tau \in [0, 1]$. The map (see (2.4))

$$\mathbf{\Phi}_g^\tau(U) := \begin{pmatrix} \Phi_g^\tau(U) \\ \overline{\Phi_g^\tau(U)} \end{pmatrix} : \mathbf{H}^s \to \mathbf{H}^s \tag{3.3}$$

is symplectic according to Definition 2.15.

Proof. The result follows by a standard energy estimate using the fact that the symbol $g(U; x, \xi)$ is *real* valued. For more details we refer to Lemma 3.22 in [7]. The map $\mathbf{\Phi}_g^\tau$ in (3.3) can be seen as the linear flow generated by the field $\mathcal{G}(U) = \mathrm{i} E \mathbb{1} G(U)$. Therefore one can check that it is symplectic by reasoning as in Lemma 2.1 in [24]. □

We set $\Phi_g(U) := \Phi^1_g(U)$ and its inverse $\Phi_g(U)^{-1} := \Phi^\tau_g(U)_{|\tau=-1}$. The following lemma holds.

Lemma 3.2. ***(Conjugation of operators under paradifferential flows).*** *Let $g \in \Sigma^n_1$ with $n < \delta$ and assume that $g(U; x, \xi)$ is a real symbol. Then the following holds.*

(i) If $a \in \Sigma^m_1$, for any fixed $N \in \mathbb{N}$, one has

$$\Phi_g(U)^{-1}\mathrm{Op}^{bw}(a)\Phi_g(U) = \mathrm{Op}^{bw}(a) + \mathrm{Op}^{bw}(b) + \mathcal{R}(U)\,,$$

where $b \in \Gamma^{m+n-\delta}_2$, $\mathcal{R} \in \mathcal{S}_2(N)$. If the symbol a is real valued, then b is real valued as well.

(ii) For any fixed $N \in \mathbb{N}$, one has (see (1.5))

$$\Phi_g(U)^{-1}\mathrm{Op}^{bw}(\Lambda)\Phi_g(U) = \mathrm{Op}^{bw}(\Lambda) + \mathrm{Op}^{bw}(\{\Lambda, g\}) + \mathrm{Op}^{bw}(b) + \mathcal{R}(U)\,,$$

where b is a real valued symbol in $\Sigma^{n+1-(\delta-n)}_1$ and $\mathcal{R} \in \mathcal{S}(N)$.

(iii) Let $\mathcal{R}$ be in $\mathcal{S}(N)$. Then $\mathcal{R}_1(U) := \Phi_g(U)^{-1}\mathcal{R}(U)\Phi_g(U)$ is in the class $\mathcal{S}(N-n)$.

Proof. *Item* (i). Using (3.1) we get, for $L \geq 3$, the Lie expansion

$$\begin{aligned}\Phi g(U)^{-1} Op^{bw}(a)\Phi_g(U) = {} & Op^{bw}(a) + \big[Op^{bw}(a), Op^{bw}(\mathrm{i}g)\big] + \\ & + \sum_{k=2}^{L} \frac{(-1)^k}{k!}\mathrm{Ad}^k_{Op^{bw}(\mathrm{i}g)}[Op^{bw}(a)] + \\ & + \frac{(-1)^{L+1}}{L!}\int_0^1 (1-\theta)^L \Phi^{-\theta}(U)\big(\mathrm{Ad}^{L+1}_{Op^{bw}(\mathrm{i}g)}[Op^{bw}(a)]\big)\Phi^{\theta}(U)d\theta\,,\end{aligned} \tag{3.4}$$

where we defined $\mathrm{Ad}_G[A] := [G, A]$ and $\mathrm{Ad}^k_G[A] := \mathrm{Ad}_G\big[\mathrm{Ad}^{k-1}_G[A]\big]$ for $k \geq 2$. By Lemma 2.19 (possibly replacing the smoothing index N by some $\tilde{N}$ chosen below large enough) and Remark 2.17 we get

$$\mathrm{Ad}_{Op^{bw}(\mathrm{i}g)}[Op^{bw}(a)] = \big[Op^{bw}(\mathrm{i}g), Op^{bw}(a)\big] = Op^{bw}\big(\{g, a\} + r_1\big), \quad r_1 \in \Sigma^{m+n-2\delta}_1\,,$$

up to a smoothing operator in $\mathcal{S}_2(\tilde{N} - m - n)$. Similarly, by induction, for $k \geq 2$ we have

$$\mathrm{Ad}^k_{Op^{bw}(\mathrm{i}g)}[Op^{bw}(a)] = Op^{bw}(b_k), \quad b_k \in \Sigma^{k(n-\delta)+m}_1\,,$$

up to a smoothing operator in $\mathcal{S}_2(\tilde{N} - m - kn)$. We choose L in such a way that $(L + 1)(\delta - n) - m \geq \rho$ and $L + 1 \geq 3$, so that the operator $Op^{bw}(b_{L+1})$ belongs to $\mathcal{S}_2(N)$. The integral Taylor remainder in (3.4) belongs to $\mathcal{S}_2(N)$ as well by item (iii) that we proved above. Then we choose $\tilde{N}$ large enough so that $\tilde{N} - m - (L + 1)n \geq N$ and the remainders are N-smoothing. Assume now that $a \in \Sigma^m_1$ is real valued. Using formula (2.15) one can check that also the symbol b constructed through the expansion above is real valued.

Item (ii) follows by reasoning as done for item by replacing a with the symbol $\Lambda(\xi) := \|\xi\|_g^2 + m$. Item (iii) follows by using estimates (2.5), (2.7) on the remainder $\mathcal{R}$ and the second estimate in (3.2) on the map $\Phi_g(U)$. This concludes the proof. □

Consider now a smooth vector field $X_{NLS} : \mathbf{H}^s \to \mathbf{H}^{s-2}$ (see (2.4)) satisfying, for $s \gg 1$,

$$
\begin{aligned}
\|X_{NLS}(U)\|_{s-2} &\lesssim_s \|u\|_s(1+\|u\|_s)\,, && \forall\, U = \begin{bmatrix} u \\ \bar{u} \end{bmatrix} \in \mathbf{H}^s\,,\\
\|dX_{NLS}(U)[H_1]\|_{s-2} &\lesssim_s \|H_1\|_s(1+\|u\|_s)\,, && \forall\, U, H_1 \in \mathbf{H}^s\,,\\
\|d^n X_{NLS}(U)[H_1,\dots,H_n]\|_{s-2} &\lesssim_s \|H_1\|_s \dots \|H_n\|_s\,, && \forall\,, U, H_1,\dots,H_n \in \mathbf{H}^s\,,\ \ n \geq 2\,.
\end{aligned}
\tag{3.5}
$$

Lemma 3.3. *Let $g \in \Sigma_1^m$ and assume that $U(t,x)$ is a solution belonging to $C^0([0,T];\mathbf{H}^s)$, $T > 0$, $s \gg 1$ of the Schrödinger equation $\partial_t U = X_{NLS}(U)$. Then $\partial_t \psi(U(t);x,\xi) = a_\psi(U(t);x,\xi)$ where the symbol $a_\psi(U;x,\xi)$ belongs to the class Σ_1^m with estimates uniform in $t \in [0,T]$.*

Proof. One has that

$$
\partial_t \psi(U(t);x,\xi) = d\psi(U(t);x,\xi)[\partial_t U] = d\psi(U(t);x,\xi)[X_{NLS}(U(t))]\,.
$$

Hence the symbol a_ψ is defined by $a_\psi(U;x,\xi) := d\psi(U;x,\xi)[X_{NLS}(U)]$. Then the result follows by using Remark 2.5 and estimates (3.5). □

Lemma 3.4. ***(Conjugation of ∂_t under paradifferential flows).*** *Let $g \in \Sigma_1^n$ with $n < \delta$ and $g(U;x,\xi)$. Consider a vector field X_{NLS} satisfying (3.5). Assume that $\partial_t U(t) = X_{NLS}(U(t))$ and $U \in C^0([0,T];\mathbf{H}^s)$ for some $T > 0$, $s \gg 1$. Then for any $N \in \mathbb{N}$*

$$
\Phi_g(U(t))^{-1} \circ \partial_t \circ \Phi_g(U(t)) = \partial_t + \mathrm{Op}^{bw}(b(U(t);x,\xi)) + \mathcal{R}(U(t))\,,
$$

where $b(U;x,\xi)$ is a purely imaginary symbol in Σ_1^n and the map $(U,w) \mapsto \mathcal{R}(U)[w]$ is in the class $\mathcal{S}(N)$.

Proof. Fix $L \geq 3$. By classical Lie expansions we obtain

$$
\begin{aligned}
\Phi g(U)^{-1}\partial_t \Phi_g(U) &= \partial_t + Op^{bw}(\mathrm{i}\partial_t g) + \sum_{k=2}^{L} \frac{(-1)^{k-1}}{k!} \mathrm{Ad}^{k-1}_{Op^{bw}(\mathrm{i}g)}[Op^{bw}(\mathrm{i}\partial_t g)] \\
&\quad + \frac{(-1)^L}{L!} \int_0^1 (1-\theta)^L \Phi^{-\theta}(U)\big(\mathrm{Ad}^L_{Op^{bw}(\mathrm{i}g)}[Op^{bw}(\mathrm{i}\partial_t g)]\big)\Phi^{\theta}(U)d\theta\,,
\end{aligned}
$$

where we used that $[\partial_t, Op^{bw}(\mathrm{i}g)] = Op^{bw}(\mathrm{i}\partial_t g)$. By Lemma 3.3 we have that $\partial_t g$ is a symbol in Σ_1^n with estimates uniform in $t \in [0,T]$. The one concludes arguing as done in Lemma 3.2. □

In our procedure, we also need to consider the maps of the form $\Phi_\psi(U) := \Phi^1_\psi(U)$, $\Phi_{\mathcal{F}}(U) := \Phi^1_{\mathcal{F}}(U)$ where $\Phi^\tau_\psi(U)$, $\Phi^\tau_{\mathcal{F}}(U)$, $\tau \in [0,1]$ are given by

$$\partial_\tau \Phi^\tau_\psi(U) = \mathrm{i}Op^{bw}\begin{pmatrix} 0 & \psi(U;x,\xi) \\ -\overline{\psi(U;x,-\xi)} & 0 \end{pmatrix}\Phi^\tau_\psi(U), \quad \Phi^0_\psi(U) = \mathbb{1}, \ \psi \in \Sigma_1^{-n}, \ n \in \mathbb{N} \tag{3.6}$$

$$\partial_\tau \Phi^\tau_{\mathcal{F}}(U) = \mathcal{F}(U)\Phi^\tau_{\mathcal{F}}(U), \quad \Phi^0_{\mathcal{F}}(U) = \mathbb{1}, \qquad \mathcal{F}(U) \in \mathcal{OS}_1(N). \tag{3.7}$$

We only state the conjugacy properties with the flows $\Phi^\tau_\psi(U)$ and $\Phi_{\mathcal{F}}(U)$. The proofs can be done arguing as in Lemmata 3.1-3.4, with the obvious modifications.

Lemma 3.5. *There are $s_0 > d/2$ and $r > 0$ such that, for any $U = \left[\begin{smallmatrix} u \\ \overline{u} \end{smallmatrix}\right]$ with $u \in H^s(\mathbb{T}^d;\mathbb{C}) \cap B_{s_0}(r)$, for any $s > s_0$ the problems (3.6) and (3.7) admit unique solutions $\Phi^\tau_\psi(U)$, $\Phi^\tau_{\mathcal{F}}(U)$ satisfying*

$$\|\Phi^\tau_\psi(U)\|_{\mathcal{L}(H^s)} \le 1 + C(s)\|u\|_{s_0}, \quad \|\Phi^\tau_\psi(U) - \mathrm{Id}\|_{\mathcal{L}(H^s, H^{s+n})} \lesssim_s \|u\|_{s_0}$$
$$\|\Phi^\tau_{\mathcal{F}}(U)\|_{\mathcal{L}(H^s)} \le 1 + C(s)\|u\|_{s_0}, \quad \|\Phi^\tau_{\mathcal{F}}(U) - \mathrm{Id}\|_{\mathcal{L}(H^s, H^{s+N})} \lesssim_s \|u\|_{s_0}, \quad \forall s \ge s_0,$$

uniformly in $\tau \in [-1, 1]$. Moreover the maps $\Phi^\tau_\psi(U)$, $\Phi^\tau_{\mathcal{F}}(U)$ are symplectic according to Definition 2.15.

Proof. It follows by standard theory of ODEs in Banach space. □

We set $\Phi_\psi(U) := \Phi^1_\psi(U)$ and $\Phi_{\mathcal{F}}(U) := \Phi^1_{\mathcal{F}}(U)$. Their inverse is given by $\Phi_\psi(U)^{-1} = \Phi^\tau_\psi(U)_{|\tau=-1}$ and $\Phi_{\mathcal{F}}(U)^{-1} = \Phi^\tau_{\mathcal{F}}(U)_{|\tau=-1}$. The following Lemmata can be deduced by reasoning exactly as done in Lemmata 3.2, 3.3 and 3.4. Hence we omit their proofs.

Lemma 3.6. (i) *Let A be a matrix valued symbol in Σ_1^m as in the Definition 2.13 and let $\Phi_\psi(U)$ as in (3.6). Then for any $N \in \mathbb{N}$,*

$$\Phi_\psi(U)^{-1}\mathrm{Op}^{bw}(A)\Phi_\psi(U) = \mathrm{Op}^{bw}(A) + \mathrm{Op}^{bw}(B) + \mathcal{R}(U)$$

where $B \in \Gamma_2^{m-n}$ and the $\mathcal{R}(U)$ belongs to the class $\mathcal{S}_2(N)$. If the matrix of symbols A satisfies the conditions (2.13), then the matrix B satisfies (2.13) as well.

(ii) *One has that*

$$\begin{aligned}\Phi_\psi(U)^{-1}\mathrm{i}E\mathrm{Op}^{bw}(\Lambda)\Phi_\psi(U) = {} & \mathrm{i}E\mathrm{Op}^{bw}(\Lambda) \\ & + Op^{bw}\begin{pmatrix} 0 & -2\Lambda(\xi)\psi(U;x,\xi) \\ 2\Lambda(-\xi)\overline{\psi(U;x,-\xi)} & 0 \end{pmatrix} \\ & + \mathrm{Op}^{bw}(B) + \mathcal{R}(U)\end{aligned}$$

where B is a matrix in Σ_1^{1-n} satisfying (2.13) and $\mathcal{R}(U)$ is in the class $\mathcal{S}(N)$.

(iii) *Assume that $U \in C^0([0,T];\mathbf{H}^s)$ for some $T > 0$, $s \gg 1$, solves $\partial_t U(t) = X_{NLS}(U(t))$ where X_{NLS} satisfies (3.5). Then for any $N \in \mathbb{N}$*

$$\Phi_\psi(U(t))^{-1} \circ \partial_t \circ \Phi_\psi(U(t)) = \partial_t + \mathrm{Op}^{bw}(B(U(t);x,\xi))) + \mathcal{R}(U(t))$$

where $B(U;x,\xi)\in\Sigma_1^{-n}$ and the map $(U,w)\mapsto\mathcal{R}(U)[w]$ is in the class $\mathcal{S}(N)$. Moreover the matrix of symbols $-\mathrm{i}EB$ satisfies the conditions (2.13).

(iv) Let $\mathcal{R}(U)$ be in the class $\mathcal{S}(N)$. Then $\Phi_\psi(U)^{-1}\mathcal{R}(U)\Phi_\psi(U)$ is in the class $\mathcal{S}(N)$.

Lemma 3.7. *Let $N\in\mathbb{N}$, $\mathcal{F}\in\mathcal{OS}_1(N)$. Then the following holds*

(i) Let $A\in\Sigma_1^m$ be a matrix valued symbol. Then for any $\tau\in[-1,1]$,

$$\Phi_{\mathcal{F}}(U)^{-1}\mathrm{Op}^{bw}(A)\Phi_{\mathcal{F}}(U)=\mathrm{Op}^{bw}(A)+\mathcal{R}(U)$$

where $\mathcal{R}(U)\in\mathcal{S}_2(N-m)$.

(ii) One has that

$$\Phi_{\mathcal{F}}(U)^{-1}\mathrm{i}E\mathrm{Op}^{bw}(\Lambda)\Phi_{\mathcal{F}}(U)=\mathrm{i}E\mathrm{Op}^{bw}(\Lambda)+[\mathrm{i}E\mathrm{Op}^{bw}(\Lambda),\mathcal{F}(U)]+\mathcal{R}(U)$$

for some $\mathcal{R}(U)$ in the class $\mathcal{S}_2(N-2)$.

(iii) Assume that $U\in C^0([0,T],H^{\rho+2})\cap C^1([0,T],H^{\rho})$ solves the equation $\partial_t U=X_{NLS}(U)$ for some $\rho\equiv\rho_N\geq N$ large enough where X_{NLS} satisfies (3.5). Then

$$\Phi_{\mathcal{F}}(U)^{-1}\circ\partial_t\circ\Phi_{\mathcal{F}}(U)=\partial_t-\mathcal{F}(\mathrm{i}E\mathrm{Op}^{bw}(\Lambda)U(t))+\mathcal{R}(U(t))$$

where $\mathcal{R}(U)$ belongs to the class $\mathcal{S}_2(N)$.

(iv) Let $N'\in\mathbb{N}$, $\mathcal{R}(U)$ be in the class $\mathcal{S}(N')$. Then

$$\Phi_{\mathcal{F}}(U)^{-1}\mathcal{R}(U)\Phi_{\mathcal{F}}(U)=\mathcal{R}(U)+\mathcal{Q}(U)$$

where $\mathcal{Q}(U)$ is in the class $\mathcal{S}_2(N+N')$.

3.2. Some calculus about smoothing operators

In this section we prove some abstract results on linear smoothing operators introduced in Definition 2.11. These results will be used in Section 5 and they are based on the estimates on the small divisors proved in Appendix A.

Lemma 3.8. *Let $\mathcal{G}\in(0,+\infty)$ be the full Lebesgue measure set given by Lemma A.1. Then for any $m\in\mathcal{G}$ the following holds. Let*

$$\mathcal{R}(u)w=\sum_{k,\xi\in\mathbb{Z}^d}r(k,\xi)\widehat{u}(k-\xi)\widehat{w}(\xi)e^{\mathrm{i}k\cdot x}$$

be in the class $\mathcal{OS}_1(N)$. Define for any $\sigma,\sigma'\in\{+1,-1\}$, the operator $\mathcal{F}_{\sigma,\sigma'}(u)$ as

$$\mathcal{F}_{\sigma,\sigma'}(u)[w]:=-\sum_{k,\xi\in\mathbb{Z}^d}\frac{r(k,\xi)}{\mathrm{i}\Big(\Lambda(k)+\sigma\Lambda(k-\xi)+\sigma'\Lambda(\xi)\Big)}\widehat{u}(k-\xi)\widehat{w}(\xi)e^{\mathrm{i}k\cdot x}\,. \tag{3.8}$$

Then $\mathcal{F}_{\sigma,\sigma'}(u)$ is in the class $\mathcal{OS}_1(N-\tau)$ and solves the equation

$$\mathrm{Op}^{bw}(\Lambda)\mathcal{F}_{\sigma,\sigma'}(u)+\sigma\mathcal{F}_{\sigma,\sigma'}(\mathrm{Op}^{bw}(\Lambda)u)+\sigma'\mathcal{F}_{\sigma,\sigma'}(u)\mathrm{Op}^{bw}(\Lambda)+\mathcal{R}(u)=0\,. \tag{3.9}$$

Proof. We prove the claimed statement in the case where $\sigma = \sigma' = -1$. The other cases can be proved similarly. It is immediate to verify that $\mathcal{F}_{\sigma,\sigma'}$ defined in (3.8) solves the equation (3.9). Since $\mathcal{R}(u)$ is in the class $\mathcal{OS}_1(N)$, one has that there exists $\rho \equiv \rho_N > N$ large enough such that for any $s \geq \rho$,

$$\sum_{k\in\mathbb{Z}^d} \langle k\rangle^{2(s+N)} \Big| \sum_{\xi\in\mathbb{Z}^d} r(k,\xi)\widehat{u}(k-\xi)\widehat{w}(\xi)\Big|^2 = \|\mathcal{R}(u)w\|_{s+N}^2 \lesssim_s \|u\|_\rho^2 \|w\|_s^2 .$$

By taking $w(x) = e^{\mathrm{i}x\cdot\xi}$, the latter estimate implies

$$\sum_{k\in\mathbb{Z}^d} \langle k\rangle^{2(s+N)} |r(k,\xi)|^2 |\widehat{u}(k-\xi)|^2 \lesssim_s \|u\|_\rho^2 \langle\xi\rangle^{2s} , \quad \forall s \geq \rho , \quad \forall u \in H^\rho . \tag{3.10}$$

Let $\mathcal{F}(u) := \mathcal{F}_{-1,-1}(u)$ (see (3.8)). One has that

$$\mathcal{F}(u)[w] = \sum_{k,\xi\in\mathbb{Z}^d} f(k,\xi)\widehat{u}(k-\xi)\widehat{w}(\xi)e^{\mathrm{i}k\cdot\xi} ,$$

$$f(k,\xi) := -\frac{r(k,\xi)}{\mathrm{i}\Big(\Lambda(k) - \Lambda(k-\xi) - \Lambda(\xi)\Big)} , \qquad k,\xi \in \mathbb{Z}^d .$$

By the bounds (A.2) given by Lemma A.1, one has that there exists $\tau = \tau(d) \gg 0$, large enough and $\gamma \in (0,1)$ small enough such that

$$|\Lambda(k) - \Lambda(k-\xi) - \Lambda(\xi)| \geq \frac{\gamma}{\langle k-\xi\rangle^\tau \langle\xi\rangle^\tau} , \quad \forall k,\xi \in \mathbb{Z}^d ,$$

and therefore

$$|f(k,\xi)| \lesssim \langle k-\xi\rangle^\tau \langle\xi\rangle^\tau |r(k,\xi)| . \tag{3.11}$$

We now estimate the norm $\|\mathcal{F}(u)w\|_{s-\tau+N}$. Take $s-\tau \geq \rho$ in such a way that (3.10) holds with $s-\tau$ instead of s. Using the Cauchy-Schwartz inequality (using that $\sum_\xi \langle k-\xi\rangle^{-2s_0} \leq C < \infty$ for $s_0 > d/2$), one has

$$\begin{aligned}
\|\mathcal{F}(u)w\|_{s-\tau+N}^2 &\leq \sum_{k\in\mathbb{Z}^d} \langle k\rangle^{2(s-\tau+N)} \Big(\sum_{\xi\in\mathbb{Z}^d} |f(k,\xi)||\widehat{u}(k-\xi)||\widehat{w}(\xi)|\Big)^2 \\
&\overset{(3.11)}{\lesssim} \sum_{k\in\mathbb{Z}^d} \langle k\rangle^{2(s-\tau+N)} \Big(\sum_{\xi\in\mathbb{Z}^d} |r(k,\xi)|\langle k-\xi\rangle^\tau |\widehat{u}(k-\xi)|\langle\xi\rangle^\tau |\widehat{w}(\xi)|\Big)^2 \\
&\lesssim \sum_{k,\xi\in\mathbb{Z}^d} \langle k\rangle^{2(s-\tau+N)} |r(k,\xi)|^2 \langle k-\xi\rangle^{2(\tau+s_0)} |\widehat{u}(k-\xi)|^2 \langle\xi\rangle^{2\tau} |\widehat{w}(\xi)|^2 \\
&= \sum_{k,\xi\in\mathbb{Z}^d} \langle k\rangle^{2(s-\tau+N)} |r(k,\xi)|^2 |\widehat{\langle D\rangle^{\tau+s_0} u}(k-\xi)|^2 |\widehat{\langle D\rangle^\tau w}(\xi)|^2
\end{aligned}$$

$$= \sum_{\xi\in\mathbb{Z}^d} |\widehat{\langle D\rangle^\tau w}(\xi)|^2 \Big(\sum_{k\in\mathbb{Z}^d} \langle k\rangle^{2(s-\tau+N)} |r(k,\xi)|^2 |\widehat{\langle D\rangle^{\tau+s_0} u}(k-\xi)|^2 \Big)$$

$$\overset{(3.10)}{\lesssim_s} \|\langle D\rangle^{\tau+s_0} u\|_\rho^2 \sum_{\xi\in\mathbb{Z}^d} \langle\xi\rangle^{2(s-\tau)} |\widehat{\langle D\rangle^\tau w}(\xi)|^2 \lesssim_s \|u\|^2_{\rho+\tau+s_0} \|w\|_s^2 .$$

The latter chain of inequalities implies that for any $s \geq \rho' := \rho + \tau + s_0$, the map

$$H^{\rho'} \to \mathcal{L}(H^s, H^{s+N-\tau}), \quad u \mapsto \mathcal{F}(u)$$

is a bounded linear map. This implies that $\mathcal{F}$ belongs to the class $\mathcal{OS}_1(N-\tau)$. $\square$

As a consequence of the latter lemma, one gets the following.

Lemma 3.9. *Let $\mathcal{G} \in (0,+\infty)$ be the full Lebesgue measure set given by Lemma A.1. Then for any $m \in \mathcal{G}$ the following holds. Let $\mathcal{R} \in \mathcal{OS}_1(N)$ be a matrix valued operator. Then there exists a matrix valued operator $\mathcal{F} \in \mathcal{OS}_1(N-\tau)$ which solves the equation*

$$-\mathcal{F}(\mathrm{i}E\mathrm{Op}^{bw}(\Lambda)U) + [\mathrm{i}E\mathrm{Op}^{bw}(\Lambda), \mathcal{F}(U)] + \mathcal{R}(U) = 0 . \tag{3.12}$$

Proof. The operator $\mathcal{R} \in \mathcal{OS}_1(N)$ has the form

$$\mathcal{R}(U) = \begin{pmatrix} \mathcal{R}_1(U) & \mathcal{R}_2(U) \\ \overline{\mathcal{R}_2(U)} & \overline{\mathcal{R}_1(U)} \end{pmatrix} = \begin{pmatrix} \mathcal{R}_1^+(u) + \mathcal{R}_1^-(\overline{u}) & \mathcal{R}_2^+(u) + \mathcal{R}_2^-(\overline{u}) \\ \overline{\mathcal{R}_2^+(u) + \mathcal{R}_2^-(\overline{u})} & \overline{\mathcal{R}_1^+(u) + \mathcal{R}_1^-(\overline{u})} \end{pmatrix} .$$

One looks for $\mathcal{F} \in \mathcal{OS}_1(N-\tau)$ of the same form, namely

$$\mathcal{F}(U) = \begin{pmatrix} \mathcal{F}_1^+(u) + \mathcal{F}_1^-(\overline{u}) & \mathcal{F}_2^+(u) + \mathcal{F}_2^-(\overline{u}) \\ \overline{\mathcal{F}_2^+(u) + \mathcal{F}_2^-(\overline{u})} & \overline{\mathcal{F}_1^+(u) + \mathcal{F}_1^-(\overline{u})} \end{pmatrix} .$$

A direct calculation shows that the equation (3.12) is equivalent to

$$\begin{aligned}
&\mathrm{i}\Big(-\mathcal{F}_1^+(\mathrm{Op}^{bw}(\Lambda)u) + [\mathrm{Op}^{bw}(\Lambda), \mathcal{F}_1^+(u)]\Big) + \mathcal{R}_1^+(u) = 0,\\
&\mathrm{i}\Big(\mathcal{F}_1^-(\mathrm{Op}^{bw}(\Lambda)\overline{u}) + [\mathrm{Op}^{bw}(\Lambda), \mathcal{F}_1^-(\overline{u})]\Big) + \mathcal{R}_1^-(\overline{u}) = 0 ,\\
&\mathrm{i}\Big(-\mathcal{F}_2^+(\mathrm{Op}^{bw}(\Lambda)u) + \mathrm{Op}^{bw}(\Lambda)\mathcal{F}_2^+(u) + \mathcal{F}_2^+(u)\mathrm{Op}^{bw}(\Lambda)\Big) + \mathcal{R}_2^+(u) = 0 ,\\
&\mathrm{i}\Big(\mathcal{F}_2^-(\mathrm{Op}^{bw}(\Lambda)\overline{u}) + \mathrm{Op}^{bw}(\Lambda)\mathcal{F}_2^-(\overline{u}) + \mathcal{F}_2^-(\overline{u})\mathrm{Op}^{bw}(\Lambda)\Big) + \mathcal{R}_2^-(\overline{u}) = 0 .
\end{aligned}$$

The claimed statement then directly follows by Lemma 3.8. $\square$

4. Paradifferential normal form

4.1. Paralinearization of the Schrödinger equation

In this section we rewrite the equation (1.1) as a paradifferential system. From now on we shall assume the following hypothesis:

- **Hypothesis on local in time solutions.** There exists $\rho \gg 0$ large enough and $T > 0$ such that

$$
\begin{aligned}
& u \in C^0\big([-T,T], H^{\rho}\big) \cap C^1([-T,T], H^{\rho-2}), \\
& \sup_{t\in[-T,T]} \|u(t)\|_{\rho} + \sup_{t\in[-T,T]} \|\partial_t u(t)\|_{\rho-2} \lesssim_{\rho} \varepsilon ,
\end{aligned}
\tag{4.1}
$$

  solves the equation (1.1).

The latter hypothesis is actually guaranteed by the local existence theorem proved in [25]. The only small difference is that the standard Laplacian on the torus is replaced by the more general elliptic operator (1.5), but the proof can be done exactly in the same way, with the obvious small modifications.

Proposition 4.1. ***(Paralinearization of NLS).*** *We have that the equation* (1.1) *is equivalent to the following system:*

$$
\partial_t U + \mathrm{i} E\, Op^{bw}\big(\Lambda(\xi)\big)U + \mathcal{A}(U)U + \mathcal{R}(U)U = 0\,, \qquad U = \begin{bmatrix} u \\ \overline{u} \end{bmatrix}, \quad \mathbb{1} := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{4.2}
$$

where E is in (2.14)*, $\Lambda(\xi)$ is in* (1.5)*, the operator $\mathcal{A}(U)$ is in $\mathcal{OB}_{\Sigma}(1)$ (see Definition* 2.9*) and has the form*

$$
\mathcal{A}(U) := \mathrm{i} E\, Op^{bw}\begin{pmatrix} a(U;x,\xi) & b(U;x,\xi) \\ \overline{b(U;x,-\xi)} & \overline{a(U;x,-\xi)} \end{pmatrix}, \tag{4.3}
$$

where

$$
a(U;x,\xi) := \sum_{j=1}^{d} \Big[\mathrm{i}(\partial_{\overline{u}u_{x_j}} f - \partial_{\overline{u}_{x_j}u} f)\xi_j + \frac{1}{2}\big(- \partial_{x_j}(\partial_{\overline{u}_{x_j}u} f) - \partial_{x_j}(\partial_{\overline{u}u_{x_j}} f)\big)\Big] + \partial_{u\overline{u}} f\,,
$$

$$
b(U;x,\xi) := - \sum_{j=1}^{d} \partial_{x_j}(\partial_{\overline{u}\overline{u}_{x_j}} f) + \partial_{u\overline{u}} f\,.
$$

The remainder $\mathcal{R}(U)$ is a matrix of smoothing operators in the class $\mathcal{O}S_1(N)$. Finally the operator $\mathcal{A}(U)$ is Hamiltonian according to Definition 2.14*.*

Proof. By paralinearizing the nonlinearity $\mathcal{Q}$ in (1.3), using the Bony paralinearization formula (see [33], [37]) and recalling the assumption (1.4), one obtains that

$$
\begin{aligned}
\mathcal{Q}(u,\overline{u}) &= Op^{bw}(\partial_{u\overline{u}}f)[u] + Op^{bw}(\partial_{\overline{u}\overline{u}}f)[\overline{u}] \\
&\quad + \sum_{j=1}^{d}\Big(Op^{bw}(\partial_{\overline{u}u_{x_j}}f)[u_{x_j}] + Op^{bw}(\partial_{\overline{u}\overline{u}_{x_j}}f)[\overline{u}_{x_j}]\Big) \\
&\quad - \sum_{j=1}^{d}\partial_{x_j}\Big[Op^{bw}(\partial_{\overline{u}_{x_j}u}f)[u] + Op^{bw}(\partial_{\overline{u}\overline{u}_{x_j}}f)[\overline{u}]\Big] + \mathcal{R}(u,\overline{u}),
\end{aligned}
\tag{4.4}
$$

where the remainder $\mathcal{R}(u,\overline{u})$ is smoothing, namely it satisfies

$$
\|\mathcal{R}(u,\overline{u})\|_{s+N} \lesssim_{s,N} \|u\|_{\rho}\|u\|_{s} \tag{4.5}
$$

for any $N \geq 1$, for some $\rho = \rho_N > N$ and $s \geq \rho$. Recall that $\partial_{x_j} = Op^{bw}(\mathrm{i}\xi_j)$, $j = 1,\dots,d$. Therefore, using Proposition 2.18 and Lemma 2.19, we have

$$
\begin{aligned}
& Op^{bw}(\partial_{\overline{u}u_{x_j}}f)[u_{x_j}] - \partial_{x_j}Op^{bw}(\partial_{\overline{u}_{x_j}u}f)[u] = \\
&\quad = Op^{bw}(\partial_{\overline{u}u_{x_j}}f)\circ Op^{bw}(\mathrm{i}\xi_j)[u] - Op^{bw}(\mathrm{i}\xi_j)\circ Op^{bw}(\partial_{\overline{u}_{x_j}u}f)[u] \\
&\quad = Op^{bw}\Big(\mathrm{i}(\partial_{\overline{u}u_{x_j}}f - \partial_{\overline{u}_{x_j}u}f)\xi_j\Big)[u] + Op^{bw}\Big(\frac{1}{2\mathrm{i}}\{\partial_{\overline{u}u_{x_j}}f, \mathrm{i}\xi_j\} - \frac{1}{2\mathrm{i}}\{\mathrm{i}\xi_j, \partial_{\overline{u}_{x_j}u}f\}\Big)[u] \\
&\quad = Op^{bw}\Big(\mathrm{i}(\partial_{\overline{u}u_{x_j}}f - \partial_{\overline{u}_{x_j}u}f)\xi_j\Big)[u] + \frac{1}{2}Op^{bw}\Big(-\partial_{x_j}(\partial_{\overline{u}_{x_j}u}f) - \partial_{x_j}(\partial_{\overline{u}u_{x_j}}f)\Big)[u],
\end{aligned}
$$

up to some smoothing remainder satisfying (4.5). Reasoning similarly we deduce

$$
\begin{aligned}
& Op^{bw}(\partial_{\overline{u}\overline{u}_{x_j}}f)[\overline{u}_{x_j}] - \partial_{x_j}Op^{bw}(\partial_{\overline{u}\overline{u}_{x_j}}f)[\overline{u}] = \\
&\quad = Op^{bw}(\partial_{\overline{u}\overline{u}_{x_j}}f)\circ Op^{bw}(\mathrm{i}\xi_j)[\overline{u}] - Op^{bw}(\mathrm{i}\xi_j)\circ Op^{bw}(\partial_{\overline{u}\overline{u}_{x_j}}f)[\overline{u}] \\
&\quad = -Op^{bw}\Big(\partial_{x_j}(\partial_{\overline{u}\overline{u}_{x_j}}f)\Big)[\overline{u}],
\end{aligned}
$$

up to some smoothing remainder satisfying (4.5). Therefore, by (4.4), we obtained

$$
\begin{aligned}
\mathcal{Q}(u,\overline{u}) &= \sum_{j=1}^{d} Op^{bw}\Big(\mathrm{i}(\partial_{\overline{u}u_{x_j}}f - \partial_{\overline{u}_{x_j}u}f)\xi_j\Big)[u] \\
&\quad + Op^{bw}\Big(\partial_{u\overline{u}}f - \frac{1}{2}\sum_{j=1}^{d}\big(\partial_{x_j}(\partial_{\overline{u}_{x_j}u}f) + \partial_{x_j}(\partial_{\overline{u}u_{x_j}}f)\big)\Big)[u] \\
&\quad + Op^{bw}\Big(\partial_{\overline{u}\overline{u}}f - \sum_{j=1}^{d}\partial_{x_j}(\partial_{\overline{u}\overline{u}_{x_j}}f)\Big)[\overline{u}] + \mathcal{R}(u,\overline{u}),
\end{aligned}
$$

where $\mathcal{R}(u,\overline{u})$ is some remainder satisfying (4.5). By writing (1.1) as a system in $U = (u,\overline{u})^T$, one gets the (4.2)-(4.3). By an explicit computation one can check that the operator $\mathcal{A}(U)$ is Hamiltonian. $\square$

4.2. Diagonalization up to a smoothing remainder

In this section we analyze the para-differential operator

$$\mathcal{P}(U) := \partial_t + \mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi)) + \mathcal{A}(U)$$

where $\mathcal{A}(U)$ is in (4.3) and the symbol $\Lambda(\xi)$ is in (1.5). We prove the following result.

Proposition 4.2. *Let $N \in \mathbb{N}$, $s_0 \gg d/2$. Then there exists $\rho = \rho_N > N$, s_0 large enough such that if (4.1) holds, then the following holds. There exists a linear symplectic invertible transformation $\Phi^{(1)}(U) : \mathbf{H}^s \to \mathbf{H}^s$ such that*

$$\mathcal{P}^{(1)}(U) := \Phi^{(1)}(U)^{-1}\mathcal{P}(U)\Phi^{(1)}(U) = \partial_t + \mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi)) + \mathcal{A}^{(1)}(U) + \mathcal{R}^{(1)}(U) \tag{4.6}$$

where

$$\mathcal{A}^{(1)}(U) = \mathrm{i}Op^{bw}\begin{pmatrix} a^{(1)}(U;x,\xi) & 0 \\ 0 & -a^{(1)}(U;x,-\xi) \end{pmatrix}$$

with $a^{(1)}(U;x,\xi)$ a real symbol in the class Σ_1^1 and the remainder $\mathcal{R}^{(1)}(U) \in \mathcal{S}(\rho, N)$ is Hamiltonian. Moreover, for any $s \geq \rho$, one has

$$\|\Phi^{(1)}(U)^{\pm 1} - \mathrm{Id}\|_{\mathcal{L}(H^s)} \lesssim_s \|u\|_\rho .$$

Proof. The proposition is proved by means of an iterative procedure. At the step n of such a procedure one has an operator

$$\mathcal{P}_n(U) = \partial_t + \mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi)) + \mathcal{A}_n(U) + \mathcal{B}_n(U) + \mathcal{R}_n(U),$$

where

$$\mathcal{A}_n(U) = \mathrm{i}Op^{bw}\begin{pmatrix} a_n(U;x,\xi) & 0 \\ 0 & -a_n(U;x,-\xi) \end{pmatrix}, \tag{4.7}$$

$$\mathcal{B}_n(U) := \mathrm{i}Op^{bw}\begin{pmatrix} 0 & b_n(U;x,\xi) \\ -\overline{b_n(U;x,-\xi)} & 0 \end{pmatrix}, \tag{4.8}$$

$a_n \in \Sigma_1^1$ and $b_n \in \Sigma_1^{-n}$. The remainder $\mathcal{R}_n(U)$ is a linear Hamiltonian operator, smoothing of order $-N$ in the class $\mathcal{S}(\rho, N)$ for some $\rho \equiv \rho_N > N$ large enough. We consider

$$\Phi_n(U) := \exp(\mathrm{i}\Psi_n(U))$$

where $\Psi_n(U)$ is a para-differential operator of the form

$$\Psi_n(U) := \mathrm{i}Op^{bw}\begin{pmatrix} 0 & \psi_n(U;x,\xi) \\ -\overline{\psi_n(U;x,-\xi)} & 0 \end{pmatrix},$$

where $\psi_n(U;x,\xi)$ is a symbol of order $-n-2$ which has to be determined appropriately. Notice that, for any n, the map $\Phi_n(U)$ has the same form of $\Phi_\psi(U)$ in (3.6). Hence it is well-posed and symplectic by Lemma 3.5. We actually choose the symbol $\psi_n(U;x,\xi)$ in such a way that

$$-2\Lambda(\xi)\psi_n(U;x,\xi)+b_n(U;x,\xi)=0\,,\quad \text{hence we set}\quad \psi_n(U;x,\xi):=\frac{b_n(U;x,\xi)}{2\Lambda(\xi)}\,. \tag{4.9}$$

Clearly, since $b_n\in\Sigma_1^{-n}$, then $\psi_n\in\Sigma_1^{-n-2}$. Since $\Phi_n(U)$ is symplectic the transformed operator

$$\mathcal{P}_{n+1}(U)=\Phi_n(U)^{-1}\mathcal{P}_n(U)\Phi_n(U)$$

is Hamiltonian. By Lemma 3.6, and using (4.9), one gets that

$$\mathcal{P}_{n+1}(U)=\partial_t+\mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi))+\mathcal{A}_{n+1}(U)+\mathcal{B}_{n+1}(U)+\mathcal{R}_{n+1}(U)$$

where $\mathcal{A}_{n+1}(U)$, $\mathcal{B}_{n+1}(U)$ are operators of the form (4.7), (4.8) with $n\rightsquigarrow n+1$ for some symbols $a_{n+1}\in\Sigma_1^1$ and $b_{n+1}\in\Sigma_1^{-n-1}$. Moreover the symbol a_{n+1} is real valued by Lemma 3.6. The remainder $\mathcal{R}_{n+1}(U)$ is a linear Hamiltonian operator, smoothing of order $-N$ in the class $\mathcal{S}(\rho,N)$ for some $\rho\equiv\rho_N>N$ large enough. The proof of the lemma is then concluded. □

4.3. Normal form on the diagonal term

In this section we prove a normal form theorem on the operator $\mathcal{P}^{(1)}(U)$ in (4.6) which is a para-differential version of the normal form procedure developed in [6]. Moreover we denote by $(\xi;k)$ the scalar product induced by the matrix G (see (1.2)), namely $(\xi;k):=G\xi\cdot k$. We start with the following definition.

Definition 4.3. (Normal form symbols). A symbol $z(x,\xi)$ in $\mathcal{N}_s^m$ is said to be *in normal form* (with parameters δ,ϵ,τ) if

$$z(x,\xi)=\sum_{k\in\mathbb{Z}^d}\widehat{z}(k,\xi)e^{\mathrm{i}k\cdot x}$$

satisfies

$$\widehat{z}(k,\xi)\neq 0\quad\Longrightarrow\quad |(\xi;k)|\le\langle\xi\rangle^{\delta}|k|^{-\tau}\ \text{and}\ |k|\le\langle\xi\rangle^{\epsilon}$$

for any $k\neq 0$, $\xi\in\mathbb{R}^d$.

We shall fix appropriately the parameters $\varepsilon,\delta\in(0,1)$, $\tau>0$ as

$$\frac{2}{3}<\delta<1\,,\quad \tau>d-1\,,\quad 0<\epsilon<\frac{\delta}{\tau+1} \tag{4.10}$$

cf. [5]. The main result of this section is the following.

Proposition 4.4. *Let $N \in \mathbb{N}$, $s_0 \gg d/2$. Then there exists $\rho = \rho_N > N$, s_0 large enough such that if (4.1) holds, then the following holds. There exists a linear symplectic invertible transformation $\Phi^{(2)}(U) : \mathbf{H}^s \to \mathbf{H}^s$ such that*

$$\mathcal{P}^{(2)}(U) := \Phi^{(2)}(U)^{-1}\mathcal{P}^{(1)}(U)\Phi^{(2)}(U) = \partial_t + \mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi)) + \mathcal{Z}(U) + \mathcal{R}^{(2)}(U)$$

where $\mathcal{P}_1(U)$ is in (4.6), the operator $\mathcal{Z}(U)$ has the form

$$\mathcal{Z}(U) = \mathrm{i}Op^{bw}\begin{pmatrix} z(U;x,\xi) & 0 \\ 0 & -z(U;x,-\xi), \end{pmatrix} \tag{4.11}$$

with $z(U;x,\xi)$ a real symbol in the class Σ_1^1 which is in normal form, according to the Definition 4.3. The remainder $\mathcal{R}^{(2)}(U)$ is Hamiltonian and it is in the class $\mathcal{S}(N)$. Moreover, for any $s \geq \rho$, one has

$$\|\Phi^{(2)}(U)^{\pm 1}\|_{\mathcal{L}(H^s)} \leq 1 + C(s)\|u\|_\rho .$$

In order to prove the proposition stated above, we need some further symbolic calculus. First of all, we consider an even smooth cut-off function $\chi : \mathbb{R} \to [0,1]$ with the property that $\chi(y) = 1$ for all y with $|y| \leq \frac{1}{2}$ and $\chi(y) = 0$ for all y with $|y| \geq 1$.

Definition 4.5. Given ϵ, δ, τ as in (4.10), define the following functions:

$$\chi_k(\xi) = \chi\left(\frac{2|k|^\tau(\xi;k)}{\langle\xi\rangle^\delta}\right), \qquad k \in \mathbb{Z}^d \setminus \{0\},$$
$$\tilde{\chi}_k(\xi) = \chi\left(\frac{|k|}{\langle\xi\rangle^\varepsilon}\right), \qquad k \in \mathbb{Z}^d \setminus \{0\} .$$

Correspondingly, given a symbol $a \in \mathcal{N}_s^m$, we decompose it as follows:

$$a = \langle a \rangle + a^{(\mathrm{nr})} + a^{(\mathrm{res})} + a^{(S)} ,$$

where $\langle a \rangle$ is the x-average of the symbol of a, namely

$$\langle a \rangle(\xi) = \frac{1}{\mu(\mathbb{T}^d)}\int_{\mathbb{T}^d} a(x,\xi)\, dx ,$$

and

$$\begin{aligned} a^{(\mathrm{res})}(x,\xi) &= \sum_{k\neq 0} \chi_k(\xi)\tilde{\chi}_k(\xi)\widehat{a}(k,\xi)e^{\mathrm{i}k\cdot x} , \\ a^{(\mathrm{nr})}(x,\xi) &= \sum_{k\neq 0} (1-\chi_k(\xi))\,\tilde{\chi}_k(\xi)\widehat{a}(k,\xi)e^{\mathrm{i}k\cdot x} , \\ a^{(S)}(x,\xi) &= \sum_{k\neq 0} (1-\tilde{\chi}_k(\xi))\widehat{a}(k,\xi)e^{\mathrm{i}k\cdot x} . \end{aligned} \tag{4.12}$$

We also define

$$g_a(x,\xi) := -\sum_{k\neq 0} \frac{1}{2(\xi;k)}\,(1-\chi_k(\xi))\,\tilde{\chi}_k(\xi)\widehat{a}(k,\xi)e^{\mathrm{i}k\cdot x}\,. \tag{4.13}$$

In [5], Lemma 5.4, we provided suitable bounds for the cut-off functions defined above.

Lemma 4.6. *For any multi-index $\alpha \in \mathbb{N}^d$, one has*

$$\begin{aligned} &|\chi_k(\xi)| \le 1\,, \quad |\partial_\xi^\alpha \chi_k(\xi)| \lesssim_\alpha |k|^{(\tau+1)|\alpha|}\langle\xi\rangle^{-\delta|\alpha|}\,,\\ &|\tilde{\chi}_k(\xi)| \le 1\,, \quad |\partial_\xi^\alpha \tilde{\chi}_k(\xi)| \lesssim_\alpha |k|^{|\alpha|}\langle\xi\rangle^{-(\varepsilon+|\alpha|)}\,, \end{aligned} \tag{4.14}$$

and

$$d_k(\xi) := \frac{1}{2(\xi;k)}(1-\chi_k(\xi))\,, \quad |\partial_\xi^\alpha d_k(\xi)| \lesssim_\alpha \frac{\langle k\rangle^{(|\alpha|+1)\tau+|\alpha|}}{\langle\xi\rangle^{\delta(|\alpha|+1)}}\,. \tag{4.15}$$

As a consequence, for any $s \ge 0$ and $k \in \mathbb{Z}^d$, one has

$$|\chi_k|_{0,s} \lesssim_s \langle k\rangle^{(\tau+1)s}\,, \quad |\tilde{\chi}_k|_{0,s} \lesssim_s \langle k\rangle^{s}\,, \quad |d_k|_{-\delta,s} \lesssim_s \langle k\rangle^{(s+1)\tau+s}\,. \tag{4.16}$$

We now prove the following Lemma.

Lemma 4.7. *Let $m \in \mathbb{R}$. Then, for any $s \ge 0$, the linear map $\mathcal{N}^m_s \to \mathcal{N}^m_s$, $a \mapsto \langle a\rangle$ is linear and continuous. For any $s \ge 0$, there exists $\sigma_s > s$ large enough such that the maps (see Definition 4.5)*

$$\begin{aligned} &\mathcal{N}^m_{\sigma_s} \to \mathcal{N}^m_s\,, \qquad a \mapsto a^{\mathrm{nr}}\,, \quad a \mapsto a^{\mathrm{res}}\,,\\ &\mathcal{N}^m_{\sigma_s} \to \mathcal{N}^{m-\delta}_s\,, \quad a \mapsto g_a\,, \end{aligned}$$

are linear and continuous. Let $N \in \mathbb{N}$. Then there exists $\rho = \rho_N > 0$ large enough, such that, for any $s \ge \rho$, the map

$$\mathcal{N}^m_\rho \to \mathcal{B}(H^s, H^{s+N})\,, \quad a \mapsto \mathcal{R}^{(S)}(a) := \mathrm{Op}^{bw}(a^{(S)})\,,$$

is linear and continuous.

Proof. Since $\langle a\rangle$ is only the space average of the symbol a it is straightforward that $|\langle a\rangle|_{m,s} \lesssim |a|_{m,s}$. We now estimate g_a in terms of a. The estimates for a^{nr} and a^{res} can be done arguing similarly. By the definitions (4.13), (4.15), one has that g_a can be written as

$$g_a(x,\xi) = -\sum_{k\neq 0} d_k(\xi)\tilde{\chi}_k(\xi)\widehat{a}(k,\xi)e^{\mathrm{i}k\cdot x}\,.$$

By applying Lemma 2.1 one gets, for any $s \ge 0$, $N \in \mathbb{N}$, for any $k \in \mathbb{Z}^d$, that

$$|\widehat{a}(k,\cdot)|_{m,s} \lesssim_N \langle k\rangle^{-N} |a|_{m,s+N}\,. \tag{4.17}$$

Fix

$$N := (s+1)\tau + 2s + d + 1\,.$$

One has that

$$\begin{aligned}|d_k\tilde{\chi}_k\widehat{a}(k,\cdot)|_{m-\delta,s} &\lesssim_s |d_k|_{-\delta,s}|\tilde{\chi}_k|_{0,s}|\widehat{a}(k,\cdot)|_{m,s} \overset{(4.16),(4.17)}{\lesssim_s} \langle k\rangle^{(s+1)\tau+2s-N}|a|_{m,s+N} \\ &\lesssim_s \langle k\rangle^{-d-1}|a|_{m,s+N}\,.\end{aligned}$$

Using that $\sum_k \langle k\rangle^{-d-1}$ is convergent, one then gets that $|g_a|_{m-\delta,s} \lesssim_s |a|_{m,s+N}$ and the claimed statement follows.

We now estimate the symbol $a^{(S)}$. By the definition of the cut off function $\tilde{\chi}_k$ in Definition 4.5, one has that

$$\mathrm{supp}(1-\tilde{\chi}_k) \subseteq \big\{\xi : \langle\xi\rangle^{\varepsilon} \le 2|k|\big\}\,,$$

hence, by the estimate (4.14), one has that, for any $N \in \mathbb{N}$, $\alpha \in \mathbb{N}^d$,

$$\langle\xi\rangle^{N+m}|1-\tilde{\chi}_k(\xi)| \lesssim \langle k\rangle^{(N+m)/\varepsilon}\,, \quad \langle\xi\rangle^{N+m+|\alpha|}|\partial_\xi^\alpha(1-\tilde{\chi}_k(\xi))| \lesssim_\alpha \langle k\rangle^{|\alpha|+\frac{N+m+|\alpha|}{\varepsilon}}\,,$$

implying that

$$|1-\tilde{\chi}|_{-N-m,s} \lesssim_s \langle k\rangle^{s+\frac{N+m+s}{\varepsilon}}\,. \tag{4.18}$$

Now fix

$$M := d+1+s+(N+m+s)\varepsilon^{-1}\,.$$

By Lemma 2.1, one has

$$\begin{aligned}|(1-\tilde{\chi}_k)\widehat{a}(k,\cdot)|_{-N,s} &\lesssim |1-\tilde{\chi}_k|_{-N-m,s}|\widehat{a}(k,\cdot)|_{m,s} \overset{(4.18)}{\lesssim_s} \langle k\rangle^{s+\frac{N+m+s}{\varepsilon}-M}|a|_{m,s+M} \\ &\lesssim_s \langle k\rangle^{-d-1}|a|_{m,s+M}\,.\end{aligned}$$

Hence, using that $\sum_k \langle k\rangle^{-d-1}$ is convergent, one gets that

$$|a^{(S)}|_{-N,s} \lesssim_s |a|_{m,s+M}\,. \tag{4.19}$$

We now consider the operator $\mathcal{R}^{(S)}(a) := \mathrm{Op}^{bw}(a^{(S)})$. Its action is given by

$$\mathcal{R}^{(S)}(a)[w] = \sum_{k,\xi\in\mathbb{Z}^d} \eta_\varepsilon\Big(\frac{|k-\xi|}{\langle k+\xi\rangle}\Big)\widehat{a}^{(S)}\Big(k-\xi,\frac{k+\xi}{2}\Big)\widehat{u}(\xi)e^{\mathrm{i}k\cdot x}$$
$$\overset{(4.12)}{=} \sum_{k,\xi\in\mathbb{Z}^d} \eta_\varepsilon\Big(\frac{|k-\xi|}{\langle k+\xi\rangle}\Big)\Big(1-\tilde{\chi}_k\Big(\frac{k+\xi}{2}\Big)\Big)\widehat{a}\Big(k-\xi,\frac{k+\xi}{2}\Big)\widehat{u}(\xi)e^{\mathrm{i}k\cdot x}\,.$$

Clearly the map $a \mapsto \mathcal{R}^{(S)}(a)$ is linear. By defining

$$\rho := s_0 + M = s_0 + d + 1 + s_0 + \frac{N+m+s_0}{\varepsilon}, \quad s_0 := \frac{d}{2}+1$$

the estimate (4.19) reads $|a^{(S)}|_{-N,s_0} \lesssim_N |a|_{m,\rho}$ and therefore, by applying Lemma 2.3, one has that, for any $s \geq \rho$,

$$\|\mathcal{R}^{(S)}(a)\|_{\mathcal{L}(H^s,H^{s+N})} \lesssim_s |a^{(S)}|_{-N,s_0} \lesssim_{s,N} |a|_{m,\rho}\,.$$

Hence the linear map

$$\mathcal{N}^m_\rho \to \mathcal{L}(H^s, H^{s+N}), \quad a \mapsto \mathcal{R}^{(S)}(a)$$

is bounded. The claimed statement has then been proved. □

Lemma 4.8. *Let $a \in \Sigma^m_1$. Then $\langle a\rangle, a^{(\mathrm{nr})}, a^{(\mathrm{res})} \in \Sigma^m_1$ and $g_a \in \Sigma^{m-\delta}_1$. Moreover, for any $N \in \mathbb{N}$, the remainder $\mathcal{R}^{(S)}(U) := \mathrm{Op}^{bw}\Big(a^{(S)}(U;x,\xi)\Big)$ belongs to the class $\mathcal{S}(N)$.*

Proof. Let $a \in \Sigma^m_1$. We show that $g_a \equiv g_a \in \Sigma^{m-\delta}_1$. The proof that $a, a^{(\mathrm{nr})}, a^{(\mathrm{res})} \in \Sigma^m_1$ is analogous. Since $a \in \Sigma^m_1$, then

$$a = a_l + a_q, \quad \text{with} \quad a_l \in O^m_1 \quad \text{and} \quad a_q \in \Gamma^m_2 \tag{4.20}$$

and according to the definitions (4.13), (4.15), one obtains a corresponding splitting $g = g_l + g_q$ where

$$g_{a_l}(x,\xi) = -\sum_{k\neq 0} d_k(\xi)\tilde{\chi}_k(\xi)\widehat{a_l}(U;x,\xi)e^{\mathrm{i}k\cdot x}\,,$$
$$g_{a_q}(x,\xi) = -\sum_{k\neq 0} d_k(\xi)\tilde{\chi}_k(\xi)\widehat{a}_q(U;x,\xi)e^{\mathrm{i}k\cdot x}\,.$$

We show that $g_{a_l} \in O^m_1$ and $g_{a_q} \in \Gamma^m_2$.

Since $a_l \in O^m_1$, then

$$\widehat{a}(U;k,\xi) = m_+(k,\xi)\widehat{u}(k) + m_-(k,\xi)\overline{\widehat{u}(-k)}$$

for some suitable multipliers m_+, m_- and therefore g_l is a symbol which is linear in U and of the same form as a_l. It remains only to show that g_l is in Γ^m_1 and $g_q \in \Gamma^m_2$. Fix $s \geq 0$ and let $\sigma_s > s$

the constant provided by Lemma 4.7. Since $a_l \in \Gamma_1^m$, $a_q \in \Gamma_2^m$ there exists a constant $\sigma_s' > \sigma_s$ and a radius $r = r(s) \in (0,1)$ such that the linear map

$$H^{\sigma_s'} \to \mathcal{N}^m_{\sigma_s}\,, \quad U \mapsto a_l(U;x,\xi)$$

is continuous and the map

$$B_{\sigma_s'}(r) \to \mathcal{N}^m_{\sigma_s}\,, \quad U \mapsto a_q(U;x,\xi)$$

is $\mathcal{C}^\infty$ and vanishes of order two at $U=0$. Since by Lemma 4.7, the map $\mathcal{N}^m_{\sigma_s} \to \mathcal{N}^m_s$, $a \mapsto g_a$ is linear and continuous, then by composition one gets that the map

$$B_{\sigma_s'}(r) \to \mathcal{N}^m_{s}\,, \quad U \mapsto g_{a_l}(U;x,\xi)$$

is linear and continuous and the map

$$B_{\sigma_s'}(r) \to \mathcal{N}^m_{s}\,, \quad U \mapsto g_{a_q}(U;x,\xi)$$

is $\mathcal{C}^\infty$ and vanishes of order two at $U=0$. This shows the claimed statement.

Analysis of the operator $\mathcal{R}^{(S)}(U) = \mathrm{Op}^{bw}(a^{(S)}(U;x,\xi))$. According to (4.20)

$$\begin{aligned} a^{(S)} &= a_l^{(S)} + a_q^{(S)}\,, \\ a_l^{(S)} &= \sum_{k\neq 0} (1-\tilde{\chi}_k(\xi))\,\widehat{a}_l(k,\xi) e^{\mathrm{i}k\cdot x} \\ &= \sum_{k\neq 0} (1-\tilde{\chi}_k(\xi))\, m_+(k,\xi)\widehat{u}(k) e^{\mathrm{i}k\cdot x} + \sum_{k\neq 0} (1-\tilde{\chi}_k(\xi))\, m_-(k,\xi)\overline{\widehat{u}(-k)} e^{\mathrm{i}k\cdot x}\,, \\ a_q^{(S)} &= \sum_{k\neq 0} (1-\tilde{\chi}_k(\xi))\,\widehat{a}_q(k,\xi) e^{\mathrm{i}k\cdot x}\,, \end{aligned} \tag{4.21}$$

and correspondingly

$$\mathcal{R}^{(S)}(U) = \mathcal{R}^{(S)}_l(U) + \mathcal{R}^{(S)}_q(U)\,, \qquad \mathcal{R}^{(S)}_l(U) := \mathrm{Op}^{bw}(a_l^{(S)})\,, \quad \mathcal{R}^{(S)}_q(U) := \mathrm{Op}^{bw}(a_q^{(S)})\,.$$

Fix $N \in \mathbb{N}$ and let $\rho \equiv \rho_N$ be the constant appearing in Lemma 4.7. Since $a_l \in O_1^m$ and $a_q \in \Gamma_2^m$, one has that for some constant $\sigma_\rho > \rho$ large enough

$$|a_l|_{m,\rho} \lesssim_\rho \|U\|_{\sigma_\rho}\,, \quad |a_q|_{m,\rho} \lesssim_\rho \|U\|^2_{\sigma_\rho}\,,$$

implying that, for any $s \geq \sigma_\rho > \rho$, one has

$$\begin{aligned} \|\mathcal{R}^{(S)}_l(U)\|_{\mathcal{L}(H^s,H^{s+N})} &\lesssim_s |a_l|_{m,\rho} \lesssim_s \|U\|_{\sigma_\rho}\,, \\ \|\mathcal{R}^{(S)}_q(U)\|_{\mathcal{L}(H^s,H^{s+N})} &\lesssim_s |a_q|_{m,\rho} \lesssim_s \|U\|^2_{\sigma_\rho}\,, \end{aligned}$$

and hence $\mathcal{R}_q^{(S)} \in \mathcal{S}_2(N)$. In order to show that $\mathcal{R}_l^{(S)}(U)$ belongs to the class $\mathcal{OS}_1(N)$ it remains only to show that it is sum of terms of the form (2.6). This follows since

$$\begin{aligned}\mathcal{R}_l^{(S)}(U)[w] &= \sum_{k,\xi\in\mathbb{Z}^d} \eta_\varepsilon\Big(\frac{|k-\xi|}{\langle k+\xi\rangle}\Big)\Big(1-\tilde{\chi}_k\Big(\frac{k+\xi}{2}\Big)\Big)\widehat{a_l}\Big(k-\xi,\frac{k+\xi}{2}\Big)\widehat{w}(\xi)e^{\mathrm{i}k\cdot x}\\ &\overset{(4.21)}{=} \sum_{k,\xi\in\mathbb{Z}^d} r_+(k,\xi)\widehat{u}(k-\xi)\widehat{w}(\xi)e^{\mathrm{i}k\cdot x} + \sum_{k,\xi\in\mathbb{Z}^d} r_-(k,\xi)\overline{\widehat{u}(\xi-k))}\widehat{w}(\xi)e^{\mathrm{i}k\cdot x}\end{aligned}$$

where

$$r_\pm(k,\xi) := \eta_\varepsilon\Big(\frac{|k-\xi|}{\langle k+\xi\rangle}\Big)\Big(1-\tilde{\chi}_k\Big(\frac{k+\xi}{2}\Big)\Big)m_\pm\Big(k-\xi,\frac{\xi+k}{2}\Big).$$

The claimed statement has then been proved. □

We are now in position to prove the Proposition 4.4.

Proof of Proposition 4.4. The Proposition is proved also inductively, hence we describe the induction step of the procedure. In the proof it is convenient to use the following notations. If $\mathcal{O}$ is one of the classes of operators defined in Section 2.1, we write $A = B + \mathcal{O}$ if $A - B$ belongs to the class $\mathcal{O}$.

We define the *gain of regularization* along the reduction procedure as

$$\mathfrak{e} := \min\big\{\delta, 3\delta-2, 2\delta-1\big\}. \tag{4.22}$$

At the n-th step, we deal with a Hamiltonian para-differential operator of the form

$$\mathcal{P}_n^{(1)}(U) := \partial_t + \mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi)) + \mathcal{Z}_n(U) + \mathcal{A}_n(U) + \mathcal{R}_n(U) \tag{4.23}$$

where

$$\begin{aligned}&\mathcal{Z}_n(U) := \mathrm{i}Op^{bw}\begin{pmatrix} z_n(U;x,\xi) & 0\\ 0 & -z_n(U;x,-\xi)\end{pmatrix}, \quad z_n\in\Sigma_1^1, \quad z_n \text{ is real and in normal form},\\ &\mathcal{A}_n(U) := \mathrm{i}Op^{bw}\begin{pmatrix} a_n(U;x,\xi) & 0\\ 0 & -a_n(U;x,-\xi)\end{pmatrix}, \quad a_n\in\Sigma_1^{1-n\mathfrak{e}}, \quad a_n \text{ is real},\\ &\mathcal{R}_n \in \mathcal{S}(N).\end{aligned} \tag{4.24}$$

By Lemma 4.8, one has that

$$\begin{aligned}&\mathrm{Op}^{bw}(a_n) = \mathrm{Op}^{bw}(\langle a_n\rangle + a_n^{(\mathrm{nr})} + a_n^{(\mathrm{res})}) + \mathcal{R}^{(S)}(a_n),\\ &\langle a_n\rangle, a_n^{(\mathrm{nr})}, a_n^{(\mathrm{res})} \in \Sigma_1^{1-n\mathfrak{e}}, \quad \mathcal{R}^{(S)}(a_n)\in\mathcal{S}(N).\end{aligned}$$

Moreover, by defining (as in (4.13))

$$g_n(U;x,\xi) := -\sum_{k\neq 0} \frac{1}{2(\xi;k)} (1-\chi_k(\xi))\, \tilde{\chi}_k(\xi)\widehat{a}_n(U;k,\xi)e^{\mathrm{i}k\cdot x}\,,$$

one has that the symbol $g_n(U;x,\xi)$ is in $\Sigma_1^{1-n\mathfrak{e}-\delta}$ and solves the equation

$$\{\Lambda, g_n\} + a_n^{(\mathrm{nr})} = 0\,. \tag{4.25}$$

We then consider the map

$$\mathbf{\Phi}_n(U) := \begin{pmatrix} \Phi_n(U) & 0 \\ 0 & \overline{\Phi_n(U)} \end{pmatrix}$$

where $\Phi_n(U)$ is the time one flow map of

$$\partial_\tau \Phi_n(U) = \mathrm{i}\mathrm{Op}^{bw}(g_n)\Phi_n^\tau(U)\,, \quad \Phi_n^0(U) = \mathrm{Id}\,.$$

The map $\mathbf{\Phi}_n(U)$ is well-posed and symplectic by Lemma 3.1. We now compute the conjugated operator

$$\mathcal{P}_{n+1}^{(1)}(U) := \mathbf{\Phi}_n(U)^{-1}\mathcal{P}_n^{(1)}(U)\mathbf{\Phi}_n(U)\,.$$

Note that for any $n \geq 0$, $1-n\mathfrak{e}-\delta < \delta < 1$, hence the conjugation Lemmas of Section 2.2 can be applied. In particular, by applying Lemmata 3.2, 3.4 (where n is replaced by $1-n\mathfrak{e}-\delta$), one gets

$$\begin{aligned}
\Phi_n(U)^{-1}\partial_t\Phi_n(U) &= \partial_t + \mathcal{OB}_\Sigma(1-n\mathfrak{e}-\delta) + \mathcal{S}(N)\,,\\
\Phi_n(U)^{-1}\mathrm{i}\mathrm{Op}^{bw}(\Lambda)\Phi_n(U) &= \mathrm{i}\mathrm{Op}^{bw}(\Lambda + \{\Lambda, g_n\}) + \mathcal{OB}_\Sigma(3-2n\mathfrak{e}-3\delta) + \mathcal{S}(N)\,,\\
\Phi_n(U)^{-1}\mathrm{i}\mathrm{Op}^{bw}(z_n)\Phi_n(U) &= \mathrm{i}\mathrm{Op}^{bw}(z_n) + \mathcal{OB}_\Sigma(2-n\mathfrak{e}-2\delta) + \mathcal{S}(N)\,,\\
\Phi_n(U)^{-1}\mathrm{i}\mathrm{Op}^{bw}(a_n)\Phi_n(U) &= \mathrm{i}\mathrm{Op}^{bw}(\langle a_n\rangle + a_n^{(\mathrm{nr})} + a_n^{(\mathrm{res})}) + \mathcal{OB}_\Sigma(2-2n\mathfrak{e}-2\delta) + \mathcal{S}(N)\,,\\
\mathbf{\Phi}_n(U)^{-1}\mathcal{R}_n(U)\mathbf{\Phi}_n(U) &= \mathcal{S}(N)\,.
\end{aligned}$$

By the definition of $\mathfrak{e}$ given in (4.22), one obtains that

$$1-n\mathfrak{e}-\delta\,, 3-2n\mathfrak{e}-3\delta\,, 2-n\mathfrak{e}-2\delta\,, 2-2n\mathfrak{e}-2\delta \leq 1-(n+1)\mathfrak{e}$$

and using that g_n solves the equation (4.25), one obtains that $\mathcal{P}_{n+1}^{(1)}(U)$ has the form (4.23) with $n \rightsquigarrow n+1$, for some $\mathcal{R}_{n+1}(U) \in \mathcal{S}(N)$ and

$$\begin{aligned}
&\mathcal{Z}_{n+1}(U) := \mathrm{i}Op^{bw}\begin{pmatrix} z_{n+1}(U;x,\xi) & 0 \\ 0 & -z_{n+1}(U;x,-\xi)\end{pmatrix}\,,\\
&\quad z_{n+1} \in \Sigma_1^1\,, \quad z_{n+1} := z_n + \langle a_n\rangle + a_n^{(\mathrm{res})}\,,\\
&\mathcal{A}_{n+1}(U) := \mathrm{i}Op^{bw}\begin{pmatrix} a_{n+1}(U;x,\xi) & 0 \\ 0 & -a_{n+1}(U;x,-\xi)\end{pmatrix}\,, \quad a_{n+1} \in \Sigma_1^{1-(n+1)\mathfrak{e}}\,.
\end{aligned}$$

Since $\Phi_n(U)$ is a linear symplectic map, the paradifferential operator $\mathcal{P}_{n+1}(U)$ is Hamiltonian, hence z_{n+1} and a_{n+1} are real symbols. Furthermore, z_{n+1} is a symbol in normal form, since z_n is in normal form by the induction hypothesis and $\langle a_n \rangle, a_n^{(\mathrm{res})}$ are in normal form by their definition. The claimed induction statement has then been proved. □

5. The Birkhoff normal form step

By Propositions 4.2, 4.4, one has that U solves the equation (4.2) if and only if $U := \Phi^{(1)}(U)\Phi^{(2)}(U)W$ solves

$$\mathcal{P}^{(3)}(U)[W] = 0\,, \quad \mathcal{P}^{(3)}(U) := \partial_t + \mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi))W + \mathcal{Z}(U) + \mathcal{Q}(U)$$

where $\mathcal{Q} \in \mathcal{S}(N)$ and $\mathcal{Z}(U)$ is the normal form operator provided in Proposition 4.4. We now perform a step of Birkhoff normal in order to remove the quadratic terms from $\mathcal{Q}(U)W$. Since $\mathcal{Q} \in \mathcal{S}(N)$, then

$$\mathcal{Q} = \mathcal{Q}_l + \mathcal{Q}_q\,, \quad \mathcal{Q}_l \in \mathcal{OS}_1(N)\,, \quad \mathcal{Q}_q \in \mathcal{S}_2(N)\,. \tag{5.1}$$

We fix the number of regularization step N as

$$N := \tau + 3\,, \tag{5.2}$$

where τ is the loss of derivatives in the small divisors estimate of Lemma A.1. We prove the following.

Proposition 5.1. *Let $\mathcal{G} \in (0, +\infty)$ be the full Lebesgue measure set given by Lemma A.1. Then for any $m \in \mathcal{G}$ the following holds. Then there exists $\rho \equiv \rho(\tau) \gg 0$ large enough such that if (4.1) is fulfilled, then the following holds. There exists a linear and invertible transformation $\Phi^{(3)}(U) : \mathbf{H}^s \to \mathbf{H}^s$ such that*

$$\mathcal{P}^{(4)}(U) := \Phi^{(3)}(U)^{-1}\mathcal{P}^{(3)}(U)\Phi^{(3)}(U) = \partial_t + \mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi)) + \mathcal{Z}(U) + \mathcal{R}^{(4)}(U) \tag{5.3}$$

where $\mathcal{Z}(U)$ is given in Proposition 4.4 and $\mathcal{R}^{(4)}(U)W$ is cubic and one-smoothing remainder, namely it satisfies, for any $s \geq \rho$, $W \in H^s$, the estimate

$$\|\mathcal{R}^{(4)}(U)W\|_{s+1} \lesssim_s \|U\|_\rho^2 \|W\|_s\,. \tag{5.4}$$

Moreover, for any $s \geq \rho$, one has

$$\|\Phi^{(3)}(U)^{\pm 1}\|_{\mathcal{L}(H^s)} \leq 1 + C(s)\|u\|_\rho\,.$$

Proof. We look for a smoothing operator $\mathcal{F} \in \mathcal{OS}_1(3)$ and we consider the flow map $\Phi^\tau_{\mathcal{F}}(U)$. We then set $\Phi^{(3)}(U) := \Phi^1_{\mathcal{F}}(U)$. By applying Lemma 3.7, one gets that

$$\begin{aligned}
\Phi^{(3)}(U)^{-1}\circ\partial_t\circ\Phi^{(3)}(U) &= \partial_t - \mathcal{F}(\mathrm{i}E\mathrm{Op}^{bw}(\Lambda)U) + \mathcal{S}_2(3)\\
\Phi^{(3)}(U)^{-1}\mathrm{i}E\mathrm{Op}^{bw}(\Lambda)\Phi^{(3)}(U) &= \mathrm{i}E\mathrm{Op}^{bw}(\Lambda) + [\mathrm{i}E\mathrm{Op}^{bw}(\Lambda), \mathcal{F}(U)] + \mathcal{S}_2(1)\\
\Phi^{(3)}(U)^{-1}\mathcal{Z}(U)\Phi^{(3)}(U) &= \mathcal{Z}(U) + \mathcal{S}_2(2)\\
\Phi^{(3)}(U)^{-1}\mathcal{Q}(U)\Phi^{(3)}(U) &\overset{(5.1),(5.2)}{=} \mathcal{Q}_l(U) + \mathcal{S}_2(\tau+3)\,.
\end{aligned}\tag{5.5}$$

By applying Lemma 3.9, since $\mathcal{Q}_l \in \mathcal{OS}_1(\tau+3)$, then there exists $\mathcal{F} \in \mathcal{OS}_1(3)$ which solves

$$-\mathcal{F}(\mathrm{i}E\mathrm{Op}^{bw}(\Lambda)U) + [\mathrm{i}E\mathrm{Op}^{bw}(\Lambda), \mathcal{F}(U)] + \mathcal{Q}_l(U) = 0\,. \tag{5.6}$$

Then (5.5), (5.6) imply that $\mathcal{P}^{(4)}(U) := \Phi^{(3)}(U)^{-1}\mathcal{P}^{(3)}(U)\Phi^{(3)}(U)$ has the form (5.3), with $\mathcal{R}^{(4)} \in \mathcal{S}_2(1)$ and hence satisfying the claimed estimate (5.4). □

6. Energy estimates and proof of Theorem 1.1 concluded

In this section we conclude the proof of the main result of the paper, namely Theorem 1.1. The main point is to provide an energy estimate for the reduced equation.

$$\partial_t W + \mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi))W + \mathcal{Z}(U)W + \mathcal{R}^{(4)}(U)W = 0\,. \tag{6.1}$$

First of all, let us consider the linear flow associated to the normal form equation

$$\partial_t W + \mathrm{i}E\mathrm{Op}^{bw}(\Lambda(\xi))W + \mathcal{Z}(U)W = 0\,,$$

which, by (4.11), is equivalent to the scalar equation

$$\partial_t w + \mathrm{i}\mathrm{Op}^{bw}\Big(\Lambda(\xi) + z(U;x,\xi)\Big)w = 0\,, \tag{6.2}$$

where z is a real symbol in normal form (see Definition 4.3). The following Lemma is proved in [6], Section 5.1.

Lemma 6.1. *For any $t,\tau \in [-T,T]$ (where T is the same as in (4.1)), the flow $\mathcal{U}_z(\tau,t)$ associated to the equation (6.2) (with $\mathcal{U}_z(\tau,\tau) = \mathrm{Id}$) is well defined as a bounded linear operator $H^s \to H^s$ and it satisfies*

$$\|\mathcal{U}_z(\tau,t)w_0\|_s \lesssim_s \|w_0\|_s\,, \quad \textit{uniformly w.r. to} \quad t,\tau \in [-T,T]\,, \tag{6.3}$$

for any $w_0 \in H^s$.

Proof. The proof is exactly the same as the one made in [6], Section 5.1. Indeed the operator $\mathrm{Op}^{bw}(z)$ is the Weil quantization of the *truncated* symbol

$$\sigma_z(U;x,\xi) = \sum_{k\in\mathbb{Z}^d} \eta_\epsilon\Big(\frac{|k|}{\langle\xi\rangle}\Big)\widehat{z}(U;k,\xi)e^{\mathrm{i}k\cdot x}$$

see (2.3). Since $\widehat{\sigma_z}(U;k,\xi)=\eta_\varepsilon\Big(\frac{|k|}{\langle\xi\rangle}\Big)\widehat{z}(U;k,\xi)$, one has that also σ_z is a symbol in normal form according to the definition (4.3). Hence the arguments developed in [6] apply. □

We then denote by

$$\mathcal{U}_{\mathcal{Z}}(\tau,t):=\begin{pmatrix}\mathcal{U}_z(\tau,t) & 0\\ 0 & \overline{\mathcal{U}_z(\tau,t)}\end{pmatrix}.$$

By Duhamel formula, solutions of (6.1) satisfy

$$W(t)=\mathcal{U}_{\mathcal{Z}}(0,t)w_0-\int_0^t \mathcal{U}_{\mathcal{Z}}(\tau,t)\mathcal{R}(U(\tau))W(\tau)\,d\tau\,, \tag{6.4}$$

and recall that, by the ansatz on $U(t)$, we have $\|U(t)\|_\rho\lesssim\varepsilon$, $\forall t\in[-T,T]$, for $\rho\gg 0$ large enough and some $T>0$. By the estimates (5.4), (6.3), one then has if $w_0\in H^\rho$,

$$\|W(t)\|_\rho\lesssim_\rho\|W_0\|_\rho+\int_0^t\|U(\tau)\|_\rho^2\|W(\tau)\|_\rho\,d\tau\lesssim_\rho\|W_0\|_\rho+\varepsilon^2\int_0^t\|W(\tau)\|_\rho\,d\tau\,.$$

By Gronwall inequality, one then gets that

$$\|W(t)\|_\rho\le C(\rho)e^{C(\rho)\varepsilon^2 t}\|W_0\|_\rho\,,\quad \forall t\in[-T,T]\,,$$

for some constant $C(\rho)\gg 0$ large enough. By Propositions 4.2, 4.4, 5.1 and

$$U(t)=\Big(\Phi^{(1)}(U(t))\circ\Phi^{(2)}(U(t))\circ\Phi^{(3)}(U(t))\Big)[W(t)]\,,$$

one deduces that

$$\|U(t)\|_\rho\sim_\rho\|W(t)\|_\rho\,,\quad\forall t\in[-T,T]$$

and therefore,

$$\|U(t)\|_\rho\le C_1(\rho)e^{C(\rho)\varepsilon^2 t}\|U_0\|_\rho\,,\quad\forall t\in[-T,T]\,,$$

for some constant $C_1(\rho)\gg 0$ large enough. By a standard bootstrap argument, the latter estimate implies that $T=T_\rho=O(\varepsilon^{-2})$ and

$$\|U(t)\|_\rho\lesssim_\rho\|U_0\|_\rho\,,\quad\forall t\in[-T_\rho,T_\rho]\,.$$

Clearly, by the smallness assumption on the initial datum $\|U_0\|_\rho\le\varepsilon$, one then gets that $\|U(t)\|_\rho\lesssim_\rho\varepsilon$, for any $t\in[-T_\rho,T_\rho]$. This is the estimate (1.6) in Theorem 1.1.

We now perform a boothstrap argument in order to show that if $s\ge\rho$ and $U_0\in\mathbf{H}^s$ (see (2.4)) then

$$U \in C^0\Big([-T_\rho, T_\rho], H^s\Big) \quad \text{with} \quad \|U(t)\|_s \le C_*(s)\|U_0\|_s\,, \quad \forall t \in [-T_\rho, T_\rho]\,,$$
$$\text{for some} \quad C_*(s) \gg 0 \quad \text{large enough}\,. \tag{6.5}$$

The latter claim implies the estimate (1.7). In order to prove the (6.5) we argue by induction on $s \ge \rho$. If $s = \rho$, then the claimed statement is proved. Assume that the statement is true for some $s > \rho$ and let us prove it for $s + 1$. Let $U_0 \in \mathbf{H}^{s+1}$. Then $W_0 \in \mathbf{H}^{s+1}$ and by the induction hypothesis $U(t)$ and then $W(t)$ is in $C^0\Big([-T_\rho, T_\rho], \mathbf{H}^s\Big)$. By applying Lemma 6.1, using that the remainder $\mathcal{R}(U)$ in (6.4) is one smoothing (see (5.4)), one has that

$$\begin{aligned}
&\mathcal{U}_{\mathcal{Z}}(\tau,t)W_0 \in \mathbf{H}^{s+1}\,, \quad \|\mathcal{U}_{\mathcal{Z}}(\tau,t)W_0\|_{s+1} \lesssim_s \|W_0\|_{s+1}\,, \quad \forall \tau, t \in \mathbb{R}\,,\\
&W(\tau) \in \mathbf{H}^s \implies \mathcal{R}(U(\tau))[W(\tau)] \in \mathbf{H}^{s+1} \implies \mathcal{U}_{\mathcal{Z}}(\tau,t)\mathcal{R}(U(\tau))[W(\tau)] \in \mathbf{H}^{s+1}\,,\\
&\Big\| \int_0^t \mathcal{U}_{\mathcal{Z}}(\tau,t)\mathcal{R}(U(\tau))W(\tau)\,d\tau \Big\|_{s+1} \lesssim_s \varepsilon^2 \int_0^t \|W(\tau)\|_s\,d\tau\,.
\end{aligned} \tag{6.6}$$

Therefore, (6.4), (6.6) imply that

$$W(t) \in \mathbf{H}^{s+1}\,, \quad \|W(t)\|_{s+1} \lesssim_s \|W_0\|_{s+1} + \varepsilon^2 \int_0^t \|W(\tau)\|_s\,d\tau\,, \quad \forall t \in [-T_\rho, T_\rho]\,. \tag{6.7}$$

Using that (by the boundedness of the normal form transformations)

$$\|U(t)\|_{s+1} \sim_s \|W(t)\|_{s+1}, \quad \|U(t)\|_s \sim_s \|W(t)\|_s\,, \quad \|W_0\|_{s+1} \sim_s \|U_0\|_{s+1}\,, \quad \|W_0\|_s \sim_s \|U_0\|_s\,,$$

we note that (6.7) implies

$$U(t) \in \mathbf{H}^{s+1}\,, \quad \|U(t)\|_{s+1} \le K(s)\Big(\|U_0\|_{s+1} + \varepsilon^2 \int_0^t \|U(\tau)\|_s\,d\tau\Big)\,, \quad \forall t \in [-T_\rho, T_\rho]\,,$$

for some constant $K(s) \gg 0$ large enough. Hence, by the induction hypothesis (6.5), the latter inequality implies that, for any $t \in [-T_\rho, T_\rho]$,

$$\|U(t)\|_{s+1} \le K(s)\|U_0\|_{s+1} + K(s)C_*(s)T_\rho\varepsilon^2\|U_0\|_s \le C_*(s+1)\|U_0\|_{s+1}$$

with $C_*(s+1) := K(s)(1 + C_*(s))$ and using that $T_\rho\varepsilon^{-2} \le 1$. The claimed statement (6.5) has then been proved for $s + 1$. The proof of Theorem 1.1 is then concluded.

Appendix A. Non-resonance conditions

In this section we verify the non resonance conditions appearing in the Birkhoff normal form. We need to provide suitable lower bounds for the *three wave interactions*

$$\phi^{\sigma,\sigma'}(\xi,k) := \Lambda(\xi+k) + \sigma\Lambda(\xi) + \sigma'\Lambda(k)\,, \quad \xi, k \in \mathbb{Z}^d\,, \quad \sigma, \sigma' \in \{\pm\}\,, \tag{A.1}$$

where $\Lambda(\xi)$ is the symbol defined in (1.5), for “most” choices of the parameter $m \in (0, +\infty)$. This is the content of the following Lemma.

Lemma A.1. *There exists a set $\mathcal{G} \subseteq (0, +\infty)$ of Lebesgue measure 1 such that for any $m \in \mathcal{G}$ there exist $\tau = \tau(d) \geq 0$ and $\gamma > 0$ such that, for all $\xi, k \in \mathbb{Z}^d$, $\sigma, \sigma' \in \{\pm\}$, one has*

$$|\phi^{\sigma,\sigma'}(\xi, k)| \geq \frac{\gamma}{\langle \xi \rangle^{\tau} \langle k \rangle^{\tau}} . \tag{A.2}$$

The rest of the section is devoted to the proof of the lemma above. Let us denote by

$$\omega_g \in \mathbb{R}^{d_*} , \quad d_* := \frac{d(d-1)}{2} + d , \tag{A.3}$$

the vector obtained by putting in a vector the matrix elements (upon the diagonal) of the matrix G in (1.2), namely

$$\omega_g := (g_{11}, \dots, g_{1d}, g_{22}, \dots, g_{2d}, \dots, g_{(d-1)(d-1)}, g_{(d-1)d}, g_{dd}) .$$

Fix $\tau_* \geq d_*$. We define the set

$$\mathcal{G} := \Big\{ m > 0 : \exists \gamma > 0 \quad \text{such that} \quad |\omega_g \cdot \ell \pm m| \geq \frac{\gamma}{\langle \ell \rangle^{\tau_*}} , \quad \forall \ell \in \mathbb{Z}^{d_*} \Big\} . \tag{A.4}$$

The following Lemma holds

Lemma A.2. *The Lebesgue measure of $(0, +\infty) \setminus \mathcal{G}$ is equal to zero.*

Proof. A direct calculation shows that

$$\mathcal{G}^c := (0, +\infty) \setminus \mathcal{G} = \cap_{\gamma > 0} \cup_{\ell \in \mathbb{Z}^{\nu}} \mathcal{R}_{\ell}(\gamma),$$
$$\mathcal{R}_{\ell}(\gamma) := \Big\{ m > 0 : |\omega_g \cdot \ell \pm m| < \frac{\gamma}{\langle \ell \rangle^{\tau_*}} \Big\} .$$

Clearly $|\mathcal{R}_{\ell}(\gamma)| \lesssim \gamma \langle \ell \rangle^{-\tau_*}$, implying that

$$| \cup_{\ell \in \mathbb{Z}^{\nu}} \mathcal{R}_{\ell}(\gamma)| \lesssim \sum_{\ell \in \mathbb{Z}^{d_*}} \gamma \langle \ell \rangle^{-\tau_*} \lesssim \gamma .$$

This implies that $|\mathcal{G}^c| = 0$. □

Proof of Lemma A.1. Clearly, the function $\phi^{\sigma,\sigma'}(\xi, k)$ in (A.1) is very easy to estimate in the case $\sigma = \sigma' = +$. Indeed G is positive definite and therefore

$$|\Lambda(\xi + k) + \Lambda(\xi) + \Lambda(k)| \geq 3m + |\xi + k|^2 + |\xi|^2 + |k|^2$$

which is bounded away from zero.

We now estimate from below $\phi^{\sigma,\sigma'}(\xi, k)$ in the cases $\sigma = -\sigma' = +$ or $\sigma = \sigma' = -$. Let $\xi, k \in \mathbb{Z}^d$. A direct calculation shows that

$$\Lambda(\xi+k)-\Lambda(\xi)-\Lambda(k)=2\big\langle G\xi\,,\,k\big\rangle-m=2\Big(\sum_{i=1}^{d}g_{ii}\xi_i k_i+\sum_{i=1}^{d}\sum_{j=1}^{i-1}g_{ij}(\xi_i k_j+\xi_j k_i)\Big)-m\,.$$

By the diophantine condition (A.4), one then obtains that for some $\gamma\in(0,1)$,

$$\begin{aligned}&|\Lambda(\xi+k)-\Lambda(\xi)-\Lambda(k)|\geq\frac{\gamma}{f(k,\xi)^{\tau_*}}\quad\text{where}\\&f(k,\xi):=1+2\sum_{i=1}^{d}|\xi_i k_i|+2\sum_{1=1}^{d}\sum_{j=1}^{i-1}|\xi_i k_j+\xi_j k_i|\,.\end{aligned}\tag{A.5}$$

We note that

$$|f(k,\xi)|\lesssim 1+\sum_{i=1}^{d}|\xi_i||k_i|+\sum_{i,j=1}^{d}|\xi_i||k_j|\leq c(d)\langle\xi\rangle\langle k\rangle$$

for some constant $c(d)\geq 1$. Hence

$$|\Lambda(\xi+k)-\Lambda(\xi)-\Lambda(k)|\geq\frac{\gamma_1}{\langle\xi\rangle^{\tau_*}\langle k\rangle^{\tau_*}}\quad\text{for some}\quad\gamma_1\ll\gamma\,.$$

Similarly, one computes for any $\xi,k\in\mathbb{Z}^d$,

$$\begin{aligned}\Lambda(\xi+k)+\Lambda(\xi)-\Lambda(k)&=2\|\xi\|_g^2+2\big\langle G\xi,k\big\rangle+m\\&=2\big\langle G\xi\,,\,\xi+k\big\rangle+m\\&=2\Big(\sum_{i=1}^{d}g_{ii}\xi_i(\xi+k)_i+\sum_{i=1}^{d}\sum_{j=1}^{i-1}g_{ij}(\xi_i(\xi+k)_j+\xi_j(\xi+k)_i)\Big)+m\,.\end{aligned}$$

Hence, using again the diophantine condition (A.4) and recalling the definition of f in (A.5), one obtains that

$$|\Lambda(\xi+k)+\Lambda(\xi)-\Lambda(k)|\geq\frac{\gamma}{f(\xi+k,\xi)^{\tau_*}}\,.$$

Moreover

$$\begin{aligned}f(\xi+k,\xi)&\lesssim 1+\sum_{i=1}^{d}|\xi_i||\xi_i+k_i|+\sum_{i,j=1}^{d}|\xi_i||\xi_j+k_j|\\&\lesssim 1+\sum_{i,j=1}^{d}|\xi_i||\xi_j|+\sum_{i,j=1}^{d}|\xi_i||k_j|\leq c(d)(\langle\xi\rangle\langle k\rangle+\langle\xi\rangle^2)\end{aligned}$$

for some constant $c(d)\geq 1$. This implies that

$$|\Lambda(\xi+k)+\Lambda(\xi)-\Lambda(k)| \geq \frac{\gamma_1}{\langle\xi\rangle^{2\tau_*}\langle k\rangle^{\tau_*}},$$

and hence Lemma A.1 follows. □

References

[1] P. Baldi, E. Haus, R. Montalto, Controllability of quasi-linear Hamiltonian NLS equations, J. Differ. Equ. 264 (3) (2018) 1789–1840.
[2] D. Bambusi, Birkhoff normal form for some nonlinear PDEs, Commun. Math. Phys. 234 (2003) 253–283.
[3] D. Bambusi, J.M. Delort, B. Grébert, J. Szeftel, Almost global existence for Hamiltonian semi-linear Klein-Gordon equations with small Cauchy data on Zoll manifolds, Commun. Pure Appl. Math. 60 (2007) 1665–1690.
[4] D. Bambusi, B. Grébert, Birkhoff normal form for partial differential equations with tame modulus, Duke Math. J. 135 (3) (2006) 507–567.
[5] D. Bambusi, B. Langella, R. Montalto, On the spectrum of the Schrödinger operator on $\mathbb{T}^d$: a normal form approach, Commun. Partial Differ. Equ. 45 (4) (2020) 303–320.
[6] D. Bambusi, B. Langella, R. Montalto, Growth of Sobolev norms for unbounded perturbations of the Laplacian on flat tori, preprint, arXiv:2012.02654.
[7] M. Berti, J.-M. Delort, Almost Global Solutions of Capillary-Gravity Water Waves Equations on the Circle, UMI Lecture Notes, ISBN 978-3-319-99486-4, 2018.
[8] M. Berti, A. Maspero, Long time dynamics of Schrödinger and wave equations on flat tori, J. Differ. Equ. 267 (2) (2019) 1167–1200.
[9] J. Bourgain, Growth of Sobolev norms in linear Schrödinger equations with smooth time dependent potentials, J. Anal. Math. 77 (1999) 315–348.
[10] J. Bourgain, Growth of Sobolev norms in linear Schrödinger equations with quasi-periodic potential, Commun. Math. Phys. 204 (1) (1999) 207–247.
[11] J. Bourgain, Global Solutions of Nonlinear Schrödinger Equations, Colloquium Publications, vol. 46, American Mathematical Society, 1999.
[12] J.M. Delort, Growth of Sobolev norms of solutions of linear Schrödinger equations on some compact manifolds, Int. Math. Res. Not. 12 (2010) 2305–2328.
[13] N. de Bouard, A. Hayashi, J.C. Saut, Global existence of small solutions to a relativistic nonlinear Schrödinger equation, Commun. Math. Phys. 189 (1997) 73–105.
[14] J. Bourgain, Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations, Geom. Funct. Anal. 6 (2) (1996) 201–230.
[15] M. Colin, On the local well-posedness of quasilinear Schrödinger equations in arbitrary space dimension, Commun. Partial Differ. Equ. 27 (2002) 325–354.
[16] M. Christ, Illposedness of a Schrödinger equation with derivative nonlinearity, preprint.
[17] J.M. Delort, On long time existence for small solutions of semi-linear Klein-Gordon equations on the torus, J. Anal. Math. 107 (2009) 161–194.
[18] 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).
[19] J.M. Delort, Quasi-Linear Perturbations of Hamiltonian Klein-Gordon Equations on Spheres, American Mathematical Society, 2015.
[20] J.M. Delort, N. Masmoudi, Long time dispersive estimates for perturbations of kink solution of one dimensional cubic wave equation, preprint, hal-02862414v2.
[21] J.M. Delort, J. Szeftel, Long-time existence for small data nonlinear Klein–Gordon equations on tori and spheres, Int. Math. Res. Not. 37 (2004).
[22] D. Fang, Q. Zhang, Long-time existence for semi-linear Klein-Gordon equations on tori, J. Differ. Equ. 249 (2010) 151–179.
[23] E. Faou, B. Grébert, A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus, Anal. PDE 6 (6) (2013).
[24] R. Feola, F. Iandoli, Local well-posedness for quasi-linear NLS with large Cauchy data on the circle, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 36 (1) (2018) 119–164.
[25] R. Feola, F. Iandoli, Local well-posedness for the Hamiltonian quasi-linear Schrödinger equation on tori, preprint, arXiv:2003.04815.

[26] R. Feola, F. Iandoli, Long time existence for fully nonlinear NLS with small Cauchy data on the circle, Ann. Sc. Norm. Super. Pisa, Cl. Sci. (2019), https://doi.org/10.2422/2036-2145.201811-003, in press.
[27] R. Feola, F. Iandoli, A non-linear Egorov theorem and Poincaré-Birkhoff normal forms for quasi-linear pdes on the circle, preprint, arXiv:2002.12448, 2020.
[28] R. Feola, B. Grébert, F. Iandoli, Long time solutions for quasi-linear Hamiltonian perturbations of Schrödinger and Klein-Gordon equations on tori, preprint, arXiv:2009.07553, 2020.
[29] R. Imekraz, Long time existence for the semi-linear beam equation on irrational tori of dimension two, Nonlinearity 29 (2007) 1–46.
[30] A.D. Ionescu, F. Pusateri, Long-time existence for multi-dimensional periodic water waves, Geom. Funct. Anal. 29 (2019) 811–870.
[31] C.E. Kenig, G. Ponce, L. Vega, The Cauchy problem for quasi-linear Schrödinger equations, Invent. Math. 158 (2004) 343–388.
[32] J. Marzuola, J. Metcalfe, D. Tataru, Quasilinear Schrödinger equations III: large data and short time, preprint, arXiv:2001.01014, 2020.
[33] G. Métivier, Para-Differential Calculus and Applications to the Cauchy Problem for Nonlinear Systems, vol. 5, Edizioni della Normale, 2008.
[34] J. Murphy, F. Pusateri, Almost global existence for cubic nonlinear Schrödinger equations in one space dimension, Discrete Contin. Dyn. Syst., Ser. A 37 (4) (2017) 2077–2102.
[35] F. Planchon, N. Tzevtkov, N. Visciglia, On the growth of Sobolev norms for NLS on $2d$ and $3d$ manifolds, in: Analysis and PDEs, vol. 10, 2017.
[36] M. Poppenberg, Smooth solutions for a class of fully nonlinear Schrödinger type equations, Nonlinear Anal., Theory Methods Appl. 45 (6) (2001) 723–741.
[37] M. Taylor, Tools for PDE, Amer. Math. Soc., 2007.
[38] Q. Zhang, Long-time existence for semi-linear Klein-Gordon equations with quadratic potential, Commun. Partial Differ. Equ. 35 (4) (2010) 630–668.