AN ABSTRACT BIRKHOFF NORMAL FORM THEOREM AND EXPONENTIAL TYPE STABILITY OF THE 1D NLS

LUCA BIASCO, JESSICA ELISA MASSETTI, AND MICHELA PROCESI

ABSTRACT. We study stability times for a family of parameter dependent nonlinear Schrödinger equations on the circle, close to the origin. Imposing a suitable Diophantine condition (first introduced by Bourgain), we prove a rather flexible Birkhoff Normal Form theorem, which implies, e.g., exponential and sub-exponential time estimates in the Sobolev and Gevrey class respectively.

CONTENTS

1. INTRODUCTION AND MAIN RESULTS

We consider families of NLS equations on the circle with external parameters of the form:

$$(1.1) \qquad \mathrm{i}u_t + u_{xx} - V * u + f(x, |u|^2)u = 0 \,,$$

where $\mathrm{i} = \sqrt{-1}$ and $V*$ is a Fourier multiplier

$$V * u = \sum_{j \in \mathbb{Z}} V_j u_j e^{\mathrm{i}jx} \,, \quad (V_j)_{j \in \mathbb{Z}} \in \mathtt{w}_q^\infty \,,$$

living in the weighted ℓ^∞ space

$$\mathbf{w}_q^\infty := \{V = (V_j)_{j \in \mathbb{Z}} \in \ell^\infty \ \mid \ |V|_q := \sup_{j \in \mathbb{Z}} |V_j| \langle j \rangle^q < \infty \}, \qquad q \geq 0,$$

where $\langle j \rangle := \max\{|j|, 1\}$, while $f(x, y)$ is 2π periodic and real analytic in x and is real analytic in y in a neighborhood of $y = 0$. We shall assume that $f(x, y)$ has a zero in $y = 0$. By analyticity, for some $\mathbf{a}, R > 0$ we have

$$(1.2) \qquad f(x, y) = \sum_{d=1}^\infty f^{(d)}(x) y^d, \quad |f|_{\mathbf{a}, R} := \sum_{d=1}^\infty |f^{(d)}|_{\mathbb{T}_\mathbf{a}} R^d < \infty,$$

where, given a real analytic function $g(x) = \sum_{j \in \mathbb{Z}} g_j e^{ijx}$, we set[1] $|g|_{\mathbb{T}_\mathbf{a}}^2 := \sum_{j \in \mathbb{Z}} |g_j|^2 e^{2\mathbf{a}|j|}$. Note that if f is independent of x (1.2) reduces to

$$(1.3) \qquad |f|_R := \sum_{d=1}^\infty |f^{(d)}| R^d < \infty.$$

Equation (1.1) is at least locally well-posed (say in a neighborhood of $u = 0$ in H^1, see e.g. Lemma 5.4) and has an elliptic fixed point at $u = 0$, so that an extremely natural question is to understand *stability times* for small initial data. One can informally state the problem as follows: let $E \subset H^1$ be some Banach space and consider (1.1) with initial datum u_0 such that $|u_0|_E \leq \delta \ll 1$. By local well posedness, the solution $u(t, x)$ of (1.1) with such initial datum exists and is in H^1.

We call *stability time* $T = T(\delta)$ the supremum of the times t such that for all $|u_0|_E \leq \delta$ one has $u(t, \cdot) \in E$ with $|u(t, \cdot)|_E \leq 2\delta$.

Computing the stability time $T(\delta)$ is out of reach, so the goal is to give lower (and possibly upper) bounds.

A good comparison is with the case of a finite dimensional Hamiltonian system with a non-degenerate elliptic fixed point, which in the standard complex symplectic coordinates $u_j = \frac{1}{\sqrt{2}}(q_j + ip_j)$ is described by the Hamiltonian

$$(1.4) \qquad \sum_{j=1}^n \omega_j |u_j|^2 + O(u^3), \quad \text{where } \omega_j \in \mathbb{R} \text{ are the } \textit{linear frequencies.}$$

Here if the frequencies ω are sufficiently non degenerate, say diophantine[2], then one can prove exponential lower bounds on $T(\delta)$ and, if the nonlinearity satisfies some suitable hypothesis (e.g. convexity or steepness), even super-exponential ones. This was proved in [MG95] (see also the recent paper [BFN15] and references therein). The strategy for obtaining exponential bounds is made of two main steps. The first one consists in the so-called Birkhoff normal form procedure: after $\mathbb{N} \geq 1$ steps the Hamiltonian (1.4) is transformed into

$$(1.5) \qquad \sum_{j=1}^n \omega_j |u_j|^2 + Z + R,$$

where Z depends only on the actions $(|u_i|^2)_{i=1}^n$ while $R = O(|u|^{2\mathbb{N}+3})$ contains terms of order at least $2\mathbb{N} + 3$ in $|u|$. It is well known that this procedure generically diverges in $\mathbb{N}$, so the second step consists in finding $\mathbb{N} = \mathbb{N}(\delta)$ which minimizes the size of the remainder R.

The problem of *long-time* stability for equations (1.1) has been studied by many authors. In the context of infinite chains with a finite range coupling, we mention [BFG88]. Regarding applications to PDEs (and particularly the NLS) the first results were given in [Bou96a] by Bourgain, who proved polynomial bounds for the stability times in the following terms: for any $\mathbb{N}$ there exists $p = p(\mathbb{N})$ such that initial data which are δ-small in the $H^{p'+p}$ norm stay small in the $H^{p'}$ norm, for times of order $\delta^{-\mathbb{N}}$. Afterwards, Bambusi in [Bam99b] proved that superanalytic initial data stay small in analytic norm, for times of order $e^{\ln(\frac{1}{\delta})^{1+b}}$, where $b > 0$.

---

[1]Namely g is a holomorphic function on the domain $\mathbb{T}_a := \{x \in \mathbb{C}/2\pi\mathbb{Z} : |\operatorname{Im} x| < a\}$ with L^2-trace on the boundary.

[2]A vector $\omega \in \mathbb{R}^n$ is called diophantine when it is badly approximated by rationals, i.e. it satisfies, for some $\gamma, \tau > 0$, $|k \cdot \omega| \geq \gamma |k|^{-\tau}, \quad \forall k \in \mathbb{Z}^n \setminus \{0\}$.

Following the strategy proposed in [Bam03] for the Klein-Gordon equation Bambusi and Grébert in [BG03] first considered equation (1.1) on $\mathbb{T}^d$ and then, in [BG06], proved polynomial bounds for a class of *tame-modulus* PDEs, which includes (1.1). Their main result is that for any $\mathbb{N} \gg 1$ there exists $p(\mathbb{N})$ (tending to infinity as $\mathbb{N} \to \infty$) such that for all $p \geq p(\mathbb{N})$ and all $\delta-$small initial data in H^p one has $T \geq C(\mathbb{N}, p)\delta^{-\mathbb{N}}$, provided $\delta < \delta_0(\mathbb{N}, p)$. Similar results were also proved for the Klein Gordon equation on tori and Zoll manifolds in [DS04], [DS06], [BDGS07]. Successively Faou and Grébert in [FG13] considered the case of analytic initial data and proved subexponential bounds of the form $T \geq e^{\ln(\frac{1}{\delta})^{1+b}}$, $b > 0$, for classes of NLS equations in $\mathbb{T}^d$ (which include (1.1) by taking $d = 1$). Regarding derivative NLS equations, the first results were in [YZ14] for the semilinear case. Recently, Feola and Iandoli in [FI] prove polynomial lower bounds for the stability times of reversible NLS equations with two derivatives in the nonlinearity.

A closely related topic is the study of orbital stability times close to periodic or quasi-periodic solutions of (1.1). In the case $E = H^1$, Bambusi in [Bam99a] proved a lower bound of the form $T \geq e^{\delta^{-b}}$, $b > 0$, for perturbations of the integrable cubic NLS close to a quasi-periodic solution. Regarding higher Sobolev norms, most results are in the periodic case. See [FGL13] (polynomial bounds for Sobolev initial data) and the preprint [MSW18] (subexponential bounds for Gevrey initial data).
A dual point of view is to construct special orbits for which the Sobolev norms grow as fast as possible (thus giving an upper bound on the stability times). As far as we are aware such results are mostly on $\mathbb{T}^2$ and in parameterless cases (for instance [CKS+10], [GK15], [GHP16]) and the time scales involved are much longer than our stability times (see [Gua14] for the instability of (1.1) on $\mathbb{T}^2$ and [Han14] for the instability of the plane wave in H^p with $p < 1$).

In this paper we propose an abstract Birkhoff normal form result (see Theorem 1.3) on weighted sequence spaces (based on ℓ^2) and deduce from it stability estimates for initial data in analytic, Gevrey and Sobolev class. An important difference of our approach with respect to the aforementioned papers and one of the main motivations of our work is that we use a different diophantine non-resonance condition on the linear frequencies, originally introduced in [Bou05] in the context of almost-periodic solutions. More precisely set

$$(1.6) \qquad \Omega_q := \left\{ \omega = (\omega_j)_{j \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}, \quad \sup_j |\omega_j - j^2| \langle j \rangle^q < 1/2 \right\}$$

and, for $\gamma > 0$, define the set of "good frequencies" as

$$(1.7) \qquad \mathtt{D}_{\gamma,q} := \left\{ \omega \in \Omega_q \ : \ |\omega \cdot \ell| > \gamma \prod_{n \in \mathbb{Z}} \frac{1}{(1 + |\ell_n|^2 \langle n \rangle^{2+q})} \cdot \quad \forall \ell \in \mathbb{Z}^{\mathbb{Z}} : |\ell| < \infty \right\},$$

It is known that $\mathtt{D}_{\gamma,q}$ is large with respect to a natural probability product measure on Ω_q (for a proof see [Bou05] or Lemma 4.1 in the present paper). It turns out that such diophantine conditions are very natural and easy to use in the context of PDEs on the circle with a superlinear dispersion law. Then from now on we shall fix $\gamma > 0$, $q \geq 0$ and assume that $\omega \in \mathtt{D}_{\gamma,q}$.

Remark 1.1. We note that some non-resonance condition on the frequencies is inevitable if one wants to prove *long-time* stability, indeed if one takes $V = 0$ and $f(x, |u|^2) = |u|^4$ then one can exhibit orbits in which the Sobolev norm is unstable in times of order δ^{-4}, see [GT12], [HP17].

At the formal level our BNF scheme is identical to the one used in finite dimensional systems, see formula (1.5). The fact that such a scheme may be applied in an infinite dimensional context follows from introducing a suitable norm (see Definition 1.2 and the comments thereafter); it turns out that our norm has explicit (and for us quite surprising) *immersion* properties (see Proposition 3.1) and allows good bounds on the solution of the *homological equation* (see Lemma 4.2). The gist of these properties is that they ensure that any vector field mapping a (neighborhood of) given Hilbert space in itself also maps (smaller neighborhoods of) *more regular* Hilbert spaces in themselves. Analogously also the vector field solving the homological equation maps *sufficiently more regular* Hilbert spaces in themselves.

To show that our procedure works in significative cases, we have computed stability times for various regularity classes. More precisely we improve the results in [FG13] on *analytic and Gevrey* initial data, see Theorem 1.1. Moreover we recover [BG06] on *Sobolev* initial data, giving an explicit control on the dependence of the stability time and of the smallness condition on the regularity, see Proposition 1.1 and the improved estimates of Theorem 1.2.

Comments on possible generalizations. In this paper we have considered the simplest possible example of dispersive PDE on the circle. One can easily see that the same strategy can be followed word by word in more general cases provided that the non-linearity does not contain derivatives and that the dispersion law is superlinear. A much more challenging question is to consider NLS models with derivatives in the non-linearity. As we have mentioned a semilinear case was discussed by [CMW]. A very promising approach to Birkhoff normal form for quasilinear PDEs is the one of [Del12]- [BD18] which was applied to fully-nonlinear reversible NLS equations in [FI]. It seems very plausible (at least in the reversible case) that one can adapt their methods (based on paralinearizations and paradifferential calculus) to our setting.

A natural generalization would be the extension to higher dimensions. While the immersion properties would work essentially in the same way, the diophantine condition should be adapted, for instance one could use the condition in [FG13].

Equation (1.1) contains infinitely many external parameters. Of course one would like to consider parameterless equations as in the very interesting recent preprint [BFG18]. In this direction a natural question would be to understand if one could impose similar diophantine conditions by tuning only one parameter such as the mass in the beam or wave equations (see, e.g., [Bam03], [BD18]).

Before explaining the abstract BNF procedure in detail let us describe our stability results.

1.1. **Stability results.**

Analytic and Gevrey initial data. Our result is similar to [FG13] in the sense that we also prove *subexponential bounds* on the time. We mention however that in [FG13] the control of the Sobolev norm in time is in a lower regularity space w.r.t. the initial datum. Recently we have been made aware of a preprint by Cong, Mi and Wang [CMW] in which the authors give subexponential bounds for Gevrey initial data of a model like (1.1), very similar to ours. A difference is that in their case the non linearity contains a derivative (see the comments after Theorem 1.1) but satisfies momentum conservation. The two results were obtained independently and contemporarily, anyway, the overall strategies of proofs are quite different. In particular our result is a consequence of the general Birkhoff Norma Form Theorem 1.3 and the non-resonance conditions are different (recall (1.7)).

To state our result, let us fix $0 < \theta < 1$, and define the function space[3]

$$(1.8) \qquad \mathtt{H}_{p,s,a} := \left\{ u(x) = \sum_{j \in \mathbb{Z}} u_j e^{\mathrm{i}jx} \in L^2 \ : \ |u|_{p,s,a}^2 := \sum_{j \in \mathbb{Z}} |u_j|^2 \langle j \rangle^{2p} e^{2a|j| + 2s\langle j \rangle^{\theta}} < \infty \right\}.$$

with the assumption $a \geq 0, s > 0, p > 1/2$. We remark that if $a > 0$ this is a space of analytic functions, while if $a = 0$ the functions have Gevrey regularity. Note that for technical reasons connected to the way in which we control the small divisors, we *cannot* deal with the purely analytic case $\theta = 1$, see Lemmas 6.1, 7.1. For this reason we denote this result as $\mathtt{G}$ (Gevrey case).

Our result, stated below, depends on some constants $\boldsymbol{\delta}_{\mathtt{G}}, \mathtt{T}_{\mathtt{G}}$, *explicitly defined* in Subsection A, and depending only on $\gamma, q, \mathtt{a}, R, |f|_{\mathtt{a},R}, p, s, a, \theta$.

Theorem 1.1 (Gevrey Stability). *Fix any $a \geq 0$, $s > 0$ such that $a + s < \mathtt{a}$ and any $p > 1/2$. For any $0 < \delta \leq \boldsymbol{\delta}_{\mathtt{G}}$ and any u_0 such that*

$$|u_0|_{p,s,a} \leq \delta \,,$$

---

[3]Actually $\mathtt{H}_{p,s,a}$ also depends on θ, however, since we think θ fixed, we omit to write explicitly the dependence on it.

the solution $u(t)$ of (1.1) with initial datum $u(0) = u_0$ exists and satisfies

$$|u(t)|_{p,s,a} \le 2\delta \quad \text{for all times} \quad |t| \le \frac{\mathtt{T_G}}{\delta^2} e^{\left(\ln \frac{\delta_{\mathtt{G}}}{\delta}\right)^{1+\theta/4}} .$$

Remark 1.2. Some comments on Theorem 1.1 are in order.

1. The main point in the proof is to verify that the abstract Birkhoff Normal Form Theorem 1.3 is applicable. Then we put the Hamiltonian of the NLS in Birkhoff normal form:

$$(1.9) \qquad \sum_{j \in \mathbb{Z}} \omega_j |u_j|^2 + Z + R ,$$

where Z depends only on the actions $(|u_i|^2)_{i \in \mathbb{Z}}$ while $R = O(|u|^{2\mathtt{N}+3})$ is analytic in a ball centered at zero of $\mathtt{h}_{p,s,a}$ and has a zero of order at least $2\mathtt{N} + 3$ in $u = 0$. Then we find $\mathtt{N} = \mathtt{N}(\delta)$ which minimizes the size of the remainder R.

2. We did not make an effort to maximize the exponent $1 + \theta/4$ in the stability time. In fact, by trivially modifying the proof, one could get $1 + \theta/(2^+)$. We remark that in [CMW], in which $\theta = 1/2$, the exponent is better, i.e. it is $1 + 1/(2^+)$.

Sobolev initial data. Here our first goal was to recover by our methods the result of [BG06], computing explicitly all the constants in the estimates. In particular it is fundamental to have a good control on the dependence of the stabiliy time T on the the regularity p. Indeed there are two natural ways of taking a small ball around zero: reducing the size δ or increasing the regularity p. A crucial point is that, in the case of Sobolev regularity, the number of BNF steps that one may perform is (apparently unavoidably) tied to the regularity p. This is clearly seen in [BG06], where the number of steps is $\sim \sqrt{p}$. It seemed an interesting point to verify how our approach worked in such a case, and wether we would see the same phenomenon.

As before, our estimates depend on some constants, denoted by $\tau_{\mathtt{S}}, \delta_{\mathtt{S}}, \mathtt{k_S}, \mathtt{T_S},$, *explicitly defined* in Appendix A. These constants depend only on $\gamma, q, \mathtt{a}, R, |f|_{\mathtt{a},R}$.

Proposition 1.1 (A quantitative version of [BG06]). *Consider equation (1.1) with f satisfying (1.2) for* $\mathtt{a}, R > 0$. *For any $p \ge 3\tau_{\mathtt{S}} + 1$ and any initial datum $u(0) = u_0$ satisfying*

$$(1.10) \qquad |u_0|_{H^p} := |u_0|_{L^2} + |\partial_x^p u_0|_{L^2} \le \delta \le \delta_{\mathtt{S}}(\mathtt{k_S} p)^{-3p}$$

the solution $u(t)$ of (1.1) with initial datum $u(0) = u_0$ exists and satisfies

$$(1.11) \qquad |u(t)|_{H_p} \le 4\delta \quad \text{for all times} \quad |t| \le \mathtt{T_S} p^{-5p} \left(\frac{\delta_{\mathtt{S}}}{\delta}\right)^{\frac{2(p-1)}{\tau_{\mathtt{S}}}} .$$

Remark 1.3. Also in this result we just have to verify the hypotheses of Theorem 1.3. However as it happens in [BG06] the maximum number $\mathtt{N}$ of steps of BNF we can perform depends on p, in particular $\mathtt{N} = [\frac{p-1}{\tau_{\mathtt{S}}}]$. This is in fact slightly better than the previously cited paper ($\mathtt{N} \sim p$ instead of $\sqrt{p}$). On the other hand it is not difficult to show that the bound $\delta \le \delta_{\mathtt{S}}(\mathtt{k_S} p)^{-3p}$ is essentially optimal (see Remark 10.1).

Looking at the proof of the Theorem or even constructing other finite-dimensional models, one can see that in the traslation invariant case, the very restrictive smallness condition in (1.10) is only due to interactions between the modes $0, 1, -1$ and all the others. It then seems natural to consider initial data for which the energy on such modes is smaller, namely $|u_0|_{L^2} \le 2^{-p}\delta$. We refer to this case as $\mathtt{M}$, the relevant constants can be found in Appendix A

Theorem 1.2. *Consider equation (1.1) with f independent of x and satisfying (1.3) for $R > 0$. For any $p > 3\tau_{\mathtt{M}} + 1$ and for any initial datum $u(0) = u_0$ satisfying*

$$(1.12) \qquad |u_0|_{H^p} \le \delta \le \frac{\delta_{\mathtt{M}}}{\sqrt{p}}, \quad |u_0|_{L^2} \le 2^{-p}\delta$$

the solution $u(t)$ of (1.1) exists and satisfies

$$(1.13) \qquad |u(t)|_{H^p} \leq 8\delta \quad for \ all \ times \quad |t| \leq \mathtt{T}_{\mathtt{M}} \left(\frac{8\delta_{\mathtt{M}}^2}{(p-1)\delta^2} \right)^{\frac{p-1}{\tau_{\mathtt{M}}}} \ .$$

Remark 1.4. Note that, since the L^2 norm is a constant of motion, one trivially has $|u(t)|_{L^2} \leq 2^{-p}\delta$. Comparing with (1.11), we see that the time estimate is more or less the same but now it holds in a *much bigger neighborhood* of zero ($\delta \leq p^{-1/2}$ instead of $\leq p^{-3p}$).

If one requires a stronger condition on the L^2 norm, i.e., $|u_0|_{L^2} \leq 3^{-p}\delta$, it turns out that the size of the perturbation is *exponentially decreasing in p* and, therefore, keeping δ fixed and sending p to infinity one immediately obtains stability.

The main difference between the Gevrey and Sobolev cases is that in the latter the number of BNF steps $\mathtt{N}$ depends on the regularity, while in the former it is independent. Thus in the Sobolev case we cannot fix both δ and p and optimize in $\mathtt{N}$. What we can do is to fix δ and find an *optimal* regularity $p(\delta)$, which maximizes the stability time. It turns out that the two cases $\mathtt{S}$ and $\mathtt{M}$ behave differently. Indeed the weaker smallness condition (1.12) allows us to take much bigger $p(\delta)$, obtaining much longer stability times. As before our statements depend on some constants, denoted by $\bar{\delta}_{\mathtt{S}}, \bar{\delta}_{\mathtt{M}}$ *explicitly defined* in Subsection A.

Corollary 1.1 (Sobolev stability: optimization)**.**
(S) *For any $0 < \delta \leq \bar{\delta}_{\mathtt{S}}$ and any u_0 such that*

$$(1.14) \qquad |u_0|_{H^p} \leq \delta \,, \qquad p = p(\delta) := 1 + \frac{\ln(\delta_{\mathtt{S}}/\delta)}{6 \ln \ln(\delta_{\mathtt{S}}/\delta)} \,,$$

the solution $u(t)$ of (1.1) with initial datum $u(0) = u_0$ exists and satisfies

$$(1.15) \qquad |u(t)|_{H^p} \leq 4\delta \quad for \ all \ times \quad |t| \leq \mathtt{T}_{\mathtt{S}} e^{\frac{\ln^2(\delta_{\mathtt{S}}/\delta)}{4\tau_{\mathtt{S}} \ln \ln(\delta_{\mathtt{S}}/\delta)}} \,.$$

(M) *Assume that f in (1.1) is independent of x. For any $0 < \delta \leq \bar{\delta}_{\mathtt{M}}$ and*

$$(1.16) \qquad \forall \, p \geq p(\delta) := \frac{\delta_{\mathtt{M}}^2}{\delta^2} \,, \quad \forall u_0 \quad s.t. \quad |u_0|_{H^p} \leq \delta \,, \quad |u_0|_{L^2} \leq 2^{-p}\delta \,,$$

the solution $u(t)$ of (1.1) with initial datum $u(0) = u_0$ exists and satisfies

$$(1.17) \qquad |u(t)|_{H^p} \leq 8\delta \quad for \ all \ times \quad |t| \leq \mathtt{T}_{\mathtt{M}} e^{\frac{\delta_{\mathtt{M}}^2}{\tau_{\mathtt{M}}\delta^2}} \,.$$

Remark 1.5. Some remarks on Corollary 1.1 are in order.

Note that (1.15) is the stability time computed in [BFG88] for short range couplings.

1. We will prove the case $\mathtt{M}$ only for $p = p(\delta)$, the general case being analogous[4] (with the same constants!) also if $p \geq p(\delta)$.

2. One can easily restate Corollary 1.1 in terms of the Sobolev exponent p, instead of δ, since the map $\delta \to p(\delta)$ is injective.

Remark 1.6 (finite dimensional examples). It is interesting to compute the stability times predicted by our theorems for initial data supported on a finite number of modes. To this purpose consider an initial datum $u^{(0)}$ uniformly distributed over the modes $1, \ldots, j$:

$$|u_i^{(0)}| = \varepsilon \,, \quad \forall i = 1, \ldots, j$$

Theorem 1.1 with $a = 0, p = 1$ states that if $\varepsilon \leq \varepsilon_{\mathtt{G}} := \delta_{\mathtt{G}} e^{-2j^{\theta}}$ then $u(t)$ stays stable, in Gevrey norm, for times of order $e^{\left(\ln \frac{\varepsilon_{\mathtt{G}}}{\varepsilon}\right)^{1+\theta/4}}$.

---

[4]Indeed, thanks to the immersion property of our norms (see Proposition 6.1 below) the canonical transformation putting the system in Birkhoff Normal Form (see Theorem 8.1 below) in the p-case is simply the restriction to H^p of the one of the $p(\delta)$-case.

Now if $\varepsilon \leq \varepsilon_{\mathsf{M}}(p) := \delta_{\mathsf{M}} j^{-p-1}/\sqrt{p}$ we have $|u_0|_{H_p} < \delta$ and $|u_0|_{L^2} < 2^{-p}\delta$; then by Theorem 1.2 the solution $u(t)$ stays stable, in H^p norm, for times of order $T \sim \left(\frac{\varepsilon_{\mathsf{M}}(p)}{\varepsilon}\right)^{2(p-1)/\tau_{\mathsf{M}}}$. Maximizing the time in p with fixed ε we get

$$(1.18) \qquad p \sim \frac{\ln(\frac{\delta_{\mathsf{M}}^2}{\varepsilon})}{2\ln j}, \quad T \sim e^{\frac{(\ln \varepsilon^{-1})^2}{\ln j}}$$

provided that $\varepsilon \lesssim j^{-7\tau_{\mathsf{M}}}$. Explicitly we get a weaker constraint on ε and a better time estimate. Of course one could play the same game directly with the estimate of Proposition 1.1. As it should be expected the time estimate is more or less the same as 1.18 but the smallness condition is much stronger, i.e. of the type $\varepsilon \lesssim e^{-2j^\theta}$.

1.2. **The abstract Birkhoff Normal Form.** We start by setting our functional framework. The main point is to introduce a *weighted majorant norm* which penalizes the terms in the Hamiltonian which do not preserve momentum, see Definition 1.1.

Let us pass to the Fourier side via the identification

$$(1.19) \qquad u(x) = \sum_{j\in\mathbb{Z}} u_j e^{\mathrm{i}jx} \;\mapsto\; u = (u_j)_{j\in\mathbb{Z}},$$

where u belongs to some complete subspace of ℓ^2. Fix the symplectic structure to be

$$(1.20) \qquad \mathrm{i}\sum_j du_j \wedge d\bar{u}_j.$$

In this framework the Hamiltonian of (1.1) is

$$(1.21) \qquad H_{\mathrm{NLS}}(u) := D_\omega + P, \qquad \text{where}$$

$$D_\omega := \sum_{j\in\mathbb{Z}} \omega_j |u_j|^2, \quad P := \int_{\mathbb{T}} F(x, |u(x)|^2)dx, \quad F(x,y) := \int_0^y f(x,s)ds.$$

We shall always work with quite regular solutions; given a real sequence $\mathsf{w} = (\mathsf{w}_j)_{j\in\mathbb{Z}}$, with $\mathsf{w}_j \geq 1$ let us set the Hilbert space[5]

$$(1.22) \qquad \mathtt{h}_{\mathsf{w}} := \left\{ u := (u_j)_{j\in\mathbb{Z}} \in \ell^2(\mathbb{C}) : \quad |u|_{\mathsf{w}}^2 := \sum_{j\in\mathbb{Z}} \mathsf{w}_j^2 |u_j|^2 < \infty \right\}.$$

As examples of $\mathtt{h}_{\mathsf{w}}$ we consider:

  G) (Gevrey case) $\mathsf{w}(p,s,a) := \left(\langle j\rangle^p e^{a|j|+s\langle j\rangle^\theta}\right)_{j\in\mathbb{Z}}$, which is isometrically isomorphic, by Fourier transform, to $\mathtt{H}_{p,s,a}$ defined in (1.8).

  S) (Sobolev case) $\mathsf{w}(p) := \mathsf{w}(p,0,0) = (\langle j\rangle^p)_{j\in\mathbb{Z}}$, which is isometrically isomorphic, by Fourier transform, to $\mathtt{H}_{p,0,0}$ defined in (1.8) and is equivalent to H^p equipped with the norm $|\cdot|_{L^2} + |\partial_x^p \cdot|_{L^2}$ with equivalence constants independent of p (see (5.28))

  M) (Modified-Sobolev case) $\mathsf{w}_j = \lfloor j\rfloor^p$, where $\lfloor j\rfloor := \max\{|j|,2\}$; this space is equivalent to H^p equipped with the norm $2^p|\cdot|_{L^2} + |\partial_x^p \cdot|_{L^2}$ with equivalence constants independent of p (see (5.30))

Here and in the following, given $r > 0$, by $B_r(\mathtt{h}_{\mathsf{w}})$ we mean the closed ball of radius r centered at the origin of $\mathtt{h}_{\mathsf{w}}$.

In the following we always consider Hamiltonians $H : B_r(\mathtt{h}_{\mathsf{w}}) \to \mathbb{R}$ such that there exists a pointwise absolutely convergent power series expansion[6]

$$H(u) = \sum_{\substack{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}^{\mathbb{Z}}, \\ |\boldsymbol{\alpha}|+|\boldsymbol{\beta}|<\infty}} H_{\boldsymbol{\alpha},\boldsymbol{\beta}} u^{\boldsymbol{\alpha}} \bar{u}^{\boldsymbol{\beta}}, \qquad u^{\boldsymbol{\alpha}} := \prod_{j\in\mathbb{Z}} u_j^{\alpha_j}$$

---

[5] Endowed with the scalar product $(u,v)_{\mathtt{h}_{\mathsf{w}}} := \sum_{j\in\mathbb{Z}} \mathsf{w}_j^2 u_j \bar{v}_j$.

[6] As usual given a vector $k \in \mathbb{Z}^{\mathbb{Z}}$, $|k| := \sum_{j\in\mathbb{Z}} |k_j|$.

with the following properties:

    (i) Reality condition:

$$(1.23) \qquad\qquad\qquad\qquad H_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \overline{H}_{\boldsymbol{\beta},\boldsymbol{\alpha}} \,,$$

        this means that H is real analytic in the real and imaginary part of u (see section 2);

    (ii) Mass conservation:

$$(1.24) \qquad\qquad\qquad\qquad H_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 0 \quad \text{if } |\boldsymbol{\alpha}| \neq |\boldsymbol{\beta}| \,,$$

        namely the Hamiltonian Poisson commutes with the mass $\sum_{j\in\mathbb{Z}} |u_j|^2$;

The Hamiltonian functions being defined modulo a constant term, we shall assume without loss of generality that $H(0) = 0$.

We say that a Hamiltonian H as above preserves momentum when

$$(1.25) \qquad\qquad H_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 0 \quad \text{if } \quad \pi(\boldsymbol{\alpha} - \boldsymbol{\beta}) := \sum_{j\in\mathbb{Z}} j(\boldsymbol{\alpha}_j - \boldsymbol{\beta}_j) \neq 0 \,,$$

namely the Hamiltonian H Poisson commutes with $\sum_{j\in\mathbb{Z}} j|u_j|^2$. Note that if the nonlinearity f in equation (1.1) does not depend on the variable x, then the Hamiltonian P in (1.21) preserves momentum.

Definition 1.1 (η-majorant analytic Hamiltonians). For $\eta \geq 0, r > 0$ let $\mathcal{A}_{r,\eta}(\mathtt{h_w})$ be the space of Hamiltonians as above such that the η-majorant

$$(1.26) \qquad\qquad \underline{H}_\eta(u) := \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}^{\mathbb{Z}}} |H_{\boldsymbol{\alpha},\boldsymbol{\beta}}| e^{\eta|\pi(\boldsymbol{\alpha}-\boldsymbol{\beta})|} u^{\boldsymbol{\alpha}} \bar{u}^{\boldsymbol{\beta}}$$

is point-wise absolutely convergent on $B_r(\mathtt{h_w})$. If we take $\eta = 0$ we denote $\underline{H}_0(u) = \underline{H}(u)$ as the majorant of H.

The exponential weight $e^{\eta|\pi(\boldsymbol{\alpha}-\boldsymbol{\beta})|}$ is added in order to ensure that the monomials which do not preserve momentum have an exponentially small coefficient.

We will say that a Hamiltonian $H(u) \in \mathcal{A}_{r,\eta}(\mathtt{h_w})$ is *η-regular* if $X_{\underline{H}_\eta} : B_r(\mathtt{h_w}) \to \mathtt{h_w}$ and is uniformly bounded, where $X_{\underline{H}_\eta}$ is the vector field associated to the η-majorant Hamiltonian in (1.26). More precisely we give the following

Definition 1.2 (η-regular Hamiltonians). For $\eta \geq 0, r > 0$ let $\mathcal{H}_{r,\eta}(\mathtt{h_w})$ be the subspace of $\mathcal{A}_{r,\eta}(\mathtt{h_w})$ of those Hamiltonians H such that

$$|H|_{\mathcal{H}_{r,\eta}(\mathtt{h_w})} = |H|_{r,\eta,\mathtt{w}} := r^{-1}\left(\sup_{|u|_{\mathtt{h_w}} \leq r} \left| X_{\underline{H}_\eta} \right|_{\mathtt{h_w}} \right) < \infty \,.$$

We shall show in Section 2 that this guarantees that the Hamiltonian flow of H exists at least locally and generates a symplectic transformation on $\mathtt{h_w}$.

Remark 1.7. Definition 1.2 with $\eta = 0$, i.e. the idea of controlling an analytic function through the sup of its *Cauchy majorant*, dates back to Cauchy-Kovalevskaya. In the context of analytic functions on Hilbert spaces, this class of functions is defined and studied, with a slightly different approach, in [Nik86] and [KP10], where it is referred to as "normally analytic" functions.

Regarding the idea of introducing a weight which penalizes monomials which do not preserve momentum, this was used already in [Bam03].

In our work the crucial point is that all the dependence on the parameters $r, \eta, \mathtt{w}$ of the norm in Definition 1.2 can be *encoded* in the coefficients

$$(1.27) \qquad c_{r,\eta,\mathtt{w}}^{(j)}(\boldsymbol{\alpha},\boldsymbol{\beta}) := r^{|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|-2} e^{\eta|\pi(\boldsymbol{\alpha}-\boldsymbol{\beta})|} \frac{\mathtt{w}_j^2}{\mathtt{w}^{\boldsymbol{\alpha}+\boldsymbol{\beta}}} \,, \qquad \mathtt{w}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} = \prod_{j\in\mathbb{Z}} \mathtt{w}_j^{\boldsymbol{\alpha}_j+\boldsymbol{\beta}_j}$$

defined for any $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}$ and $j \in \mathbb{Z}$ (see formula (3.1) and Lemma 3.1). This allows us to give a simple and explicit condition which guarantees the *immersion* $\mathcal{H}_{r,\eta}(\mathtt{h_w}) \subseteq \mathcal{H}_{r',\eta'}(\mathtt{h_{w'}})$ in terms of the ratio of the coefficients

$c^{(j)}_{r,\eta,\mathtt{w}}(\boldsymbol{\alpha},\boldsymbol{\beta})$, $c^{(j)}_{r',\eta',\mathtt{w}'}(\boldsymbol{\alpha},\boldsymbol{\beta})$, see Proposition 3.1.

As it is well known a Birkhoff Normal Form is achieved by an iterative procedure. Let us describe the general step. Given a Hamiltonian

$$H = \sum_{j\in\mathbb{Z}} \omega_j |u_j|^2 + Z + R \,, \tag{1.28}$$

where Z is a normal form and R has a zero of degree say $2\mathtt{d}+2$ (with $\mathtt{d} \geq 1$) at $u = 0$, we look for a change of variables, which conjugates H to a Hamiltonian $\sum_{j\in\mathbb{Z}} \omega_j |u_j|^2 + Z' + R'$ so that now R has a zero of degree at least $2\mathtt{d}+4$. The desired change of variables is generated by the time one flow of a Hamiltonian S which solves the *homological equation*[7]

$$\{\sum_{j\in\mathbb{Z}} \omega_j |u_j|^2, S\} = R \,.$$

As for the immersion properties, given[8] $r' \leq r, \eta' \leq \eta$ and $\mathtt{w}' \geq \mathtt{w}$ such that $\mathcal{H}_{r,\eta}(\mathtt{h}_{\mathtt{w}}) \subseteq \mathcal{H}_{r',\eta'}(\mathtt{h}_{\mathtt{w}'})$, in Proposition 4.2 and Lemma 5.2 we give a simple and explicit condition -in terms of the ratio of the coefficients $c^{(j)}_{r,\eta,\mathtt{w}}(\boldsymbol{\alpha},\boldsymbol{\beta})$, $c^{(j)}_{r',\eta',\mathtt{w}'}(\boldsymbol{\alpha},\boldsymbol{\beta})$- which ensures that if $R \in \mathcal{H}_{r,\eta}(\mathtt{h}_{\mathtt{w}})$ is appropriately small, then S is well defined and generates a close to identity change of variables $B_{r'}(\mathtt{h}_{\mathtt{w}'}) \to \mathtt{h}_{\mathtt{w}'}$. With this procedure we start in some phase space $\mathtt{h}_{\mathtt{w}}$ and then show the existence of the Birkhoff change of variables on a ball which *not only* has a smaller radius but is taken *in the stronger toplogy* $\mathtt{h}_{\mathtt{w}'}$. Note that this is *not* a smoothing change of variables: it is defined *from the smaller space to itself.*

Starting with a Hamiltonian as in (1.28) with a zero of order 4, in order to reach the form (1.9) we need to perform $\mathtt{N}$ steps of BNF. To this purpose we make the following

Assumption 1. *We say that $\eta \geq 0$ and two weights $\mathtt{w}_0 \leq \mathtt{w}$ satisfy the Birkhoff assumption at step $\mathtt{N} \geq 1$ if the following holds. The exists a sequence of weights $\mathtt{w}_0 \leq \mathtt{w}_1 \leq \cdots \leq \mathtt{w}_{\mathtt{N}} = \mathtt{w}$ such that*

$$\mathfrak{C} := \max\left\{1, \sup_{0\leq n<\mathtt{N}} \sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta}\\\boldsymbol{\alpha}_j+\boldsymbol{\beta}_j\neq 0}} \frac{c^{(j)}_{\varrho_n^*,\eta_{n+1},\mathtt{w}_{n+1}}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{\varrho_n,\eta_n,\mathtt{w}_n}(\boldsymbol{\alpha},\boldsymbol{\beta})}\right\} < \infty \,,$$

$$\mathfrak{K} := \max\left\{1, \sup_{0\leq n<\mathtt{N}} \sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta}\\\boldsymbol{\alpha}_j+\boldsymbol{\beta}_j\neq 0}} \frac{c^{(j)}_{\varrho_n^*,\eta_{n+1},\mathtt{w}_{n+1}}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{\varrho_n,\eta_n,\mathtt{w}_n}(\boldsymbol{\alpha},\boldsymbol{\beta})|\omega\cdot(\boldsymbol{\alpha}-\boldsymbol{\beta})|}\right\} < \infty \,,$$

$$\mathfrak{K}^{\sharp} := \max\left\{1, \sup_{0\leq n<\mathtt{N}} \sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta}\\\boldsymbol{\alpha}_j+\boldsymbol{\beta}_j\neq 0}} \frac{c^{(j)}_{\varrho_n^*,\eta_{\mathtt{N}},\mathtt{w}_{\mathtt{N}}}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{\varrho_n,\eta_n,\mathtt{w}_n}(\boldsymbol{\alpha},\boldsymbol{\beta})|\omega\cdot(\boldsymbol{\alpha}-\boldsymbol{\beta})|}\right\} < \infty \,, \tag{1.29}$$

where

$$\varrho_n = (2 - \frac{n}{\mathtt{N}})\,, \quad \eta_n = (1 - \frac{n}{\mathtt{N}})\eta\,, \quad 0 \leq n \leq \mathtt{N}\,, \qquad \varrho_n^* = \frac{\varrho_{n+1}+\varrho_n}{2}\,, \quad 0 \leq n < \mathtt{N}\,.$$

Informally speaking $\mathfrak{C} < \infty$ guarantees the immersion properties at each step, while $\mathfrak{K} < \infty$ guarantees that one can solve the homological equation at each step. Finally $\mathfrak{K}^{\sharp} < \infty$ guarantees that the *composition* of the changes of variables of all steps is well defined and close to identity on some ball $B_r(\mathtt{h}_{\mathtt{w}_{\mathtt{N}}})$.

Let

$$\mathcal{K}_r(\mathtt{h}_{\mathtt{w}}) := \{H \in \mathcal{H}_{r,0}(\mathtt{h}_{\mathtt{w}}) \ | \ H = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^{\mathbb{Z}}} H_{\boldsymbol{\alpha},\boldsymbol{\alpha}}|u|^{2\boldsymbol{\alpha}}\} \tag{1.30}$$

be the subspace of normal form Hamiltonians.

---

[7]Since $\omega \in \mathtt{D}_{\gamma,q}$, w.l.o.g. we may assume that R is in the range of the operator $\{\sum_{j\in\mathbb{Z}} \omega_j |u_j|^2, \cdot\}$.

[8]As usual $\mathtt{w} \leq \mathtt{w}'$ means that $\mathtt{w}_j \leq \mathtt{w}'_j$ for every $j \in \mathbb{Z}$.

Theorem 1.3 (Abstract Birkhoff Normal Form)**.** *Consider a Hamiltonian of the form*

$$(1.31) \qquad\qquad H = D_\omega + G\,, \quad D_\omega = \sum_j \omega_j |u|_j^2$$

with $\omega \in \mathtt{D}_{\gamma,q}$ and $G \in \mathcal{H}_{\bar{r},\eta}(\mathtt{h}_{\mathtt{w}_0})$, for some $\bar{r} > 0$, $\eta \geq 0$. Assume moreover that G has a zero of order at least 4 at $u = 0$. Consider $\mathtt{N} \geq 1$ and $\mathtt{w} \geq \mathtt{w}_0$ such that $\eta, \mathtt{w}_0, \mathtt{w}$ satisfy the Birkhoff assumption at step $\mathtt{N}$. Set

$$(1.32) \qquad \widehat{\mathtt{r}} := \min\left\{ \frac{r_\star}{\sqrt{\mathtt{N}\max\{\mathfrak{C}\mathfrak{K}, \mathfrak{K}^\sharp\}}}, \, \frac{\bar{r}}{2} \right\}\,, \qquad where \qquad r_\star := \bar{r}\sqrt{\frac{\gamma}{2^{11}e|G|_{\bar{r},\eta,\mathtt{w}_0}}}\,.$$

Then for all $0 < r \leq \widehat{\mathtt{r}}$ there exists an invertible[9] symplectic change of variables

$$(1.33) \qquad \Psi: \quad B_r(\mathtt{h}_{\mathtt{w}}) \mapsto B_{2r}(\mathtt{h}_{\mathtt{w}})\,, \qquad \sup_{u\in B_r(\mathtt{h}_{\mathtt{w}})} |\Psi(u) - u|_{\mathtt{w}} \leq \widehat{\mathsf{C}}_1 r^3 \leq \frac{r}{8}\,, \qquad \widehat{\mathsf{C}}_1 := \frac{\mathfrak{K}^\sharp}{2^7 e r_\star^2}\,,$$

such that in the new coordinates

$$H \circ \Psi = D_\omega + Z + R\,, \qquad Z \in \mathcal{K}_r(\mathtt{h}_{\mathtt{w}})\,,$$

where

$$(1.34) \quad |Z|_{r,0,\mathtt{w}} \leq \widehat{\mathsf{C}}_2 r^2\,, \qquad |R|_{r,0,\mathtt{w}} \leq \widehat{\mathsf{C}}_3 r^{2(\mathtt{N}+1)}\,, \qquad with \qquad \widehat{\mathsf{C}}_2 := \frac{8|G|_{\bar{r},\eta,\mathtt{w}_0}}{\bar{r}^2}\,, \qquad \widehat{\mathsf{C}}_3 := \frac{\gamma}{2^9 e r_\star^2}\left(\frac{\mathfrak{C}\mathfrak{K}\mathtt{N}}{4 r_\star^2}\right)^{\mathtt{N}}\,.$$

The theorem follows by a straightforward iteration, see Section 5.

As it is well known the bounds (1.34) imply a lower bound on the stability time; we discuss this in Corollary 5.1 where we show that the solution $u(t)$ of the Hamiltonian flow of (1.31) with initial datum $u(0) = u_0$ such that $|u_0|_{\mathtt{w}} \leq \frac{3r}{8}$ exists and satisfies

$$|u(t)|_{\mathtt{w}} \leq r \quad \text{for all times} \quad |t| \leq \frac{1}{8\widehat{\mathsf{C}}_3 r^{2(\mathtt{N}+1)}}\,.$$

By Theorem 1.3 and Corollary 5.1, in order to prove the stability results we only need to define suitable sequence spaces verifying Assumption 1. In particular we consider the three applications $\mathtt{G}, \mathtt{S}, \mathtt{M}$ introduced at page 7. Another interesting example (suggested to us by Z. Hani) could be the space

$$\left\{ (u_j) \in L^2 : \quad |u|^2 := \sum_j |u_j|^2 e^{\ln(\lfloor j \rfloor)^2} < \infty \right\}\,,$$

where $\lfloor j \rfloor = \max\{|j|, 2\}$. In this case one may get $T \approx \delta^{\ln(\ln(1/\delta))}$.

---

[9] in the sense that there exists a symplectic change of variables $\Phi: B_r(\mathtt{h}_{\mathtt{w}}) \mapsto B_{2r}(\mathtt{h}_{\mathtt{w}})$ such that $\Psi \circ \Phi u = \Phi \circ \Psi u = u$, $\forall u \in B_{\frac{7}{8}r}(\mathtt{h}_{\mathtt{w}})$.

Part 1. An abstract framework for Birkhoff normal form on sequences spaces

2. SYMPLECTIC STRUCTURE AND HAMILTONIAN FLOWS

Spaces of Hamiltonians. As explained in the Introduction our wheighted spaces $\mathtt{h_w}$ are contained in $\ell^2(\mathbb{C})$, so we endow them with the standard symplectic structure coming from the Hermitian product on $\ell^2(\mathbb{C})$. We identify $\ell^2(\mathbb{C})$ with $\ell^2(\mathbb{R}) \times \ell^2(\mathbb{R})$ through $u_j = (x_j + iy_j)/\sqrt{2}$ and induce on $\ell^2(\mathbb{C})$ the structure of a real symplectic Hilbert space[10] by setting, for any $(u^{(1)}, u^{(2)}) \in \ell^2(\mathbb{C}) \times \ell^2(\mathbb{C})$,

$$\langle u^{(1)}, u^{(2)} \rangle = \sum_j \left(x_j^{(1)} x_j^{(2)} + y_j^{(1)} y_j^{(2)} \right), \quad \omega(u^{(1)}, u^{(2)}) = \sum_j \left(y_j^{(1)} x_j^{(2)} - x_j^{(1)} y_j^{(2)} \right),$$

which are the standard scalar product and symplectic form $\Omega = \sum_j dy_j \wedge dx_j$.
For convenience and to keep track of the complex structure, one often writes the vector fields and the differential forms in complex notation, that is

$$\Omega = i \sum_j du_j \wedge d\bar{u}_j, \quad X_H^{(j)} = i \frac{\partial}{\partial \bar{u}_j} H$$

where the one form and vector field are defined through the identification between $\mathbb{C}$ and $\mathbb{R}^2$, given by

$$du_j = \frac{1}{\sqrt{2}} (dx_j + idy_j), \quad d\bar{u}_j = \frac{1}{\sqrt{2}} (dx_j - idy_j),$$

$$\frac{\partial}{\partial u_j} = \frac{1}{\sqrt{2}} \left(\frac{\partial}{\partial x_j} - i \frac{\partial}{\partial y_j} \right), \quad \frac{\partial}{\partial \bar{u}_j} = \frac{1}{\sqrt{2}} \left(\frac{\partial}{\partial x_j} + i \frac{\partial}{\partial y_j} \right).$$

Remark 2.1. By mass conservation and since $H(0) = 0$, it is straightforward to prove that the norm $|\cdot|_{r,\eta,\mathtt{w}}$ is increasing in the radius parameter r (see also Proposition3.1).

Note that if $|H|_{r,\eta,\mathtt{w}} < \infty$ then H admits an analytic extension $\widehat{H}$, that is

$$(u_+, u_-) \in B_r(\ell^2(\mathbb{C})) \times B_r(\ell^2(\mathbb{C})) \to \widehat{H}(u_+, u_-) : \qquad H(u) = \widehat{H}(u, \bar{u}),$$

whose Taylor series expansion is

$$\widehat{H}(u_+, u_-) = \sum_{\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}}^* H_{\boldsymbol{\alpha}, \boldsymbol{\beta}} u_+^{\boldsymbol{\alpha}} u_-^{\boldsymbol{\beta}}.$$

where we denote by $\sum^*$ the sum restricted to those $\boldsymbol{\alpha}, \boldsymbol{\beta} : |\boldsymbol{\alpha}| = |\boldsymbol{\beta}| < \infty$.
One can see that

$$\frac{\partial}{\partial \bar{u}_j} H(u) = \frac{\partial \widehat{H}(u_+, u_-)}{\partial u_{-,j}} \bigg|_{u_+ = \bar{u}_- = u}.$$

Poisson structure and hamiltonian flows. The scale $\{\mathcal{H}_{r,\eta}(\mathtt{h_w})\}_{r>0}$ is a Banach-Poisson algebra in the following sense

Proposition 2.1. *For $0 < \rho \leq r$ and $\eta > 0$ we have*

(2.1) $$|\{F, G\}|_{r,\eta,\mathtt{w}} \leq 4 \left(1 + \frac{r}{\rho} \right) |F|_{r+\rho,\eta,\mathtt{w}} |G|_{r+\rho,\eta,\mathtt{w}}.$$

---

[10]We recall that given a complex Hilbert space H with a Hermitian product $(\cdot, \cdot)$, its realification is a real symplectic Hilbert space with scalar product and symplectic form given by

$$\langle u, v \rangle = 2\mathrm{Re}(u, v), \quad \omega(u, v) = 2\mathrm{Im}(u, v).$$

Proof. It is essentially contained in [BBP13]. See in particular Lemma 2.16 of [BBP13] with $n = 0$ (no action variables here) and no s and s' (no actions variable here). Note that the constant in Lemma 2.16 is 8, instead of 4 in the present paper, because of the presence there of action variables which scale different from the cartesian ones (namely $(2r)^2$ instead of $2r$). Recall also the required properties of the space E (named $\mathbf{h_w}$ in the present paper) mentioned after Definition 2.5. $\square$

The following Lemma is a simple corollary and its proof is postponed to the appendix

Lemma 2.1 (Hamiltonian flow)**.** *Let* $0 < \rho < r$, *and* $S \in \mathcal{H}_{r+\rho,\eta}(\mathbf{h_w})$ *with*

$$(2.2) \qquad |S|_{r+\rho,\eta,\mathbf{w}} \leq \delta := \frac{\rho}{8e(r+\rho)}.$$

Then the time 1-*Hamiltonian flow* $\Phi_S^1 : B_r(\mathbf{h_w}) \to B_{r+\rho}(\mathbf{h_w})$ *is well defined, analytic, symplectic with*

$$(2.3) \qquad \sup_{u \in B_r(\mathbf{h_w})} \left| \Phi_S^1(u) - u \right|_{\mathbf{h_w}} \leq (r+\rho)|S|_{r+\rho,\eta,\mathbf{w}} \leq \frac{\rho}{8e}.$$

For any $H \in \mathcal{H}_{r+\rho,\eta}(\mathbf{h_w})$ *we have that* $H \circ \Phi_S^1 = e^{\{S,\cdot\}} H \in \mathcal{H}_{r,\eta}(\mathbf{h_w})$ *and*

$$(2.4) \qquad \left| e^{\{S,\cdot\}} H \right|_{r,\eta,\mathbf{w}} \leq 2|H|_{r+\rho,\eta,\mathbf{w}},$$

$$(2.5) \qquad \left| \left(e^{\{S,\cdot\}} - \mathrm{id} \right) H \right|_{r,\eta,\mathbf{w}} \leq \delta^{-1}|S|_{r+\rho,\eta,\mathbf{w}}|H|_{r+\rho,\eta,\mathbf{w}},$$

$$(2.6) \qquad \left| \left(e^{\{S,\cdot\}} - \mathrm{id} -\{S,\cdot\} \right) H \right|_{r,\eta,\mathbf{w}} \leq \frac{1}{2}\delta^{-2}|S|_{r+\rho,\eta,\mathbf{w}}^2|H|_{r+\rho,\eta,\mathbf{w}}$$

More generally for any $h \in \mathbb{N}$ *and any sequence* $(c_k)_{k\in\mathbb{N}}$ *with* $|c_k| \leq 1/k!$, *we have*

$$(2.7) \qquad \left| \sum_{k \geq h} c_k \, \mathrm{ad}_S^k(H) \right|_{r,\eta,\mathbf{w}} \leq 2|H|_{r+\rho,\eta,\mathbf{w}} \left(|S|_{r+\rho,\eta,\mathbf{w}}/2\delta \right)^h,$$

where $\mathrm{ad}_S(\cdot) := \{S,\cdot\}$.

3. Immersions for spaces of Hamiltonians.

Given two positive sequences $\mathbf{w} = (\mathbf{w}_j)_{j\in\mathbb{Z}}, \mathbf{w}' = \left(\mathbf{w}'_j\right)_{j\in\mathbb{Z}}$ we write that $\mathbf{w} \leq \mathbf{w}'$ if the inequality holds point wise, namely

$$\mathbf{w} \leq \mathbf{w}' \quad :\iff \quad \mathbf{w}_j \leq \mathbf{w}'_j, \quad \forall j \in \mathbb{Z}.$$

In this way if $r' \leq r$ and $\mathbf{w} \leq \mathbf{w}'$ then $B_{r'}(\mathbf{h_{w'}}) \subseteq B_r(\mathbf{h_w})$. Consequently if $r' \leq r, \eta' \leq \eta$ and $\mathbf{w} \leq \mathbf{w}'$ then $\mathcal{A}_{r,\eta}(\mathbf{h_w}) \subseteq \mathcal{A}_{r',\eta'}(\mathbf{h_{w'}})$.

We thus wish to study conditions on $(r,\eta,\mathbf{w}), (r^*,\eta',\mathbf{w}')$ (with $r^* \leq r$) which ensure that $\mathcal{H}_{r,\eta}(\mathbf{h_w}) \subseteq \mathcal{H}_{r^*,\eta'}(\mathbf{h_{w'}})$. Note that this is not obvious at all, since we are asking that an Hamiltonian vector field of $X_H \in \mathcal{H}_{r,\eta}(\mathbf{h_w})$, when restricted to the smaller domain $B_{r^*}(\mathbf{h_{w'}})$ belongs to the smaller space $\mathbf{h_{w'}}$.

The coefficients $c_{r,\eta,\mathbf{w}}^{(j)}(\alpha,\beta)$**.** Let us start by rewriting the norm $|\cdot|_{r,\eta,\mathbf{w}}$ in a more *adimensional* way. In this way all the dependence on the parameters $r,\eta,\mathbf{w}$ of the norm $|\cdot|_{r,\eta,\mathbf{w}}$ is *encoded* in the coefficients (1.27).

Definition 3.1. For any $H \in \mathcal{H}_{r,\eta}(\mathbf{h_w})$ we define a map

$$B_1(\ell^2) \to \ell^2, \quad y = (y_j)_{j\in\mathbb{Z}} \mapsto \left(Y_H^{(j)}(y;r,\eta,\mathbf{w}) \right)_{j\in\mathbb{Z}}$$

by setting

$$(3.1) \qquad Y_H^{(j)}(y; r, \eta, \mathtt{w}) := \sum_* |H_{\boldsymbol{\alpha},\boldsymbol{\beta}}| \frac{(\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j)}{2} c_{r,\eta,\mathtt{w}}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) y^{\boldsymbol{\alpha}+\boldsymbol{\beta}-e_j}$$

where e_j is the j-th basis vector in $\mathbb{N}^{\mathbb{Z}}$, while the coefficient $c_{r,\eta,\mathtt{w}}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ was defined in (1.27). For brevity, we set

$$\sum_* := \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}:|\boldsymbol{\alpha}|=|\boldsymbol{\beta}|} .$$

The momentum $\pi(\cdot)$ was defined in (1.25).

The vector field Y_H is a majorant analytic function on ℓ^2 which has the *same norm as H*. Since the majorant analytic functions on a given space have a natural ordering this gives us a natural criterion for immersions, as formalized in the following Lemma.

Lemma 3.1. *Let $r, r^* > 0$, $\eta, \eta' \geq 0$, $\mathtt{w}, \mathtt{w}' \in \mathbb{R}_+^{\mathbb{Z}}$. The following properties hold.*

(i) The norm of H can be expressed as

$$(3.2) \qquad |H|_{r,\eta,\mathtt{w}} = \sup_{|y|_{\ell^2} \leq 1} |Y_H(y; r, \eta, \mathtt{w})|_{\ell^2}$$

(ii) Given $H^{(1)} \in \mathcal{H}_{r^*,\eta',\mathtt{w}'}$ and $H^{(2)} \in \mathcal{H}_{r,\eta,\mathtt{w}}$,
   such that for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}$ and $j \in \mathbb{Z}$ with $\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0$ one has

$$|H_{\boldsymbol{\alpha},\boldsymbol{\beta}}^{(1)}| c_{r^*,\eta',\mathtt{w}'}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq c |H_{\boldsymbol{\alpha},\boldsymbol{\beta}}^{(2)}| c_{r,\eta,\mathtt{w}}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta}),$$

   for some $c > 0$, then

$$|H^{(1)}|_{r^*,\eta',\mathtt{w}'} \leq c |H^{(2)}|_{r,\eta,\mathtt{w}}.$$

Proof. See appendix B. $\qquad \square$

As a corollary we get the following "immersion theorem" for spaces of Hamiltonians

Proposition 3.1 (Immersion)**.** *Let $r, r^* > 0$, $\eta, \eta' \geq 0$, $\mathtt{w}, \mathtt{w}' \in \mathbb{R}_+^{\mathbb{Z}}$. If*

$$(3.3) \qquad C := \sup_{\substack{j \in \mathbb{Z}, \, \boldsymbol{\alpha},\boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}} \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0}} \frac{c_{r^*,\eta',\mathtt{w}'}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta})}{c_{r,\eta,\mathtt{w}}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta})} < \infty,$$

then $\mathcal{H}_{r,\eta}(\mathtt{h_w}) \subseteq \mathcal{H}_{r^*,\eta'}(\mathtt{h_{w'}})$, with

$$(3.4) \qquad |H|_{r^*,\eta',\mathtt{w}'} \leq C |H|_{r,\eta,\mathtt{w}}.$$

In particular $|\cdot|_{r,\eta,\mathtt{w}}$ is increasing in r and η, namely if $r^* \leq r$ and $\eta' \leq \eta$ then

$$|H|_{r^*,\eta',\mathtt{w}} \leq |H|_{r,\eta,\mathtt{w}}.$$

Moreover, if $r^* \leq r$, $\mathtt{w} \leq \mathtt{w}'$ and $H \in \mathcal{K}_{r,\eta}(\mathtt{h_w})$ then

$$(3.5) \qquad |H|_{r^*,\eta',\mathtt{w}'} \leq |H|_{r,\eta,\mathtt{w}}.$$

Furthermore, if H preserves momentum then

$$(3.6) \qquad |H|_{r^*,\eta',\mathtt{w}'} \leq C_0 |H|_{r,\eta,\mathtt{w}},$$

where

$$(3.7) \qquad C_0 := \sup_{\substack{j \in \mathbb{Z},\, \boldsymbol{\alpha},\boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}, \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0, \\ \sum_i i(\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i)=0}} \frac{c^{(j)}_{r^*,\eta',\mathtt{w}'}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{r,\eta,\mathtt{w}}(\boldsymbol{\alpha},\boldsymbol{\beta})} < \infty\,,$$

Proof. Inequality (3.4) directly follows from Lemma 3.1 (ii), while (3.5) follows directly by (1.27) since in the kernel $\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0$ implies $\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \geq 2$. The momentum preserving case follows analogously. $\qquad\square$

Remark 3.1. The above immersion properties, with different norm and in a different context, were implicitly used by Bourgain in [Bou05].

4. Small divisors and homological equation

Let us consider the set of frequencies

$$(4.1) \qquad \Omega_q := \left\{ \omega = (\omega_j)_{j \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}, \quad \sup_j |\omega_j - j^2|\langle j \rangle^q < 1/2 \right\};$$

this set is isomorphic to $[-1/2, 1/2]^{\mathbb{Z}}$ via the identification

$$(4.2) \qquad \xi \mapsto \omega(\xi)\,, \quad \text{where} \quad \omega_j(\xi) = j^2 + \frac{\xi_j}{\langle j \rangle^q}\,.$$

We endow Ω_q with the probability measure μ induced[11] by the product measure on $[-1/2, 1/2]^{\mathbb{Z}}$.

We now define the set of *Diophantine* frequencies, the following definition is a slight generalization of the one given by Bourgain in [Bou05].

Definition 4.1. Given $\gamma > 0$ and $q \geq 0$, we denote by $\mathsf{D}_{\gamma,q} \equiv \mathsf{D}^{\mu_1,\mu_2}_{\gamma,q}$ the set of μ_1, μ_2, γ-*Diophantine frequencies*

$$(4.3) \qquad \mathsf{D}^{\mu_1,\mu_2}_{\gamma,q} := \left\{ \omega \in \Omega_q \,:\, |\omega \cdot \ell| > \gamma \prod_{n \in \mathbb{Z}} \frac{1}{(1 + |\ell_n|^{\mu_1}\langle n \rangle^{\mu_2+q})}\,, \quad \forall \ell \in \mathbb{Z}^{\mathbb{Z}} : 0 < |\ell| < \infty \right\}.$$

Now we have that

Lemma 4.1. *For $\mu_1, \mu_2 > 1$ the exists a positive constant $C_{\mathtt{meas}}(\mu_1, \mu_2)$ such that*

$$\mu\big(\Omega_q \setminus \mathsf{D}^{\mu_1,\mu_2}_{\gamma,q}\big) \leq C_{\mathtt{meas}}(\mu_1, \mu_2)\gamma\,.$$

Proof. In Appendix C $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

This means that, for all $\mu_1, \mu_2 > 1$, Diophantine frequencies are *typical* in Ω_q in the sense that they have full measure. Here and in the following we shall always assume that

$$(4.4) \qquad 0 < \gamma \leq 1\,, \qquad \omega \in \mathsf{D}^{2,2}_{\gamma,q} = \mathsf{D}_{\gamma,q}\,.$$

---

[11]Denoting by μ the measure in Ω_q and by ν the product measure on $[-1/2, 1/2]^{\mathbb{Z}}$, then $\mu(A) = \nu(\omega^{(-1)}(A))$ for all sets $A \subset \Omega_q$ such that $\omega^{(-1)}(A)$ is ν-measurable.

In the remaining part of this section, on appropriate source and target spaces, we will study the invertibility of the "Lie derivative" operator

$$(4.5) \qquad L_\omega : \quad H \mapsto L_\omega H := \sum_* \mathrm{i}(\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})) H_{\boldsymbol{\alpha},\boldsymbol{\beta}} u^{\boldsymbol{\alpha}} \bar{u}^{\boldsymbol{\beta}},$$

which is nothing but the action of the Poisson bracket $\left\{ \sum_j \omega_j |u_j|^2, \cdot \right\}$ on H.

Recalling the definition of $\mathcal{K}_r(\mathbf{h_w})$ in (1.30) we give the following

Definition 4.2. Let

$$(4.6) \qquad \mathcal{R}_{r,\eta}(\mathbf{h_w}) := \{ H \in \mathcal{H}_{r,\eta}(\mathbf{h_w}) \ | \ \ H = \sum_{\boldsymbol{\alpha} \neq \boldsymbol{\beta}} H_{\boldsymbol{\alpha},\boldsymbol{\beta}} u^{\boldsymbol{\alpha}} \bar{u}^{\boldsymbol{\beta}} \}$$

Then we have the decomposition $\mathcal{H}_{r,\eta}(\mathbf{h_w}) = \mathcal{R}_{r,\eta}(\mathbf{h_w}) \oplus \mathcal{K}_r(\mathbf{h_w})$ and the continuous projections[12]

$$(4.7) \qquad |\Pi_{\mathcal{K}} H|_{r,\eta,\mathbf{w}}, |\Pi_{\mathcal{R}} H|_{r,\eta,\mathbf{w}} \leq |H|_{r,\eta,\mathbf{w}}.$$

Obviously for diophantine frequency $\mathcal{R}_{r,\eta}(\mathbf{h_w})$ and $\mathcal{K}_r(\mathbf{h_w})$ represent the range and kernel of L_ω.

For any $r, \eta, \mathbf{w}$ and $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}$ recall the coefficient defined in (1.27)

$$c_{r,\eta,\mathbf{w}}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) := r^{|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|-2} e^{\eta|\pi(\boldsymbol{\alpha}-\boldsymbol{\beta})|} \frac{\mathbf{w}_j^2}{\mathbf{w}^{\boldsymbol{\alpha}+\boldsymbol{\beta}}} .$$

In following Lemma we consider $R \in \mathcal{R}_{r,\eta}(\mathbf{h_w})$ and state sufficient conditions which ensure that $L_\omega^{-1} R \in \mathcal{R}_{r^*,\eta'}(\mathbf{h_{w'}})$.

Lemma 4.2 (Homological equation). *Fix $\omega \in \mathtt{D}_{\gamma,q}$. Consider two ordered weights $0 < r^* \leq r, 0 \leq \eta' \leq \eta, \mathbf{w}' \geq \mathbf{w}$, such that*

$$(4.8) \qquad K := \gamma \sup_{\substack{j \in \mathbb{Z}, \, \boldsymbol{\alpha} \neq \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}} \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0}} \frac{c_{r^*,\eta',\mathbf{w}'}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta})}{c_{r,\eta,\mathbf{w}}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) |\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|} < \infty,$$

then for any $R \in \mathcal{R}_{r,\eta}(\mathbf{h_w})$ the homological equation

$$L_\omega S = R$$

has a unique solution $S = L_\omega^{-1} R$ in $\mathcal{R}_{r^*,\eta'}(\mathbf{h_{w'}})$, which satisfies

$$(4.9) \qquad \left| L_\omega^{-1} R \right|_{r^*,\eta',\mathbf{w}'} \leq \gamma^{-1} K |R|_{r,\eta,\mathbf{w}}.$$

Similarly, if R preserves momentum, assuming only

$$(4.10) \qquad K_0 := \gamma \sup_{\substack{j \in \mathbb{Z}, \, \boldsymbol{\alpha} \neq \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}} \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0 \\ \sum_i i(\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i) = 0}} \frac{c_{r^*,\eta',\mathbf{w}'}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta})}{c_{r,\eta,\mathbf{w}}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) |\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|} < \infty,$$

we have that S also preserves momentum and

$$(4.11) \qquad \left| L_\omega^{-1} R \right|_{r^*,\eta',\mathbf{w}'} \leq \gamma^{-1} K_0 |R|_{r,\eta,\mathbf{w}}.$$

---

[12]Explicitly $\Pi_{\mathcal{K}} H := \sum_{\boldsymbol{\alpha}=\boldsymbol{\beta}} H_{\boldsymbol{\alpha},\boldsymbol{\beta}} u^{\boldsymbol{\alpha}} \bar{u}^{\boldsymbol{\beta}}, \Pi_{\mathcal{R}} H := \sum_{\boldsymbol{\alpha} \neq \boldsymbol{\beta}} H_{\boldsymbol{\alpha},\boldsymbol{\beta}} u^{\boldsymbol{\alpha}} \bar{u}^{\boldsymbol{\beta}}.$

Proof. Given any Hamiltonian $R \in \mathcal{R}$, the formal solution of $L_S = R$ is given by

$$(4.12) \qquad L_\omega^{-1} R = \sum_{|\boldsymbol{\alpha}| = |\boldsymbol{\beta}|, \, \boldsymbol{\alpha} \neq \boldsymbol{\beta}} \frac{1}{\mathrm{i}(\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta}))} R_{\boldsymbol{\alpha},\boldsymbol{\beta}} u^{\boldsymbol{\alpha}} \bar{u}^{\boldsymbol{\beta}} \, ,$$

where $u \in B_{r^*}(\mathbf{h}_{\mathbf{w}'})$. By Lemma 3.1 (ii) (applied to $H^{(1)} = L_\omega^{-1} R$ and $H^{(2)} = R$) and (4.8), we get (4.9). The momentum preserving case is analogous. $\qquad\square$

5. Abstract Birkhoff Normal Form

In this section we prove the abstract Birkoff normal form Theorem 1.3. We start by defining a degree decomposition which endows $\mathcal{H}_{r,\eta}(\mathbf{h}_{\mathbf{w}})$ with a graded Poisson algebra structure.

Definition 5.1 (minimal scaling degree). We say that H has *minimal* scaling degree $\mathtt{d} = \mathtt{d}(H)$ (at zero) if

$$\begin{aligned} H_{\boldsymbol{\alpha},\boldsymbol{\beta}} &= 0 \,, \quad \forall\, \boldsymbol{\alpha}, \boldsymbol{\beta} : \quad |\boldsymbol{\alpha}| = |\boldsymbol{\beta}| \leq \mathtt{d} \,, \\ H_{\boldsymbol{\alpha},\boldsymbol{\beta}} &\neq 0 \,, \quad \text{for some } \boldsymbol{\alpha}, \boldsymbol{\beta} : \quad |\boldsymbol{\alpha}| = |\boldsymbol{\beta}| = \mathtt{d} + 1 \,. \end{aligned}$$

We say that $\mathtt{d}(0) = +\infty$.

Essentially H has scaling degree $\mathtt{d}$ if and only if it has a zero of order $2\mathtt{d} + 2$ at zero, we prefer this notation because we find it more intrinsic, it produces a graded Poisson algebra structure and one has the following

Lemma 5.1. *If $H \in \mathcal{H}_{r,\eta}(\mathbf{h}_{\mathbf{w}})$ with $\mathtt{d}(H) \geq \mathtt{d}$, then for all $r^* \leq r$ one has*

$$|H|_{r^*,\eta,\mathbf{w}} \leq \left(\frac{r^*}{r} \right)^{2\mathtt{d}} |H|_{r,\eta,\mathbf{w}} \,.$$

Proof. Recalling (1.27), we have

$$\frac{c_{r^*,\eta,\mathbf{w}}^{(j)}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c_{r,\eta,\mathbf{w}}^{(j)}(\boldsymbol{\alpha},\boldsymbol{\beta})} = \left(\frac{r^*}{r} \right)^{|\boldsymbol{\alpha}| + |\boldsymbol{\beta}| - 2} \,.$$

Since $|\boldsymbol{\alpha}| + |\boldsymbol{\beta}| - 2 \geq 2\mathtt{d}$, the inequality follows by Proposition 3.1. $\qquad\square$

The normal form will be proved iteratively by means of the following Lemma, which constitutes the main step of the procedure.

Basically we start with a Hamiltonian $H = D_\omega + Z + R$ with $Z \in \mathcal{K}_r(\mathbf{h}_{\mathbf{w}})$ in normal form and $R \in \mathcal{R}_{r,\eta}(\mathbf{h}_{\mathbf{w}})$ of minimal degree $\mathtt{d}$, and we consider $r' \leq r, \eta' \leq \eta, \mathbf{w}' \geq \mathbf{w}$ so that $\mathcal{H}_{r,\eta}(\mathbf{h}_{\mathbf{w}}) \subseteq \mathcal{H}_{r',\eta'}(\mathbf{h}_{\mathbf{w}'})$. Then we give a sufficient condition which ensures the existence of a change of variables $\Phi : B_{r'}(\mathbf{h}_{\mathbf{w}'}) \to B_r(\mathbf{h}_{\mathbf{w}'})$ such that

$$H \circ \Phi = D_\omega + Z' + R' \,,$$

with $Z', R' \in \mathcal{H}_{r',\eta'}(\mathbf{h}_{\mathbf{w}'})$ and R' of minimal degree $\mathtt{d} + 1$.

Lemma 5.2. *Fix $\omega \in \mathtt{D}_{\gamma,q}$. Let $r > r' > 0, \eta \geq \eta' \geq 0, \mathbf{w} \leq \mathbf{w}'$. Consider*

$$H = D_\omega + Z + R \,, \quad Z \in \mathcal{K}_r(\mathbf{h}_{\mathbf{w}}) \,, \quad R \in \mathcal{R}_{r,\eta}(\mathbf{h}_{\mathbf{w}}) \,, \quad \mathtt{d}(Z) \geq 1 \,, \quad \mathtt{d}(R) \geq \mathtt{d} \geq 1 \,.$$

Assume that (3.3) and (4.8) hold and that[13]

$$(5.1) \qquad |R|_{r,\eta,\mathbf{w}} \leq \frac{\gamma \delta}{K} \,, \qquad \text{with} \quad \delta := \frac{r - r'}{16er} \,.$$

---

[13]K is the constant in (4.8).

Then there exists a change of variables

$$(5.2) \qquad\qquad \Phi \ : \ B_{r'}(\mathtt{h}_{\mathtt{w}'}) \ \to \ B_r(\mathtt{h}_{\mathtt{w}'}),$$

such that

$$H \circ \Phi = D_\omega + Z' + R', \quad Z' \in \mathcal{K}_{r',\eta'}(\mathtt{h}_{\mathtt{w}'}), \quad R' \in \mathcal{R}_{r',\eta'}(\mathtt{h}_{\mathtt{w}'}), \quad \mathtt{d}(Z') \geq 1, \quad \mathtt{d}(R') \geq \mathtt{d}+1.$$

Moreover[14]

$$
\begin{aligned}
(5.3) \qquad
|Z'|_{r',\eta',\mathtt{w}'} &\leq |Z|_{r,\eta,\mathtt{w}} + (\gamma\delta)^{-1}K|R|_{r,\eta,\mathtt{w}}(C|R|_{r,\eta,\mathtt{w}} + |Z|_{r,\eta,\mathtt{w}}), \\
|R'|_{r',\eta',\mathtt{w}'} &\leq (\gamma\delta)^{-1}K|R|_{r,\eta,\mathtt{w}}(C|R|_{r,\eta,\mathtt{w}} + |Z|_{r,\eta,\mathtt{w}}).
\end{aligned}
$$

Finally, for $\mathtt{w}^\sharp \geq \mathtt{w}'$, *assume the further conditions*

$$(5.4) \qquad \gamma \sup_{\substack{j\in\mathbb{Z},\, \boldsymbol{\alpha}\neq\boldsymbol{\beta}\in\mathbb{N}^{\mathbb{Z}} \\ \boldsymbol{\alpha}_j+\boldsymbol{\beta}_j\neq 0}} \frac{c^{(j)}_{r^*,\eta',\mathtt{w}^\sharp}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{r,\eta,\mathtt{w}}(\boldsymbol{\alpha},\boldsymbol{\beta})|\omega\cdot(\boldsymbol{\alpha}-\boldsymbol{\beta})|} =: K^\sharp < \infty, \qquad r^* := \frac{r'+r}{2}$$

and

$$(5.5) \qquad\qquad\qquad |R|_{r,\eta,\mathtt{w}} \leq \frac{\gamma\delta}{K^\sharp}.$$

Then

$$\Phi\big|_{B_{r'}(\mathtt{h}_{\mathtt{w}^\sharp})} \ : \ B_{r'}(\mathtt{h}_{\mathtt{w}^\sharp}) \ \to \ B_r(\mathtt{h}_{\mathtt{w}^\sharp}),$$

$$(5.6) \qquad \sup_{u\in B_{r'}(\mathtt{h}_{\mathtt{w}^\sharp})} |\Phi(u) - u|_{\mathtt{h}_{\mathtt{w}^\sharp}} \leq r\gamma^{-1}K^\sharp|R|_{r,\eta,\mathtt{w}}.$$

Moreover if R *preserves momentum, assuming only that*

$$(5.7) \qquad K^\sharp_0 := \gamma \sup_{\substack{j\in\mathbb{Z},\, \boldsymbol{\alpha}\neq\boldsymbol{\beta}\in\mathbb{N}^{\mathbb{Z}} \\ \boldsymbol{\alpha}_j+\boldsymbol{\beta}_j\neq 0, \\ \sum_i i(\boldsymbol{\alpha}_i-\boldsymbol{\beta}_i)=0}} \frac{c^{(j)}_{r^*,\eta',\mathtt{w}^\sharp}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{r,\eta,\mathtt{w}}(\boldsymbol{\alpha},\boldsymbol{\beta})|\omega\cdot(\boldsymbol{\alpha}-\boldsymbol{\beta})|} < \infty$$

and that (5.1)*,* (5.5) *hold with* $K_0, K^\sharp_0$ *instead of* $K, K^\sharp$ *we have that* R' *preserves momentum and* (5.6) *holds with* $K^\sharp_0$ *instead of* $K^\sharp$.

Proof. By Lemma 4.2 let $S = L_\omega^{-1}R$ in $\mathcal{R}_{r^*,\eta'}(\mathtt{h}_{\mathtt{w}'})$ be the unique solution of the homological equation $L_\omega S = R$ on $B_{r^*}(\mathtt{h}_{\mathtt{w}'})$. Note that $\mathtt{d}(S) \geq \mathtt{d}$. We have

$$(5.8) \qquad\qquad\qquad |S|_{r^*,\eta',\mathtt{w}'} \leq \gamma^{-1}K|R|_{r,\eta,\mathtt{w}}.$$

We now apply Lemma 2.1 with $(r,\eta,\mathtt{w}) \rightsquigarrow (r',\eta',\mathtt{w}')$ and $\rho := r^* - r'$. Note that (5.1) and (5.8) imply (2.2). We define $\Phi := \Phi^1_S$ and compute

$$H' := H \circ \Phi = D_\omega + Z + (e^{\{S,\cdot\}} - \mathrm{id} - \{S,\cdot\})D_\omega + (e^{\{S,\cdot\}} - \mathrm{id})(Z + R) =$$

$$= D_\omega + Z - \sum_{j=2}^\infty \frac{(\mathrm{ad}S)^{j-1}}{j!}R + (e^{\{S,\cdot\}} - \mathrm{id})(Z + R).$$

We now set

$$Z' = \Pi_\mathcal{K}H' - D_\omega, \quad R' = \Pi_\mathcal{R}H'.$$

Since the scaling degree is additive w.r.t. Poisson brackets, we have that $\mathtt{d}(Z') \geq 1$ and $\mathtt{d}(R') \geq \mathtt{d}+1$. By (2.7)

$$
\begin{aligned}
|Z'|_{r',\eta',\mathtt{w}'} &\leq |Z|_{r',\eta',\mathtt{w}'} + (\gamma\delta)^{-1}K|R|_{r,\eta,\mathtt{w}}(|R|_{r^*,\eta',\mathtt{w}'} + |Z|_{r^*,\eta',\mathtt{w}'}), \\
|R'|_{r',\eta',\mathtt{w}'} &\leq (\gamma\delta)^{-1}K|R|_{r,\eta,\mathtt{w}}(|R|_{r^*,\eta',\mathtt{w}'} + |Z|_{r^*,\eta',\mathtt{w}'}).
\end{aligned}
$$

---

[14]C is defined in (3.3).

Since (4.8) holds we can apply Proposition 3.1: by (3.4) and (3.5) we get

$$|R|_{r^*,\eta',\mathbf{w}'} \leq C|R|_{r,\eta,\mathbf{w}}\,, \qquad |Z|_{r^*,\eta',\mathbf{w}'} \leq |Z|_{r,\eta,\mathbf{w}}\,.$$

(5.3) follows.

Finally assume (5.5) and (5.4). By Lemma 4.2 let $S^\sharp = L_\omega^{-1}R$ in $\mathcal{R}_{r^*,\eta'}(\mathbf{h}_{\mathbf{w}^\sharp})$ be the solution of the homological equation $L_\omega S^\sharp = R$ on $B_{r^*}(\mathbf{h}_{\mathbf{w}^\sharp}) \subseteq B_{r^*}(\mathbf{h}_{\mathbf{w}'})$. Since S and $S^\sharp$ solve the same linear equation on $B_{r^*}(\mathbf{h}_{\mathbf{w}^\sharp})$, we have that

$$S^\sharp = S\big|_{B_{r^*}(\mathbf{h}_{\mathbf{w}^\sharp})}\,.$$

By (4.9) we get

$$(5.9) \qquad\qquad |S|_{r^*,\eta',\mathbf{w}^\sharp} \leq \gamma^{-1}K^\sharp|R|_{r,\eta,\mathbf{w}}\,.$$

We now apply Lemma 2.1 with $(r,\eta,\mathbf{w}) \rightsquigarrow (r',\eta',\mathbf{w}^\sharp)$ and $\rho := r^* - r'$. Note that (5.5) and (5.9) imply (2.2). Then (5.6) follows by (2.3) and (5.9).

The momentum preserving case is analogous. □

Theorem 1.3 follows Given $\eta \geq 0$ and a sequence of weights $\mathbf{w}_0 \leq \mathbf{w}_1 \leq \cdots \leq \mathbf{w}_N = \mathbf{w}$. For any given $r > 0$ we set

$$(5.10) \qquad r_n = \left(2 - \frac{n}{N}\right)r\,, \quad \eta_n = \left(1 - \frac{n}{N}\right)\eta\,, \quad 0 \leq n \leq N\,, \qquad r_n^* = \frac{r_{n+1}+r_n}{2}\,, \quad 0 \leq n < N\,.$$

From Assumption 1 and (1.27) we have[15]

$$(5.11) \qquad\qquad \max\left\{1,\ \sup_{0 \leq n < N}\ \sup_{\substack{j,\boldsymbol\alpha,\boldsymbol\beta \\ \boldsymbol\alpha_j + \boldsymbol\beta_j \neq 0}} \frac{c^{(j)}_{r_n^*,\eta_{n+1},\mathbf{w}_{n+1}}(\boldsymbol\alpha,\boldsymbol\beta)}{c^{(j)}_{r_n,\eta_n,\mathbf{w}_n}(\boldsymbol\alpha,\boldsymbol\beta)}\right\} = \mathfrak{C} < \infty\,,$$

$$(5.12) \qquad\qquad \max\left\{1,\ \sup_{0 \leq n < N}\ \sup_{\substack{j,\boldsymbol\alpha,\boldsymbol\beta \\ \boldsymbol\alpha_j + \boldsymbol\beta_j \neq 0}} \frac{c^{(j)}_{r_n^*,\eta_{n+1},\mathbf{w}_{n+1}}(\boldsymbol\alpha,\boldsymbol\beta)}{c^{(j)}_{r_n,\eta_n,\mathbf{w}_n}(\boldsymbol\alpha,\boldsymbol\beta)|\omega \cdot (\boldsymbol\alpha - \boldsymbol\beta)|}\right\} = \mathfrak{K} < \infty\,,$$

$$(5.13) \qquad\qquad \max\left\{1,\ \sup_{0 \leq n < N}\ \sup_{\substack{j,\boldsymbol\alpha,\boldsymbol\beta \\ \boldsymbol\alpha_j + \boldsymbol\beta_j \neq 0}} \frac{c^{(j)}_{r_n^*,\eta_N,\mathbf{w}_N}(\boldsymbol\alpha,\boldsymbol\beta)}{c^{(j)}_{r_n,\eta_n,\mathbf{w}_n}(\boldsymbol\alpha,\boldsymbol\beta)|\omega \cdot (\boldsymbol\alpha - \boldsymbol\beta)|}\right\} = \mathfrak{K}^\sharp < \infty\,.$$

For brevity we set

$$(5.14) \qquad \mathbf{h}_n := \mathbf{h}_{\mathbf{w}_n}\,, \quad \mathcal{H}_n := \mathcal{H}_{r_n,\eta_n}(\mathbf{h}_n)\,, \quad 0 \leq n \leq N\,, \quad \mathcal{H}_{n,*} := \mathcal{H}_{r_n^*,\eta_{n+1}}(\mathbf{h}_{n+1})\,, \quad 0 \leq n < N\,,$$

and, correspondingly, $\mathcal{R}_n, \mathcal{K}_n, \mathcal{R}_{n,*}, \mathcal{K}_{n,*}$ and

$$(5.15) \qquad\qquad |\cdot|_n := |\cdot|_{r_n,\eta_n,\mathbf{w}_n}\,, \qquad |\cdot|_{n,*} := |\cdot|_{r_n^*,\eta_{n+1},\mathbf{w}_{n+1}}\,.$$

Lemma 5.3. *By Assumption* (5.11) *we have the immersion properties*

$$(5.16) \qquad\qquad \mathcal{H}_0 \subseteq \mathcal{H}_{0,*} \subseteq \cdots \subseteq \mathcal{H}_n \subseteq \mathcal{H}_{n,*} \subseteq \mathcal{H}_{n+1} \subseteq \cdots \subseteq \mathcal{H}_N\,,$$

with estimates

$$\begin{aligned}
H \in \mathcal{H}_n &\implies & |H|_{n,*} &\leq \widehat{C}|H|_n\,, & 0 \leq n \leq i \leq N-1 \\
(5.17) \qquad H \in \mathcal{K}_n &\implies & |H|_{n,*} &\leq |H|_n\,, & 0 \leq n \leq i \leq N-1\,.
\end{aligned}$$

---

[15] we are just using the fact that the ratio $c^{(j)}_{r,\eta,\mathbf{w}}(\boldsymbol\alpha,\boldsymbol\beta)/c^{(j)}_{r',\eta',\mathbf{w}'}(\boldsymbol\alpha,\boldsymbol\beta)$ depends on r,r' only through their ratio.

Proof. We apply Proposition 3.1 with

$$r, \eta, \mathtt{w} \rightsquigarrow r_n, \eta_n, \mathtt{w}_n, \quad r^*, \eta', \mathtt{w}' \rightsquigarrow r_n^*, \eta_{n+1}, \mathtt{w}_{n+1},$$

by noting that the bound (3.3) follows from (5.11). The bounds in (5.17) follow form (3.4) and (3.5). The chain of inclusions (5.16) follows. $\qquad\square$

Proof of Theorem 1.3. We will prove the thesis inductively. Let us start by noticing that

$$\widehat{\mathtt{r}} = \min \left\{ \frac{\bar{r}}{8\sqrt{|G|_{\bar{r},\eta,\mathtt{w}_0}}} \sqrt{\frac{\gamma\hat{\delta}}{\max\{\mathfrak{C}\mathfrak{K}, \mathfrak{K}^\sharp\}}}, \frac{\bar{r}}{2} \right\}, \qquad \hat{\delta} := \frac{1}{32e\mathtt{N}}$$

and, for all $0 < r \le \widehat{\mathtt{r}}$, let us set

$$\varepsilon := \gamma^{-1} \left(\frac{2r}{\bar{r}}\right)^2 |G|_{\bar{r},\eta,\mathtt{w}_0} = \frac{1}{2^9 e} \left(\frac{r}{r_\star}\right)^2.$$

From definition (1.32) we thus deduce that

(5.18) $$8\,\varepsilon \max\{\mathfrak{C}\mathfrak{K}, \mathfrak{K}^\sharp\}\hat{\delta}^{-1} \le 1.$$

Recalling the notations introduced in (5.10)-(5.15), by Lemma (5.1) we have

$$\gamma^{-1}|G|_0 \le \varepsilon,$$

hence, setting $Z^{(0)} := \Pi_{\mathcal{K}}G$ and $R^{(0)} := \Pi_{\mathcal{R}}G$, from (4.7) it follows that

$$\gamma^{-1}|Z^{(0)}|_0, \ \gamma^{-1}|R^{(0)}|_0 \le \varepsilon.$$

We perform an iterative procedure producing a sequence of Hamiltonians, for $n = 0, \ldots, \mathtt{N}$

$$H^{(n)} = D_\omega + Z^{(n)} + R^{(n)},$$
$$Z^{(n)} \in \mathcal{K}_n, \quad R^{(n)} \in \mathcal{R}_n, \quad \mathtt{d}(Z^{(n)}) \ge 1, \quad \mathtt{d}(R^{(n)}) \ge n+1,$$
(5.19) $$\gamma^{-1}|Z^{(n)}|_n \le \varepsilon \sum_{h=0}^{n} 2^{-h}, \quad \gamma^{-1}|R^{(n)}|_n \le \varepsilon^{n+1} \left(4\mathfrak{C}\mathfrak{K}\hat{\delta}^{-1}\right)^n \overset{(5.18)}{\le} 2^{-n}\varepsilon.$$

Fix any $k < \mathtt{N}$. Let us assume that we have constructed $H^{(0)}, \ldots, H^{(k)}$ satisfying (5.19) for all $0 \le n \le k$. We want to apply Lemma 5.2 with

$$H, r, \eta, \mathtt{w} \ \rightsquigarrow \ H^{(k)}, r_k, \eta_k, \mathtt{w}_k \quad \text{and} \quad r', \eta', \mathtt{w}', \mathtt{w}^\sharp, \mathtt{d} \ \rightsquigarrow \ r_{k+1}, \eta_{k+1}, \mathtt{w}_{k+1}, \mathtt{w}_{\mathtt{N}}, k+1.$$

By construction the bounds (3.3), (4.8) and (5.4) hold since $C \le \mathfrak{C}$, $K \le \mathfrak{K}$, $K^\sharp \le \mathfrak{K}^\sharp$, where $\widehat{C}, \widehat{K}, \widehat{K}^\sharp$ were defined in (5.11),(5.12),(5.13). We just have to verify that (5.1) holds, namely

$$|R^{(k)}|_k \le \frac{\gamma}{\mathfrak{K}} \frac{r_k - r_{k+1}}{16er_k}.$$

In fact, by applying the inductive hypothesis (5.19) and the smallness condition (5.18), we get

$$|R^{(k)}|_k \le \gamma\left(4\mathfrak{C}\mathfrak{K}\hat{\delta}^{-1}\right)^k \varepsilon^{k+1} \le \frac{\gamma\varepsilon}{2^k} \le \frac{\gamma}{16e\mathfrak{K}(2\mathtt{N}-k)} = \frac{\gamma}{\mathfrak{K}} \frac{r_k - r_{k+1}}{16er_k}.$$

The verification of (5.5) is completely analogous.

So, by applying Lemma 5.2 we construct a change of variable Φ_k as in (5.2) with

$$\Phi_k \ : \ B_{r_{k+1}}(\mathtt{h}_{\mathtt{w}_{k+1}}) \ \to \ B_{r_k}(\mathtt{h}_{\mathtt{w}_{k+1}}).$$

Let us now set

$$H^{(k+1)} = D_\omega + Z^{(k+1)} + R^{(k+1)} := H_k \circ \Phi_k$$

with $Z^{(k+1)} \in \mathcal{K}_{k+1}, R^{(k+1)} \in \mathcal{R}_{k+1}$ and $\mathtt{d}(Z^{(k+1)}) \ge 1$, $\mathtt{d}(R^{(k+1)}) \ge k+2$. It remains to prove the bounds in the second line of (5.19) (with $n = k+1$). By (5.3) we have

$$
\begin{aligned}
|Z^{(k+1)}|_{k+1} &\leq |Z^{(k)}|_k + (\gamma\hat{\delta})^{-1}\mathfrak{K}|R^{(k)}|_k(\mathfrak{C}|R^{(k)}|_k + |Z^{(k)}|_k)\,, \\
|R^{(k+1)}|_{k+1} &\leq (\gamma\hat{\delta})^{-1}\mathfrak{K}|R^{(k)}|_k(\mathfrak{C}|R^{(k)}|_k + |Z^{(k)}|_k)\,.
\end{aligned}
\tag{5.20}
$$

By substituting the inductive hypothesis (5.19), we have the following chain of inequalities

$$
\begin{aligned}
\gamma^{-1}|R^{(k+1)}|_{k+1} &\leq \hat{\delta}^{-1}\varepsilon^2\mathfrak{K}(4\mathfrak{C}\mathfrak{K}\hat{\delta}^{-1}\varepsilon)^k(\mathfrak{C}(4\mathfrak{C}\mathfrak{K}\hat{\delta}^{-1}\varepsilon)^k + 2) \\
&\stackrel{(5.18)}{\leq} \hat{\delta}^{-1}\varepsilon^2\mathfrak{K}(4\mathfrak{C}\mathfrak{K}\hat{\delta}^{-1}\varepsilon)^k(\mathfrak{C} + 2) \\
&\leq (4\mathfrak{C}\mathfrak{K}\hat{\delta}^{-1})^{k+1}\varepsilon^{k+2} = (4\mathfrak{C}\mathfrak{K}\hat{\delta}^{-1}\varepsilon)^{k+1}\varepsilon,
\end{aligned}
$$

which proves the bound on $R^{(n)}$ in (5.19) for any n.
En passant, we note that

$$
\gamma\varepsilon\left(4\mathfrak{C}\mathfrak{K}\hat{\delta}^{-1}\varepsilon\right)^{\mathbb{N}} = \frac{\gamma}{2^9 er_\star^2}\left(\frac{\mathfrak{C}\mathfrak{K}\mathbb{N}}{4r_\star^2}\right)^{\mathbb{N}} r^{2(\mathbb{N}+1)}\,.
\tag{5.21}
$$

Finally, using the same strategy as above, we also get

$$
\gamma^{-1}|Z^{(k+1)}|_{k+1} \leq \varepsilon\left(\sum_{h=0}^{k} 2^{-h} + (4\mathfrak{C}\mathfrak{K}\hat{\delta}^{-1})^{k+1}\varepsilon^{k+1}\right) \stackrel{(5.18)}{\leq} \varepsilon\sum_{h=0}^{k+1} 2^{-h}\,,
$$

which completes the proof of the inductive hypothesis (5.19), and remark that

$$
\varepsilon\sum_{h=0}^{\mathbb{N}} 2^{-h} = \frac{r^2}{2^8 er_\star^2}\left(1 - 2^{-\mathbb{N}-1}\right).
\tag{5.22}
$$

By (5.6) we have

$$
\Phi_k \ : \ B_{r_{k+1}}(\mathtt{h}_{\mathtt{w}_{\mathbb{N}}}) \ \to \ B_{r_k}(\mathtt{h}_{\mathtt{w}_{\mathbb{N}}})\,,
$$
$$
\sup\nolimits_{u\in B_{r_{k+1}}(\mathtt{h}_{\mathtt{w}_{\mathbb{N}}})} |\Phi_k(u) - u|_{\mathtt{w}_{\mathbb{N}}} \leq r_k\gamma^{-1}\mathfrak{K}^\sharp|R^{(k)}|_k\,.
\tag{5.23}
$$

In conclusion we define

$$
\Psi := \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{\mathbb{N}-1} \ : \ B_r(\mathtt{h}_{\mathbb{N}}) \to B_{2r}(\mathtt{h}_{\mathbb{N}}).
$$

Since we have

$$
\begin{aligned}
&\Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{\mathbb{N}-1} - \mathrm{id} \\
&= (\Phi_0 - \mathrm{id}) \circ \Phi_1 \circ \cdots \circ \Phi_{\mathbb{N}-1} + (\Phi_1 - \mathrm{id}) \circ \Phi_2 \circ \cdots \circ \Phi_{\mathbb{N}-1} + \ldots \Phi_{\mathbb{N}-1} - \mathrm{id}\,.
\end{aligned}
$$

By (5.23) we get

$$
\sup_{u\in B_r(\mathtt{h}_{\mathtt{w}_{\mathbb{N}}})} |\Psi(u) - u|_{\mathtt{w}_{\mathbb{N}}} \leq \sum_{k=0}^{\mathbb{N}-1} r_k\gamma^{-1}\mathfrak{K}^\sharp|R^{(k)}|_k \stackrel{(5.19)}{\leq} 2r\varepsilon\mathfrak{K}^\sharp\sum_{k=0}^{\mathbb{N}-1} 2^{-k} \leq 4r\mathfrak{K}^\sharp\varepsilon\,,
$$

proving the first bound in (1.33). The second bound in (1.33) can be written as $8\widehat{C}_1 r^2 \leq 1$, which follows from $r \leq \widehat{\mathfrak{r}}$. We finally set $Z = Z_{\mathbb{N}}, R = R_{\mathbb{N}}$ and the estimates (1.34) follow by (5.21)-(5.22). Of course the same reasoniong can be applied in order to construct the inverse, i.e. a symplectic change of variables $\Phi : B_r(\mathtt{h}_{\mathtt{w}}) \mapsto B_{2r}(\mathtt{h}_{\mathtt{w}})$ such that

$$
\Psi \circ \Phi u = \Phi \circ \Psi u = u\,, \quad \forall u \in B_{\frac{7}{8}r}(\mathtt{h}_{\mathtt{w}})\,.
\tag{5.24}
$$

$\square$

When the nonlinearity G preserves momentum Theorem 1.3 can be reformulated under slightly weaker assumptions. More precisely, setting $\eta = 0$

$$\mathfrak{C}_0 := \max\left\{ 1, \sup_{0 \leq n < \mathtt{N}} \sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta}, \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0, \\ \pi(\boldsymbol{\alpha}-\boldsymbol{\beta})=0}} \frac{c^{(j)}_{\varrho_n^*,0,\mathtt{w}_{n+1}}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{\varrho_n,0,\mathtt{w}_n}(\boldsymbol{\alpha},\boldsymbol{\beta})} \right\} < \infty\,,$$

$$\mathfrak{K}_0 := \max\left\{ 1, \sup_{0 \leq n < \mathtt{N}} \sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta}, \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0, \\ \pi(\boldsymbol{\alpha}-\boldsymbol{\beta})=0}} \frac{c^{(j)}_{\varrho_n^*,0,\mathtt{w}_{n+1}}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{\varrho_n,0,\mathtt{w}_n}(\boldsymbol{\alpha},\boldsymbol{\beta})|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|} \right\} < \infty\,,$$

$$(5.25) \qquad \mathfrak{K}_0^\sharp := \max\left\{ 1, \sup_{0 \leq n < \mathtt{N}} \sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta}, \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0, \\ \pi(\boldsymbol{\alpha}-\boldsymbol{\beta})=0}} \frac{c^{(j)}_{\varrho_n^*,0,\mathtt{w}_{\mathtt{N}}}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{\varrho_n,0,\mathtt{w}_n}(\boldsymbol{\alpha},\boldsymbol{\beta})|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|} \right\} < \infty\,,$$

the following holds

Proposition 5.1. *If G preserves momentum Theorem 1.3 holds word by word with $\mathfrak{C}_0, \mathfrak{K}_0, \mathfrak{K}_0^\sharp$ instead of $\mathfrak{C}, \mathfrak{K}, \mathfrak{K}^\sharp$. Moreover also the new perturbation R preserves momentum.*

. We note that in the case that G preserves momentum, the same result holds with $\mathfrak{C}_0, \mathfrak{K}_0, \mathfrak{K}_0^\sharp$ instead of $\mathfrak{C}, \mathfrak{K}, \mathfrak{K}^\sharp$; moreover also R preserves momentum.

We finally give the following abstract stability result, whose proof is postponed to the Appendix B

Lemma 5.4. *On the Hilbert space $\mathtt{h}_{\mathtt{w}}$ consider the dynamical system*

$$\dot{v} = X_{\mathcal{N}} + X_R\,, \qquad v(0) = v_0\,, \qquad |v_0|_{\mathtt{w}} \leq \frac{3}{4} r\,,$$

where $\mathcal{N} \in \mathcal{A}_{r,0}(\mathtt{h}_{\mathtt{w}})$ and $R \in \mathcal{H}_{r,\eta}(\mathtt{h}_{\mathtt{w}})$ for some $r > 0, \eta \geq 0$. Assume that

$$\mathrm{Re}(X_{\mathcal{N}}, v)_{\mathtt{h}_{\mathtt{w}}} = 0\,.$$

Then

$$(5.26) \qquad \left| |v(t)|_{\mathtt{w}} - |v_0|_{\mathtt{w}} \right| < \frac{r}{8}\,, \qquad \forall\, |t| \leq \frac{1}{8|R|_{r,\eta,\mathtt{w}}}\,.$$

Corollary 5.1. *Under the same assumptions of Theorem 1.3, the solution $u(t)$ of the Hamiltonian flow of (1.31) with initial datum $u(0) = u_0$ such that $|u_0|_{\mathtt{w}} \leq \frac{3r}{8}$ exists and satisfies*

$$(5.27) \qquad |u(t)|_{\mathtt{w}} \leq r \quad \text{for all times} \quad |t| \leq \frac{1}{8\widehat{\mathfrak{C}}_3 r^{2(\mathtt{N}+1)}}\,.$$

Proof. Let us consider Hamiltonian (1.31), take an initial datum $|u_0|_{\mathtt{w}} := r < \frac{3}{8}\widehat{\mathtt{r}}$ and apply the change of variables of Theorem 1.3. Denoting by $v(0) = \Psi(u_0)$ we are under the hypotheses of Lemma 5.4 with $\eta = 0$ and we conclude

$$|v(t)|_{\mathtt{w}} \leq \frac{7}{8} r\,, \qquad \forall\, |t| \leq \frac{1}{8|R|_{r,0,\mathtt{w}}}\,.$$

Now we can apply (5.24) in order to return to the original variables and deduce that $u(t) = \Phi v(t)$ satisfies

$$|u(t)|_{\mathtt{w}} \leq r\,, \qquad \forall\, |t| \leq \frac{r^{-2(\mathtt{N}+1)}}{8\widehat{\mathfrak{C}}_3} \leq \frac{1}{8|R|_{r,0,\mathtt{w}}}\,.$$

$\square$

Part 2. Applications to Gevrey and Sobolev cases

In Part 2 we show how to apply the abstract BNF to Gevrey and Sobolev cases. Following the notations given in the introduction we work in the three sequence spaces defined for the applications G, S, M, see page 7. As explained in the introduction, in order to prove the estimates on the stability times we just need to verify that Assumption 1 holds. This is the content of the next sections.

Let us start by setting some notations.

Case G) In the case $\mathtt{w}(p, s, a) = \left(\langle j \rangle^p e^{s \langle j \rangle^\theta + a|j|} \right)_{j \in \mathbb{Z}}$ we denote $\mathtt{h}_{\mathtt{w}(p,s,a)} = \mathtt{h}_{p,s,a}$, same notation for the norm of vectors $|\cdot|_{p,s,a}$. Regarding the norm of Hamiltonians we write $|\cdot|_{r,\eta,\mathtt{w}(p,s,a)}$, consistently with Definition 1.2. Of course, for any $0 \le p \le p', 0 \le s \le s', 0 \le a \le a'$ we have

$$\mathtt{h}_{p',s',a'} \subseteq \mathtt{h}_{p,s,a}, \qquad |v|_{p,s,a} \le |v|_{p',s',a'}, \quad \forall v \in \mathtt{h}_{p',s',a'}.$$

Case S) If $a = s = 0$ we denote $\mathtt{h}_{p,0,0} = \mathtt{h}_p$, same notation for the norm of vectors $|\cdot|_p$ and hamiltonians $|\cdot|_{r,\eta,\mathtt{w}(p)}$.

Remark 5.1. Note that, via the usual Fourier identification one has:

$$(5.28) \qquad |u|_p \le |u(x)|_{L^2} + |\partial_x^p u(x)|_{L^2} \le 2|u|_p$$

Case M) In the case $\mathtt{w}_j = \lfloor j \rfloor^p$ where

$$\lfloor j \rfloor := \max\{|j|, 2\}$$

we denote the norm of vectors as

$$(5.29) \qquad \|u\|_p^2 = \|u\|_{\mathtt{w}}^2 := \sum_{j \in \mathbb{Z}} \lfloor j \rfloor^{2p} |u_j|^2.$$

Remark 5.2. Note that $\mathtt{h}_{\mathtt{w}}$ in M) and $\mathtt{h}_p$ are the same vector space endowed with two equivalent norms. Moreover one has

$$(5.30) \qquad \|u\|_p \le 2^p |u(x)|_{L^2} + |\partial_x^p u(x)|_{L^2} \le 2\|u\|_p.$$

Definition 5.2 (momentum preserving regular Hamiltonians). Given $r > 0, p \ge 0$ let $\mathcal{H}^{r,p}$ be the space of point-wise absolutely convergent Hamiltonians on $\|u\|_p \le r$ which preserves momentum and such that

$$(5.31) \qquad \|H\|_{r,p} := r^{-1} \left(\sup_{\|u\|_p \le r} \|X_{\underline{H}}\|_p \right) < \infty,$$

namely.[16]

$$\|\cdot\|_{r,p} = |\cdot|_{\mathcal{H}_{r,0}(\mathtt{h}_{\mathtt{w}})}, \qquad \mathtt{w}_j = \lfloor j \rfloor^p$$

We now verify that the nonlinearities in (1.1) are bounded in the norm $|\cdot|_{r,\eta,\mathtt{w}}$ in the cases S, M, G.

Proposition 5.2. *Consider the correction term $P = \int_{\mathbb{T}} F(x, |u|^2) dx$ in the NLS Hamiltonian (1.21), where the argument f in F satisfies*(1.2). *Let $p > 1/2$.*

    (i) For any $a, s, \eta \ge 0$ such that $a + \eta < \mathtt{a}$ and any $r > 0$ such that[17] *$(C_{\mathtt{alg}}(p)r)^2 \le R$, we have*

$$(5.32) \qquad |P|_{r,\eta,\mathtt{w}(p,s,a)} \le C_{\mathtt{Nem}}(p, s, \mathtt{a} - a - \eta) \frac{(C_{\mathtt{alg}}(p)r)^2}{R} |f|_{\mathtt{a},R} < \infty.$$

    where f and $|f|_{\mathtt{a},R}$ are defined in 1.2.

---

[16]Note that on the preserving momentum subspace $\mathcal{H}_{r,\eta}(\mathtt{h}_{\mathtt{w}})$ coincides with $\mathcal{H}_{r,0}(\mathtt{h}_{\mathtt{w}})$ for every η.

[17]R is defined in (1.2) and the constants in Appendix A.

(ii) If F is independent of[18] *x, for $(C_{\mathtt{alg},\mathtt{M}}(p)r)^2 \leq R$ we have*

$$(5.33) \qquad \|P\|_{r,p} \leq 2^p \frac{(C_{\mathtt{alg},\mathtt{M}}(p)r)^2}{R}|f|_R < \infty.$$

This Proposition follows directly from the fact that the corresponding sequence spaces $\mathtt{h}_{\mathtt{w}}$ are closed w.r.to convolution.

Let $\star : \mathtt{h}_{p,s,a} \times \mathtt{h}_{p,s,a} \to \mathtt{h}_{p,s,a}$ be the convolution operation defined as

$$(f,g) \mapsto f \star g := \left(\sum_{j_1,j_2 \in \mathbb{Z}, \, j_1+j_2=j} f_{j_1} g_{j_2} \right)_{j \in \mathbb{Z}}.$$

The map $\star : (f,g) \mapsto f \star g$ is continuous in the following sense:

Lemma 5.5. *For $p > 1/2$ we have*

$$(5.34) \qquad |f \star g|_{p,s,a} \leq C_{\mathtt{alg}}(p)|f|_{p,s,a}|g|_{p,s,a}, \qquad \|f \star g\|_p \leq C_{\mathtt{alg},\mathtt{M}}(p)\|f\|_p\|g\|_p.$$

The proof is given in Appendix B.

Proof of Proposition 5.2. By definition (recall (1.2) and (1.21))

$$(5.35) \qquad F(x,y) = \int_0^y f(x,s)ds = \sum_{d=2}^\infty \frac{f^{(d-1)}(x)}{d}y^d =: \sum_{d=2}^\infty F^{(d)}(x)y^d$$

therefore we have

$$P = \int_{\mathbb{T}} F(x,|u|^2)dx = \sum_{d \geq 2} \left(F^{(d)} \star \underbrace{u \star \cdots \star u}_{d \text{ times}} \star \underbrace{\bar{u} \star \cdots \star \bar{u}}_{d \text{ times}} \right)_0.$$

To each analytic function $F^{(d)}(x)$ we associate its Fourier coefficients; we have $\left(F_j^{(d)} \right)_{j \in \mathbb{Z}} \in \mathtt{h}_{p,s,a_0}$ for $a_0 := a + \eta < \mathtt{a}$ and $s, p \geq 0$. Indeed

$$|F^{(d)}|_{p,s,a_0}^2 := \sum_j e^{2a_0|j|+2s\langle j \rangle^\theta} \langle j \rangle^{2p} |F_j^{(d)}|^2 \overset{(5.35)}{=} \sum_j e^{2a_0|j|+2s\langle j \rangle^\theta} \langle j \rangle^{2p} \frac{|f_j^{(d-1)}|^2}{d^2}$$

$$\leq \frac{c^2(p,s,\mathtt{a}-a_0)}{d^2} \sum_j e^{2\mathtt{a}|j|} |f_j^{(d-1)}|^2 = \frac{c^2(\mathtt{a}-a_0,s,p)}{d^2} |f^{(d-1)}|_{\mathbb{T}_\mathtt{a}}^2$$

with

$$c(p,s,t) := e^s + \sup_{x \geq 1} x^p e^{-tx+sx^\theta}.$$

Now condition (1.2) ensures that (B.12) holds and our claim follows, by Lemma B.2, setting $a_0 = a + \eta$.
(ii) Follows from (B.14). $\qquad \square$

---

[18]i.e. P preserves momentum and we are assuming (1.3).

6. Immersions

The following proposition gathers the immersion properties of the norm $|\cdot|_{r,\eta,\mathtt{w}(p,s,a)}$ with respect to the parameters p, s, a.

Proposition 6.1. *The following inequalities hold:*

(1) *Variations w.r.t. the paramater p. For any $0 < \rho < r$, $0 < \sigma < \eta$ and $p_1 > 0$ we have*

$$|H|_{r-\rho,\eta-\sigma,\mathtt{w}(p+p_1,s,a)} \leq C_{\mathtt{mon}}(r/\rho,\sigma,p_1)|H|_{r,\eta,\mathtt{w}(p,s,a)} \,.$$

(2) *Variation w.r.t. the parameter s. For any $0 < \sigma < \eta$ we have*

(6.1) $$|H|_{r,\eta-\sigma,\mathtt{w}(p,s+\sigma,a)} \leq |H|_{r,\eta,\mathtt{w}(p,s,a)}$$

(3) *Variation w.r.t. the parameter a. For any $0 < \sigma < \eta$*

(6.2) $$|H|_{e^{-\sigma}r,\eta-\sigma,\mathtt{w}(p,s,a+\sigma)} \leq e^{2\sigma}|H|_{r,\eta,\mathtt{w}(p,s,a)}$$

Remark 6.1. All the items in the previous Proposition describe immersion properties of $\mathcal{H}_{r,\eta}(\mathtt{h}_{p,s,a})$ w.r.t variations of the parameters.

In item (1) we say that if $H \in \mathcal{H}_{r,\eta}(\mathtt{h}_{p,s,a})$ (i.e. if its vector field maps $B_r(\mathtt{h}_{p,s,a}) \to \mathtt{h}_{p,s,a}$) then it is also in $\mathcal{H}_{r-\rho,\eta-\sigma}(\mathtt{h}_{p+p_1,s,a})$ for any $\rho, \sigma, p_1 > 0$. Note however that the norm of H in the latter space is in general much larger, we denote this constant by $C_{\mathtt{mon}}$.

In item (3) we have essentially the same phenomenon, only in order to increase the analiticity parameter $a \rightsquigarrow a + \sigma$, we need to decrease the radius to $e^{-\sigma}r$.

Item (2) gives the best bound, indeed not only $\mathcal{H}_{r,\eta-\sigma}(\mathtt{h}_{p,s+\sigma,a}) \subseteq \mathcal{H}_{r,\eta}(\mathtt{h}_{p,s,a})$ but the norm of H in the latter space *does not increase.*

To prove this Proposition we show that the hypotheses of Proposition 3.1 hold. In order to prove this, in turn we strongly rely on some notation and results introduced by Bourgain in [Bou05] and extended later on by Cong-Li-Shi-Yuan in [CLSY] (Definition 6.1 and Lemma 6.1 below). The definitions and lemmata given below are the key technical arguments. Many of the ideas come from Bourgain in [Bou05] in the case of Gevrey regularity and for momentum preserving Hamiltonians, here we give a detailed presentation adapted to our more general setting and covering also the case of Sobolev regularity.

Definition 6.1. Given a vector $v = (v_i)_{i\in\mathbb{Z}}$ $v_i \in \mathbb{N}$, $|v| < \infty$ we denote by $\widehat{n} = \widehat{n}(v)$ the vector $(\widehat{n}_l)_{l\in I}$ (where $I \subset \mathbb{N}$ is finite) which is the decreasing rearrangement of

$$\{\mathbb{N} \ni h > 1 \quad \text{repeated } v_h + v_{-h} \text{ times}\} \cup \{1 \quad \text{repeated } v_1 + v_{-1} + v_0 \text{ times}\}$$

Remark 6.2. A good way of envisioning this list is as follows. Given $v = (v_i)_{i\in\mathbb{Z}}$ consider the monomial $x^v := \prod_i x_i^{v_i}$. We can write uniquely

$$x^v = \prod_i x_i^{v_i} = x_{j_1} x_{j_2} \cdots x_{j_{|v|}}$$

then $\widehat{n}(v)$ is the decreasing rearrangement of the list $(\langle j_1 \rangle, \ldots, \langle j_{|v|} \rangle)$.

As an example, consider the case $v \neq 0$. Then, by construction there exists a unique $J \geq 0$ such that $v_j = 0$ for all $|j| > J$ and $v_J + v_{-J} \neq 0$ hence

$$v = (\ldots, 0, v_{-J}, \ldots, v_0, \ldots, v_J, 0 \ldots) \,.$$

If $J = 0$ then

$$\widehat{n} = (\underbrace{1, \ldots, 1}_{v_0 \text{ times}})$$

otherwise we have

$$\widehat{n} = (\ \underbrace{J, \ldots, J}_{v_J + v_{-J} \text{ times}}, \underbrace{J-1, \ldots, J-1}_{v_{J-1} + v_{-J+1} \text{ times}}, \ldots, \ \underbrace{1, \ldots, 1}_{v_1 + v_{-1} + v_0 \text{ times}}\ )$$

Given $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}$ with $1 \le |\boldsymbol{\alpha}| = |\boldsymbol{\beta}| < \infty$, from now on we define

$$\widehat{n} = \widehat{n}(\boldsymbol{\alpha} + \boldsymbol{\beta})\,.$$

We set the *even* number

$$N := |\boldsymbol{\alpha}| + |\boldsymbol{\beta}|\,,$$

which is the cardinality of $\widehat{n}$. We observe that, given

$$\pi = \sum_{i \in \mathbb{Z}} i(\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i) = \sum_{h > 0} h\big(\boldsymbol{\alpha}_h - \boldsymbol{\beta}_h - \boldsymbol{\alpha}_{-h} + \boldsymbol{\beta}_{-h}\big)\,,$$

there exists a choice of $\sigma_i = \pm 1, 0$ such that

$$(6.3) \qquad\qquad \pi = \sum_l \sigma_l \widehat{n}_l.$$

with $\sigma_l \ne 0$ if $\widehat{n}_l \ne 1$. Hence,

$$(6.4) \qquad\qquad \widehat{n}_1 \le |\pi| + \sum_{l \ge 2} \widehat{n}_l.$$

Indeed, if $\sigma_1 = \pm 1$, the inequality follows directly from (6.3); if $\sigma_1 = 0$, then $\widehat{n}_1 = 1$ and consequently $\widehat{n}_l = 1\ \forall l$. Since the mass is conserved, the list $\widehat{n}$ has at least two elements, and the inequality is achieved.

Lemma 6.1. *Given $\boldsymbol{\alpha}, \boldsymbol{\beta}$ such that $\sum_i i(\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i) = \pi \in \mathbb{Z}$, we have that setting $\widehat{n} = \widehat{n}(\boldsymbol{\alpha} + \boldsymbol{\beta})$*

$$(6.5) \qquad \sum_i \langle i \rangle^\theta (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) = \sum_{l \ge 1} \widehat{n}_l^\theta \ge 2\widehat{n}_1^\theta + (2 - 2^\theta) \sum_{l \ge 3} \widehat{n}_l^\theta - \theta|\pi|.$$

Proof. In appendix C. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The lemma above was proved in the simpler case of momentum preserving Hamiltonians in [Bou05] for $\theta = \frac{1}{2}$ and for general θ in [CLSY]. It is fundamental in discussing the properties of $\mathcal{H}_{r,\eta}(\mathbf{h}_{p,s,a})$ with $s > 0$, indeed it implies

$$(6.6) \qquad \sum_i \langle i \rangle^\theta (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) - 2\langle j \rangle^\theta + |\pi(\boldsymbol{\alpha} - \boldsymbol{\beta})| \ge (1 - \theta)\left(\sum_{l \ge 3} \widehat{n}_l^\theta + |\pi|\right) \ge 0$$

for all $\boldsymbol{\alpha}, \boldsymbol{\beta}$ such that $\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \ne 0$.

Proof of Proposition 6.1. In all that follows we shall use systematically the fact that our Hamiltonians preserve the mass and are zero at the origin. These facts imply that $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}| \ge 1$.

Let us start by proving **Item** (2), which is the simplest case. We need to show that

$$(6.7) \qquad \frac{c_{r,\eta-\sigma,\mathtt{w}(p,s+\sigma,a)}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta})}{c_{r,\eta,\mathtt{w}(p,s,a)}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta})} = \exp(-\sigma(\sum_i \langle i \rangle^\theta (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) - 2\langle j \rangle^\theta + |\pi(\boldsymbol{\alpha} - \boldsymbol{\beta})|) \le 1\,.$$

The last inequality follows by (6.6) of Lemma 6.1

Item (1) First we assume that $\rho \le r/2$. By Proposition 3.1 for any $0 < \rho \le r/2$, $0 < \sigma < \eta$ and $p_1 > 0$ we need to compute

$$(6.8) \qquad C_{\mathtt{mon}} := \sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta} \\ \boldsymbol{\alpha}_j+\boldsymbol{\beta}_j\neq 0}} \frac{c^{(j)}_{r-\rho,\eta-\sigma,\mathtt{w}(p+p_1,s,a)}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{r,\eta,\mathtt{w}(p,s,a)}(\boldsymbol{\alpha},\boldsymbol{\beta})} = \sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta} \\ \boldsymbol{\alpha}_j+\boldsymbol{\beta}_j\neq 0}} \left(\frac{\langle j\rangle^2}{\prod_i \langle i\rangle^{\boldsymbol{\alpha}_i+\boldsymbol{\beta}_i}} \right)^{p_1} e^{-\sigma|\pi|} \left(\frac{r-\rho}{r} \right)^{|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|-2}.$$

We use the notations of Definition 6.1, with $\widehat{n}(\boldsymbol{\alpha}+\boldsymbol{\beta}) \equiv \widehat{n}$. Since $\boldsymbol{\alpha}_j+\boldsymbol{\beta}_j \neq 0$ we have that $\langle j\rangle \le \widehat{n}_1$. Note that

$$(6.9) \qquad \prod_i \langle i\rangle^{\boldsymbol{\alpha}_i+\boldsymbol{\beta}_i} = \prod_{l\ge 1} \widehat{n}_l .$$

Hence

$$\frac{\langle j\rangle^2}{\prod_i \langle i\rangle^{\boldsymbol{\alpha}_i+\boldsymbol{\beta}_i}} \le \frac{\widehat{n}_1}{\prod_{l\ge 2} \widehat{n}_l}$$

Let us call $N = |\boldsymbol{\alpha}| + |\boldsymbol{\beta}| \ge 2$. By (6.4) we have that

$$(6.10) \qquad \sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta} \\ \boldsymbol{\alpha}_j+\boldsymbol{\beta}_j\neq 0}} \frac{\langle j\rangle^2}{\prod_i \langle i\rangle^{\boldsymbol{\alpha}_i+\boldsymbol{\beta}_i}} \le \frac{\widehat{n}_1}{\prod_{l\ge 2} \widehat{n}_l} \le \frac{\sum_{l=2}^N \widehat{n}_l + |\pi|}{\prod_{l=2}^N \widehat{n}_l} \le \frac{(N-1)\widehat{n}_2 + |\pi|}{\prod_{l=2}^N \widehat{n}_l} \le \frac{N+|\pi|}{\prod_{l=3}^N \widehat{n}_l} .$$

We have shown that

$$\sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta} \\ \boldsymbol{\alpha}_j+\boldsymbol{\beta}_j\neq 0}} \frac{\langle j\rangle^2}{\prod_i \langle i\rangle^{\boldsymbol{\alpha}_i+\boldsymbol{\beta}_i}} \le N + |\pi| .$$

Since $(N + |\pi|)^{p_1} \le 2^{p_1}(N^{p_1} + |\pi|^{p_1})$, denoting $L := \ln (r/r - \rho)$ we repeatedly use Lemma C.1 in order to control

$$(6.11) \qquad \sup_{N\ge 2, \pi\in\mathbb{Z}} (N + |\pi|)^{p_1} e^{-\sigma|\pi|} \left(\frac{r-\rho}{r} \right)^{N-2}$$

$$\le 2^{p_1} \left(\sup_{N\ge 2, \pi\in\mathbb{Z}} N^{p_1} e^{-\sigma|\pi|-L(N-2)} + \sup_{N\ge 2, \pi\in\mathbb{Z}} |\pi|^{p_1} e^{-\sigma|\pi|-L(N-2)} \right)$$

$$\le 2^{p_1} \left(\max\left\{ \left(\frac{p_1}{L} \right)^{p_1}, 1 \right\} + \left(\frac{p_1}{\sigma} \right)^{p_1} \right) \le 2^{p_1+1} \max\left\{ \left(\frac{p_1}{L} \right)^{p_1}, \left(\frac{p_1}{\sigma} \right)^{p_1}, 1 \right\}$$

$$\le 2^{p_1+1} p_1^{p_1} \max\left\{ \left(\frac{2r}{\rho} \right)^{p_1}, \left(\frac{1}{\sigma} \right)^{p_1}, 1 \right\} = C_{\mathtt{mon}} ,$$

using that

$$L \ge \ln(1 + \rho/r) \ge 2\ln(3/2)\rho/r \ge \rho/2r ,$$

which holds since we are in the case $\rho \le r/2$. This completes the proof in the case $\rho \le r/2$.

Consider now the case $r/2 < \rho < r$. Using the monotonicity of the norm w.r.t. r and the already proved case with $\rho = r/2$, we have

$$|H|_{r-\rho,\eta-\sigma,\mathtt{w}(p+p_1,s,a)} \le |H|_{r/2,\eta-\sigma,\mathtt{w}(p+p_1,s,a)} \le 2^{p_1+1} \max\left\{ (4p_1)^{p_1}, \left(\frac{p_1}{e\sigma} \right)^{p_1}, 1 \right\} |H|_{r,\eta,\mathtt{w}(p,s,a)}$$

$$\le 2^{p_1+1} p_1^{p_1} \max\left\{ \left(\frac{2r}{\rho} \right)^{p_1}, \left(\frac{1}{\sigma} \right)^{p_1}, 1 \right\} |H|_{r,\eta,\mathtt{w}(p,s,a)} ,$$

proving (1) also in the case $r/2 < \rho < r$.

Item (3) We proceed as in item $(1) - (2)$,

$$(6.12) \qquad \frac{c^{(j)}_{e^{-\sigma}r,\eta-\sigma,\mathtt{w}(p,s,a+\sigma)}(\boldsymbol{\alpha},\boldsymbol{\beta})}{c^{(j)}_{r,\eta,\mathtt{w}(p,s,a)}(\boldsymbol{\alpha},\boldsymbol{\beta})} = \exp(-\sigma(\sum_i \langle i\rangle^\theta(\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) - 2\langle j\rangle^\theta + |\pi(\boldsymbol{\alpha} - \boldsymbol{\beta})| - (|\boldsymbol{\alpha}| + |\boldsymbol{\beta}| - 2)) \le e^{2\sigma} .$$

our claim follows since, by formula (6.4), one has

$$\sum_i (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i)|i| - 2|j| + |\pi| \geq \sum_{l \geq 2} \widehat{n}_l - \widehat{n}_1 + |\pi| - |\boldsymbol{\alpha}_0 + \boldsymbol{\beta}_0| \geq -(|\boldsymbol{\alpha}| + |\boldsymbol{\beta}|). \tag{6.13}$$

$\square$

Remark 6.3. Note that a key point in Items (1) and (2) are the estimates (6.10) and (6.6) where we control the ratio of the coefficient (1.27) in terms of $\{\widehat{n}_l\}_{l \geq 3}$ (namely uniformly with respect to $\widehat{n}_1$ and $\widehat{n}_2$). This means that if $\widehat{n}_3$ is "big", then the norm of the Hamiltonian is correspondingly small: polynomially in the Sobolev case and subexponentially in the Gevrey one. This is a seminal property which appears in different flavors throughout the literature; in Proposition 6.1 we do not really need to exploit it. Instead, it will be heavily used for a sharp control on the small divisors appearing in the Homological equation (see proof of Proposition 7.1).

Incidentally we note that norm $|\cdot|_{r,\eta,\mathtt{w}(p,s,a)}$ possesses the tameness property.

Proposition 6.2.

$$\sup_{|u|_{p_0,s,a} \leq r-\rho} \frac{|X_H|_{p,s,a}}{|u|_{p,s,a}} \leq C_{\mathtt{tame}}(\rho, \eta, p)|H|_{r,\eta,\mathtt{w}(p_0,s,a)}$$

Proof. In Appendix B. $\square$

Proposition 6.3. *The norm* $\|\cdot\|_{r,p}$ *is monotone decreasing in p, namely* $\|\cdot\|_{r,p+p_1} \leq \|\cdot\|_{r,p}$ *for any $p_1 > 0$.*

Proof. For the norm $\|\cdot\|_{r,p}$ the quantity in (1.27) becomes (recall that in the norm of a momentum preserving hamiltonian there is need of introducing the parameter η)

$$\mathsf{c}_{r,p}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) := r^{|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|-2} \left(\frac{\lfloor j \rfloor^2}{\prod_{i \in \mathbb{Z}} \lfloor i \rfloor^{(\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i)}} \right)^p. \tag{6.14}$$

By Lemma 3.1 item (ii) we only need to show that

$$c_{r,p+p_1}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq c_{r,p}^{(j)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \tag{6.15}$$

for all $j, \boldsymbol{\alpha}, \boldsymbol{\beta}$ with $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}| \geq 1$ and $\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \geq 1$ (recall the momentum conservation), namely we have to prove that

$$\sup_{\substack{j,\boldsymbol{\alpha},\boldsymbol{\beta} \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \geq 1}} \frac{\lfloor j \rfloor^2}{\prod_i \lfloor i \rfloor^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}} \leq 1. \tag{6.16}$$

We first show that the inequality holds in the case $j = 0, \pm 1$. Indeed we have

$$\prod_i \lfloor i \rfloor^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i} \geq \prod_i 2^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i} = 2^{\sum_i \boldsymbol{\alpha}_i + \boldsymbol{\beta}_i} \geq 4$$

since $\sum_i \boldsymbol{\alpha}_i + \boldsymbol{\beta}_i \geq 2$ (by the fact that $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}| \geq 1$).

Consider now the case $|j| = \lfloor j \rfloor \geq 2$. Since $\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \geq 1$, inequality (6.16) follows by

$$\sup_{j,\boldsymbol{\alpha},\boldsymbol{\beta}} \frac{|j|}{\prod_{i \neq j} \lfloor i \rfloor^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}} \leq 1. \tag{6.17}$$

By momentum conservation we have

$$(6.18) \qquad |j| \leq \sum_{i \neq j} |i|(\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) \leq \sum_{i \neq j} \lfloor i \rfloor (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i)$$

and (6.17) follows if we show that

$$(6.19) \qquad \sup_{j,\boldsymbol{\alpha},\boldsymbol{\beta}} \frac{\sum_{i \neq j} \lfloor i \rfloor (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i)}{\prod_{i \neq j} \lfloor i \rfloor^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}} \leq 1 \,,$$

where we can restrict the sum and the product to the indexes i such that $\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i \geq 1$. This last estimates follows by the fact that given $x_k \geq 1$

$$\frac{\sum_{2 \leq k \leq n} k x_k}{\prod_{2 \leq k \leq n} k^{x_k}} \leq 1 \,,$$

as it can be easly proved by induction over n (noting that $n^x \geq nx$ for $n \geq 2$, and any $x \geq 1$). $\qquad \square$

7. Homological equation

Now we give estimates on the solution of the homological equation

$$L_\omega S := \{D_\omega, S\} = R$$

The constants $\mathcal{C}_1, \mathcal{C}_2(r,\sigma,t)$ are defined in Appendix A. Note that $\mathcal{C}_1$ depends only on θ.

Proposition 7.1. *Let* $\omega \in \mathtt{D}_{\gamma,q}$ *and let* $0 < \sigma < \eta$, $0 < \rho < r/2$. *For any* $R \in \mathcal{R}_{r,\eta}(\mathtt{h}_{p,s,a})$, *the Homological equation* $L_\omega S = R$ *has a unique solution* $S = L_\omega^{-1} R$, *which satisfies the following two bounds:*

$$(\mathtt{G}) \qquad \left| L_\omega^{-1} R \right|_{r,\eta-\sigma,\mathtt{w}(p,s+\sigma,a)} \leq \gamma^{-1} e^{\mathcal{C}_1 \sigma^{-\frac{3}{\theta}}} |R|_{r,\eta,\mathtt{w}(p,s,a)}$$

$$(\mathtt{S}) \qquad \left| L_\omega^{-1} R \right|_{r-\rho,\eta-\sigma,\mathtt{w}(p+\tau,s,a)} \leq \gamma^{-1} \mathcal{C}_2(r/\rho,\sigma,\tau) |R|_{r,\eta,\mathtt{w}(p,s,a)}.$$

hence $L_\omega^{-1} R \in \mathcal{R}_{r,\eta-\sigma}(\mathtt{h}_{p,s+\sigma,a}) \cap \mathcal{R}_{r-\rho,\eta-\sigma}(\mathtt{h}_{p+\tau,s,a})$.

If R *preserves momentum* $R \in \mathcal{R}_{r,0}(\mathtt{h}_\mathtt{w})$, *with* $\mathtt{w}_j = \lfloor j \rfloor^p$, *the unique solution of the Homological equation preserves momentum and satisfies*

$$(\mathtt{M}) \qquad \|L_\omega^{-1} R\|_{r,p+\tau_1} \leq \gamma^{-1} 6^{\tau_1} (4^6 e^{27})^{2+q} \|R\|_{r,p},$$

so $S = L_\omega^{-1} R \in \mathcal{R}_{r,0}(\mathtt{h}_{\mathtt{w}'})$, *with* $\mathtt{w}'_j = \lfloor j \rfloor^{p+\tau_1}$.

Remark 7.1. As in the abstract case we assume that X_R maps $B_r(\mathtt{h}_{p,s,a}) \to \mathtt{h}_{p,s,a}$ and then show that S maps some smaller ball (because it has smaller radius or is in a stronger topology) to itself. This can be done in two ways: if we increase the Gevrey regularity index $s \rightsquigarrow s + \sigma$ (case $\mathtt{G}$) then the increase can be arbitrarily small, at the price of an exponential increase in the bound.

If we want to keep s fixed (say that we start with $s = a = 0$ and want to stay in the Sobolev class) then we have to increase the regularity p *by a fixed amount.* The main difference between the cases $\mathtt{S}$ and $\mathtt{M}$ is that in the first case one has to decrease r, η and the bound on S diverges as $\rho, \sigma \to 0$. In the second case, instead we have to increase the regularity p by a slightly larger amount but then we get a uniform bound for S.

Note that, differently from Proposition 6.1, we cannot consider the purely analytic case (s, p fixed say to $0, 1$). This is due to the fact that in (6.13) we have a much weaker bound for the ratio of the coefficients in (6.12), w.r.t. the one afforded by (6.6) and (6.10) for the Gevrey and Sobolev cases.

The following Lemma is the key point in the control of the small divisors appearing in the solution of the Homological equation. Here we strongly use the fact that we are working with a dispersive PDE on the circle with superlinear dispersion law.

Lemma 7.1. *Consider* $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}$ *with* $1 \leq |\boldsymbol{\alpha}| = |\boldsymbol{\beta}| < \infty$. *If*

$$(7.1) \qquad \left| \sum_i (\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i) i^2 \right| \leq 10 \sum_i |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i|,$$

then for all j *such that* $\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0$ *one has*

$$(7.2) \qquad \sum_i |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \langle i \rangle^{\theta/2} \leq C_* \left(\sum_i (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) \langle i \rangle^\theta - 2 \langle j \rangle^\theta + |\pi| \right), \qquad C_* = \frac{13}{1 - \theta}$$

$$(7.3) \qquad \prod_i (1 + |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \langle i \rangle) \leq e^{27} (1 + |\pi|)^3 N^6 \prod_{l=3}^N \widehat{n}_l^{\tau_0}.$$

where $N = |\boldsymbol{\alpha}| + |\boldsymbol{\beta}|$ *and* $\pi = \sum_i i(\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i)$ *(recall (1.25).*

Proof. In appendix C $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Note that

$$(7.4) \qquad \left| \sum_i (\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i) i^2 \right| \geq 10 \sum_i |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \qquad \implies \qquad |\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})| \geq 1.$$

Indeed denoting $\omega_j = j^2 + \xi_j \langle j \rangle^{-q}$ with $|\xi_j| \leq \frac{1}{2}$,

$$|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})| \geq 10 \sum_j |\boldsymbol{\alpha}_j - \boldsymbol{\beta}_j| - \frac{1}{2} \sum_j |\boldsymbol{\alpha}_j - \boldsymbol{\beta}_j| \geq 1.$$

Proof. In the following, we will compute for each item the corresponding K, K_0 defined in (4.8) and (4.10), and show their finiteness in order to apply Lemma 4.2 and give the explicit upper bounds entailed in Proposition 7.1 (G)-(S)-(M).

Item G) In this case by (6.7)

$$K = \gamma \sup_{j: \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0} \frac{e^{-\sigma \left(\sum_i \langle i \rangle^\theta (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) - 2 \langle j \rangle^\theta + |\pi| \right)}}{|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|}.$$

There are two cases.

If (7.1) does not hold, then by (7.4) $|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})| \geq 1$ and by (6.5) and (4.4) we get

$$\gamma \frac{e^{-\sigma \left(\sum_i \langle i \rangle^\theta (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) - 2 \langle j \rangle^\theta + |\pi| \right)}}{|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|} \leq 1$$

and the bound is trivially achieved.

Otherwise, let us consider the case in which (7.1) holds. By applying Lemma 7.1, since $\omega \in \mathsf{D}_{\gamma,q}$ we get:

$$
\gamma \frac{e^{-\sigma\left(\sum_i \langle i \rangle^\theta (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) - 2\langle j \rangle^\theta + |\pi|\right)}}{|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|}
$$

$$
\leq e^{-\frac{\sigma}{C_*} \sum_i |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \langle i \rangle^{\frac{\theta}{2}}} \prod_i \left(1 + (\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i)^2 \langle i \rangle^{2+q}\right)
$$

$$
\leq \exp \sum_i \left[-\frac{\sigma}{C_*} |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \langle i \rangle^{\frac{\theta}{2}} + \ln\left(1 + (\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i)^2 \langle i \rangle^{2+q}\right)\right]
$$

(7.5)
$$
= \exp \sum_i f_i(|\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i|)
$$

where, for $0 < \sigma \leq 1$, $i \in \mathbb{Z}$ and $x \geq 0$, we defined

$$
f_i(x) := -\frac{\sigma}{C_*} x \langle i \rangle^{\frac{\theta}{2}} + \ln\left(1 + x^2 \langle i \rangle^{2+q}\right).
$$

In order to bound (7.5), we need the following lemma, whose proof is postponed to Appendix C.

Lemma 7.2. *Setting*

$$
i_\sharp := \left(\frac{8C_*(q+3)}{\sigma\theta} \ln \frac{4C_*(q+3)}{\sigma\theta}\right)^{\frac{2}{\theta}},
$$

we get

(7.6)
$$
\sum_i f_i(|\ell_i|) \leq 7(q+3)i_\sharp \ln i_\sharp - \frac{\sigma}{2C_*}\left(\widehat{n}_1(\ell)\right)^{\frac{\theta}{2}}
$$

for every $\ell \in \mathbb{Z}^{\mathbb{Z}}$ with $|\ell| < \infty$.

The inequality (G) follows from plugging (7.6) into (7.5) and evaluating the constant.

Item S) In this case K in (4.8) is (recall (6.8))

(7.7)
$$
K = \gamma \sup_{j:\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0} \left(1 - \frac{\rho}{r}\right)^{N-2} \left(\frac{\langle j \rangle^2}{\prod_i \langle i \rangle^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}}\right)^\tau \frac{e^{-\sigma|\pi|}}{|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|},
$$

where $N = |\boldsymbol{\alpha}| + |\boldsymbol{\beta}|$.

As before we consider two cases.

If (7.1) is not satisfied then(7.4) holds and the right hand side of (7.7) is bounded by the quantity in (6.8) and it is estimated analogusly.

If (7.1) holds instead, by applying formula (6.10), Lemma 7.1 and the fact that $\omega \in \mathsf{D}_{\gamma,q}$ we get:

$$
\left(\frac{\langle j \rangle^2}{\prod_i \langle i \rangle^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}}\right)^\tau \frac{1}{|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|} \leq \left(\frac{\langle j \rangle^2}{\prod_i \langle i \rangle^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}}\right)^\tau \prod_i \left(1 + |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i|^2 \langle i \rangle^{2+q}\right)
$$

$$
\leq \left(\frac{N + |\pi|}{\prod_{l=3}^N \widehat{n}_l}\right)^\tau \left(e^{27}(1 + |\pi|)^3 N^6 \prod_{l \geq 3} \widehat{n}_l^{\tau_0}\right)^{2+q}
$$

$$
\leq e^{27(2+q)}(N + |\pi|)^{\tau + 9(2+q)} \leq e^{27(2+q)}(N + |\pi|)^{3\tau}.
$$

By using Lemma C.1 (just like explained in detail in formula (6.11) with $p_1 = 3\tau$), K in (7.7) is bounded by

$$
e^{27(2+q)}(N + |\pi|)^{3\tau}\left(1 - \frac{\rho}{r}\right)^{N-2} e^{-\sigma|\pi|}
$$

$$
\leq e^{27(2+q)} 2^{3\tau+1}(3\tau)^{3\tau} \max\left\{\left(\frac{2r}{\rho}\right)^{3\tau}, \left(\frac{1}{\sigma}\right)^{3\tau}, 1\right\}
$$

Item M) Note that in this case the constant in (4.8) amounts to

$$K_0 = \gamma \sup_{\substack{j \in \mathbb{Z}, \, \boldsymbol{\alpha} \neq \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}} \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \neq 0, \, \sum_i i(\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i) = 0}} \left(\frac{\lfloor j \rfloor^2}{\prod_i \lfloor i \rfloor^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}} \right)^{\tau_1} \frac{\gamma}{|\omega \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})|} \, .$$

We have two cases. If (7.4) holds $K_0 \leq \gamma$ by (6.16).

Otherwise (7.1) holds and, therefore, (7.3) (note that here $\pi = 0$) applies, giving

$$\begin{aligned}
K_0 &\leq \sup \left(\frac{\lfloor j \rfloor^2}{\prod_i \lfloor i \rfloor^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}} \right)^{\tau_1} \prod_i \left(1 + |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i|^2 \langle i \rangle^{2+q} \right) \\
&\leq \sup \left(\frac{\lfloor j \rfloor^2}{\prod_i \lfloor i \rfloor^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}} \right)^{\tau_1} e^{27(2+q)} N^{6(2+q)} \prod_{l=3}^N \widehat{n}_l^{\tau_0(2+q)}
\end{aligned}$$

since $\omega \in \mathtt{D}_{\gamma,q}$. We claim that

$$(7.8) \qquad\qquad N \leq 4 \prod_{l=3}^N \lfloor \widehat{n}_l \rfloor^{\frac{1}{4 \ln 2}} \, .$$

Indeed if $N = 2$, the inequality is trivial. Since N is even we have to consider only the case $N \geq 4$, which follows by Lemma C.1. Recalling (6.9) we have

$$(7.9) \qquad\qquad \prod_i \lfloor i \rfloor^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i} = \prod_{l \geq 1} \lfloor \widehat{n}_l \rfloor \, .$$

Then

$$\sup_{\substack{j, \boldsymbol{\alpha}, \boldsymbol{\beta} \\ \boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \geq 1}} \frac{\lfloor j \rfloor^2}{\prod_i \lfloor i \rfloor^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i}} \leq \frac{\lfloor \widehat{n}_1 \rfloor^2}{\prod_{l \geq 1} \lfloor \widehat{n}_l \rfloor} = \frac{\lfloor \widehat{n}_1 \rfloor}{\prod_{l \geq 2} \lfloor \widehat{n}_l \rfloor} \leq \frac{\sum_{l \geq 2} \lfloor \widehat{n}_l \rfloor}{\prod_{l \geq 2} \lfloor \widehat{n}_l \rfloor} = \frac{1}{\prod_{l \geq 3} \lfloor \widehat{n}_l \rfloor} + \frac{\sum_{l \geq 3} \lfloor \widehat{n}_l \rfloor}{\prod_{l \geq 2} \lfloor \widehat{n}_l \rfloor} \, ,$$

where the last inequality holds by momentum conservation. Then[19]

$$\begin{aligned}
K_0 &\leq 2^{\tau_1 - 1} \left(\frac{1}{\prod_{l \geq 3} \lfloor \widehat{n}_l \rfloor^{\tau_1}} + \frac{(\sum_{l \geq 3} \lfloor \widehat{n}_l \rfloor)^{\tau_1}}{\prod_{l \geq 2} \lfloor \widehat{n}_l \rfloor^{\tau_1}} \right) (4^6 e^{27})^{2+q} \prod_{l \geq 3} \lfloor \widehat{n}_l \rfloor^{\tau_1/2} \\
&\leq 2^{\tau_1 - 1} (4^6 e^{27})^{2+q} \left(1 + \frac{(\sum_{l \geq 3} \lfloor \widehat{n}_l \rfloor)^{\tau_1}}{\lfloor \widehat{n}_2 \rfloor^{\tau_1} \prod_{l \geq 3} \lfloor \widehat{n}_l \rfloor^{\tau_1/2}} \right) \\
&\leq 2^{\tau_1 - 1} (4^6 e^{27})^{2+q} \left(1 + \frac{(\lfloor \widehat{n}_3 \rfloor^{1/2} + 4)^{\tau_1}}{\lfloor \widehat{n}_2 \rfloor^{\tau_1}} \right)
\end{aligned}$$

by Lemma C.2 with $a = 1/2$. The estimate on K_0, hence inequality (M) follows. $\square$

8. Birkhoff normal form

We are now ready to apply Theorem 1.3 to the three applications $\mathtt{G}, \mathtt{S}, \mathtt{M}$, defined in page 7. We start by verifying the assumptions.

Lemma 8.1. *The following holds*

$\mathtt{G}$) *Let $s > 0$, $p > 1/2$ and $a \geq 0$. Then for all $\mathtt{N} \geq 1$, $0 < \eta \leq s$, $\mathtt{w} := \mathtt{w}(p, s, a)$ and $\mathtt{w}_0 := \mathtt{w}(p, s - \eta, a)$ satisfy the Birkhoff assumption at step $\mathtt{N}$ and in (1.29) we can take*

$$\mathfrak{C} = 1, \quad \mathfrak{K}, \, \mathfrak{K}^\sharp \leq e^{C_1 \left(\frac{\mathtt{N}}{\eta} \right)^{\frac{3}{\theta}}} \, .$$

---

[19]Using that $(a + b)^{\tau_1} \leq 2^{\tau_1 - 1}(a^{\tau_1} + b^{\tau_1})$ for $a, b \geq 0$, $\tau_1 \geq 1$.

S$)$ *Let $\tau_{\mathtt{S}} = \tau$, $s, a \geq 0$, $p \geq 3\tau_{\mathtt{S}} + 1$ and set*[20] $\mathtt{N} := [\frac{p-1}{\tau_{\mathtt{S}}}]$. *Then $\eta > 0$, $\mathtt{w} := \mathtt{w}(p, s, a)$ and*[21] $\mathtt{w}_0 := \mathtt{w}(p - \mathtt{N}\tau_{\mathtt{S}}, s, a)$ *satisfy the Birkhoff assumption at step $\mathtt{N}$ and in (1.29) we can take*

$$\mathfrak{C} \leq C_{\mathtt{mon}}(4\mathtt{N}, \eta/\mathtt{N}, \tau_{\mathtt{S}}), \quad \mathfrak{K} \leq \mathcal{C}_2(4\mathtt{N}, \eta/\mathtt{N}, \tau_{\mathtt{S}}), \quad \mathfrak{K}^{\sharp} \leq \mathcal{C}_2(4\mathtt{N}, \eta/\mathtt{N}, \mathtt{N}\tau_{\mathtt{S}}).$$

M$)$ *Let $\tau_{\mathtt{M}} = \tau_1$, $p \geq 3\tau_{\mathtt{M}} + 1$ and*[22] *set $\mathtt{N} := [\frac{p-1}{\tau_{\mathtt{M}}}]$. Then $\eta = 0$, $\mathtt{w} := (\lfloor j \rfloor^p)_{j \in \mathbb{Z}}$ and $\mathtt{w}_0 := (\lfloor j \rfloor^{p-\mathtt{N}\tau_{\mathtt{M}}})_{j \in \mathbb{Z}}$ satisfy the "momentum preserving" Birkhoff assumption at step $\mathtt{N}$ and in (5.25) we can take*

$$\mathfrak{C}_0 = 1, \quad \mathfrak{K}_0, \mathfrak{K}_0^{\sharp} \leq 6^{\tau_{\mathtt{M}}}(4^6 e^{27})^{2+q}.$$

Proof. G$)$ Set

$$\mathtt{w}_{n,j} := \mathtt{w}_{0,j} e^{\frac{n\eta}{\mathtt{N}} \langle j \rangle^{\theta}}, \qquad \forall n = 1, \ldots, \mathtt{N}.$$

The computation of $\mathfrak{C}$ follows from (6.1); the ones of $\mathfrak{K}, \mathfrak{K}^{\sharp}$ from Proposition 7.1.
S$)$ Set

$$\mathtt{w}_{n,j} := \mathtt{w}_{0,j} \langle j \rangle^{n\tau_{\mathtt{S}}}, \qquad \forall n = 1, \ldots, \mathtt{N}.$$

The computation of $\mathfrak{C}$ follows from (6.8); the ones of $\mathfrak{K}, \mathfrak{K}^{\sharp}$ from Proposition 7.1.
M$)$ Set

$$\mathtt{w}_{n,j} := \mathtt{w}_{0,j} \lfloor j \rfloor^{n\tau_{\mathtt{M}}}, \qquad \forall n = 1, \ldots, \mathtt{N}.$$

The computation of $\mathfrak{C}_0$ follows from (6.15); the ones of $\mathfrak{K}_0, \mathfrak{K}_0^{\sharp}$ again from Proposition 7.1. $\qquad\square$

We now state the Birkhoff Normal Form Theorem (1.3) for the Hamiltonian in (1.21) in the usual three cases. First we define

$$\mathtt{r}(\mathtt{G}) := \min\left\{ \frac{\delta_{\mathtt{G}}}{\sqrt{\mathtt{N}} e^{\frac{1}{2}\mathcal{C}_1\left(\frac{\mathtt{N}}{\eta_{\mathtt{G}}}\right)^{\frac{3}{\theta}}}}, \frac{\sqrt{R}}{2C_{\mathtt{alg}}(p)} \right\}, \quad \delta_{\mathtt{G}} := \frac{\sqrt{\gamma R}}{C_{\mathtt{alg}}(p)\sqrt{2^{11} e C_{\mathtt{Nem}}(p, s - \eta_{\mathtt{G}}, \mathtt{a} - a - \eta_{\mathtt{G}})|f|_{\mathtt{a}, R}}}$$

$$(8.1) \quad \mathtt{C}_1(\mathtt{G}) := \frac{e^{\mathcal{C}_1\left(\frac{\mathtt{N}}{\eta_{\mathtt{G}}}\right)^{\frac{3}{\theta}}}}{2^7 e \delta_{\mathtt{G}}^2}, \qquad \mathtt{C}_2(\mathtt{G}) := \frac{\gamma}{2^8 e \delta_{\mathtt{G}}^2}, \qquad \mathtt{C}_3(\mathtt{G}) := \frac{\gamma}{2^9 e \delta_{\mathtt{G}}^2}\left(\frac{\mathtt{N}e^{\mathcal{C}_1\left(\frac{\mathtt{N}}{\eta_{\mathtt{G}}}\right)^{\frac{3}{\theta}}}}{4\delta_{\mathtt{G}}^2}\right)^{\mathtt{N}}.$$

$$\mathtt{r}(\mathtt{S}) := \min\left\{ \frac{\mathtt{d}_{\mathtt{S}}}{\sqrt{\mathtt{N}\mathtt{C}_2(4\mathtt{N}, \mathtt{a}/2\mathtt{N}, \mathtt{N}\tau_{\mathtt{S}})}}, \frac{\sqrt{R}}{5 \cdot 2^{\tau_{\mathtt{S}}+2}} \right\}, \quad \text{where} \quad \mathtt{d}_{\mathtt{S}} := \frac{\sqrt{\gamma R}}{\sqrt{2^{17} C_{\mathtt{Nem}}(p - \mathtt{N}\tau_{\mathtt{S}}, 0, \mathtt{a}/2)|f|_{\mathtt{a}, R}}},$$

$$\mathtt{C}_1(\mathtt{S}) := \frac{\mathcal{C}_2(4\mathtt{N}, \mathtt{a}/2\mathtt{N}, \mathtt{N}\tau_{\mathtt{S}})}{2^7 e \mathtt{d}_{\mathtt{S}}^2}, \qquad \mathtt{C}_2(\mathtt{S}) := \frac{\gamma}{2^8 e \mathtt{d}_{\mathtt{S}}^2},$$

$$(8.2) \quad \mathtt{C}_3(\mathtt{S}) := \frac{2^8 C_{\mathtt{Nem}}(p - \mathtt{N}\tau_{\mathtt{S}}, 0, \mathtt{a}/2)|f|_{\mathtt{a}, R}}{eR}\left(\frac{\mathtt{N}C_{\mathtt{mon}}(4\mathtt{N}, \mathtt{a}/2\mathtt{N}, \tau_{\mathtt{S}})\mathcal{C}_2(4\mathtt{N}, \mathtt{a}/2\mathtt{N}, \tau_{\mathtt{S}})}{4\mathtt{d}_{\mathtt{S}}^2}\right)^{\mathtt{N}},$$

$$\mathtt{r}(\mathtt{M}) := \min\left\{ \frac{\delta_{\mathtt{M}}}{\sqrt{\mathtt{N}}}, \sqrt{\frac{R}{2^{\tau_{\mathtt{M}}+6}}} \right\}, \qquad \delta_{\mathtt{M}} := \frac{\sqrt{\gamma R}}{\sqrt{2^{17} e 12^{\tau_{\mathtt{M}}}(4^6 e^{27})^{2+q}|f|_R}},$$

$$\mathtt{C}_1(\mathtt{M}) := \frac{1}{2^9 \delta_{\mathtt{M}}^2}, \qquad \mathtt{C}_2(\mathtt{M}) := \frac{2^{\tau_{\mathtt{M}}+8}|f|_R}{R}, \qquad \mathtt{C}_3(\mathtt{M}) := 2^{\tau_{\mathtt{M}}+8}\frac{|f|_R}{R}\left(\frac{\mathtt{N}}{8\delta_{\mathtt{M}}^2}\right)^{\mathtt{N}},$$

---

[20]$[\cdot]$ is the integer part.
[21]Note that $1 \leq p - \mathtt{N}\tau_{\mathtt{S}} < 1 + \tau_{\mathtt{S}}$.
[22]Note that $1 \leq p - \mathtt{N}\tau_{\mathtt{M}} < 1 + \tau_{\mathtt{M}}$.

Theorem 8.1 (Birkhoff Normal Form). *Under the same assumptions of Lemma 8.1 the following holds. Consider the Hamiltonian* (1.21), *assuming, only in the case* M, *that f does not depend on x (momentum conservation). Then for any $0 < r \le \mathtt{r}$ there exists two close to identity invertible symplectic change of variables*

$$\Psi, \Psi^{-1} : \quad B_r(\mathtt{h_w}) \mapsto \mathtt{h_w}, \quad \sup_{|u|_{\mathtt{w}} \le r} |\Psi^{\pm 1}(u) - u|_{\mathtt{w}} \le \mathtt{C}_1 r^3 \le \tfrac{1}{8} r,$$

(8.3)
$$\Psi \circ \Psi^{-1} u = \Psi^{-1} \circ \Psi u = u, \quad \forall u \in B_{\frac{7}{8}r}(\mathtt{h_w})$$

such that in the new coordinates

$$H \circ \Psi = D_\omega + Z + R,$$

for suitable majorant analytic Hamiltonians $Z, R \in \mathcal{A}_r(\mathtt{h_w})$, $Z \in \mathcal{K}$, satisfying the estimate

(8.4)
$$\sup_{|u|_{\mathtt{w}} \le r} |X_{\underline{Z}}|_{\mathtt{w}} \le \mathtt{C}_2 r^3, \quad \sup_{|u|_{\mathtt{w}} \le r} |X_{\underline{R}}|_{\mathtt{w}} \le \mathtt{C}_3 r^{2\mathtt{N}+3},$$

$X_{\underline{Z}}$ (resp. $X_{\underline{R}}$), being the hamiltonian vector field generated by the the majorant of Z (resp. R). Moreover, in the case M, *R preserves momentum.*

Proof. We use Theorem 1.3 with $G \rightsquigarrow P$.

G) Setting

(8.5)
$$\eta = \eta_{\mathtt{G}} := \min\left\{ \frac{\mathtt{a} - a}{2}, s \right\}, \quad \bar{r} := \frac{\sqrt{R}}{C_{\mathtt{alg}}(p)},$$

we have that

(8.6)
$$|P|_{\bar{r}, \eta, \mathtt{w}_0} = |P|_{\bar{r}, \eta, \mathtt{w}(p, s - \eta, a)} \overset{(5.32)}{\le} C_{\mathtt{Nem}}(p, s - \eta, \mathtt{a} - a - \eta) |f|_{\mathtt{a}, R}$$

By (1.32) $r_\star \ge \delta_{\mathtt{G}} \ge \boldsymbol{\delta}_{\mathtt{G}}$ (see Appendix A). Then, recalling (1.32) and Lemma 8.1, one can verify that

$$\hat{\mathtt{r}} \ge \mathtt{r}(\mathtt{G}), \quad \widehat{\mathtt{C}}_1 \le \mathtt{C}_1(\mathtt{G}), \quad \widehat{\mathtt{C}}_2 \le \mathtt{C}_2(\mathtt{G}), \quad \widehat{\mathtt{C}}_3 \le \mathtt{C}_3(\mathtt{G}), .$$

S) Set

(8.7)
$$\eta := \mathtt{a}/2, \quad \bar{r} := \frac{\sqrt{R}}{C_{\mathtt{alg}}(p - \mathtt{N}\tau_{\mathtt{S}})}.$$

Then Assumption 1 is satisfied by Lemma 8.1 with the same choice of $\mathtt{N}, \mathtt{w}_0, \mathtt{w}$. We have that

(8.8)
$$|P|_{\bar{r}, \eta, \mathtt{w}_0} = |P|_{\bar{r}, \mathtt{a}/2, \mathtt{w}(p - \mathtt{N}\tau_{\mathtt{S}})} \overset{(5.32)}{\le} C_{\mathtt{Nem}}(p - \mathtt{N}\tau_{\mathtt{S}}, 0, \mathtt{a}/2) |f|_{\mathtt{a}, R}$$

For the various constants we refer to Appendix A. Recalling $\tau_{\mathtt{S}} = \tau$, we note that

$$C_{\mathtt{mon}}(4\mathtt{N}, \eta/\mathtt{N}, \tau) = 2^{2\tau+1} \tau^\tau \mathtt{N}^\tau \max\left\{4, (1/2\eta)\right\}^\tau = 2(4\tau \max\left\{4, (1/2\eta)\right\} \mathtt{N})^\tau,$$
$$\mathcal{C}_2(4\mathtt{N}, \eta/\mathtt{N}, \tau) = 2e^{27(2+q)} (12\tau \max\left\{4, (1/2\eta)\right\} \mathtt{N})^{3\tau}$$
$$\mathcal{C}_2(4\mathtt{N}, \eta/\mathtt{N}, \mathtt{N}\tau) = 2e^{27(2+q)} (6\mathtt{N}\tau)^{3\mathtt{N}\tau} \max\left\{(8\mathtt{N})^{3\mathtt{N}\tau}, (\mathtt{N}/\eta)^{3\mathtt{N}\tau}\right\}$$

(8.9)
$$= 2e^{27(2+q)} (12\tau \max\left\{4, (2\eta)^{-1}\right\} \mathtt{N}^2)^{3\mathtt{N}\tau},$$

we have that for $\mathtt{N} \ge 3$

(8.10)
$$C_{\mathtt{mon}}(4\mathtt{N}, \eta/\mathtt{N}, \tau) \mathcal{C}_2(4\mathtt{N}, \eta/\mathtt{N}, \tau) \le \sqrt{\mathcal{C}_2(4\mathtt{N}, \eta/\mathtt{N}, \mathtt{N}\tau)}.$$

By (1.32)

$$r_\star \ge \frac{\sqrt{\gamma R}}{C_{\mathtt{alg}}(p - \mathtt{N}\tau_{\mathtt{S}})\sqrt{2^{11} e C_{\mathtt{Nem}}(p - \mathtt{N}\tau_{\mathtt{S}}, 0, \mathtt{a}/2) |f|_{\mathtt{a}, R}}} \ge \mathtt{d}_{\mathtt{S}}.$$

Then, recalling (1.32) and (8.10) one has $\hat{\mathtt{r}} \ge \mathtt{r}(\mathtt{S})$. Moreover (recall (1.34))

(8.11)
$$\widehat{\mathtt{C}}_1 \le \mathtt{C}_1(\mathtt{S}), \quad \widehat{\mathtt{C}}_2 \le \mathtt{C}_2(\mathtt{S}), \quad \widehat{\mathtt{C}}_3 \le \mathtt{C}_3(\mathtt{S}).$$

Finally the last inequality in (8.3) follows from the second bound in (1.33).

M) Set

$$(8.12) \qquad \eta := 0, \quad \bar{r} := \frac{\sqrt{R}}{2^{\tau_{\mathtt{M}}/2} C_{\mathtt{alg},\mathtt{M}} (p - \tau_{\mathtt{M}} \mathtt{N})} \geq \sqrt{\frac{R}{10 \cdot 2^{\tau_{\mathtt{M}}}}}.$$

Then Assumption 1 is satisfied by Lemma 8.1, case M, with the same choice of $\mathtt{N}, \mathtt{w}_0, \mathtt{w}$. We have that

$$(8.13) \qquad |P|_{\bar{r},0,\mathtt{w}_0} = \|P\|_{\bar{r},p-\tau_{\mathtt{M}}\mathtt{N}} \overset{(5.33)}{\leq} 2|f|_R$$

By (1.32)

$$r_\star \geq \frac{\sqrt{\gamma R}}{\sqrt{2^{\tau_{\mathtt{M}}+17}|f|_R}}$$

Then

$$\widehat{\mathtt{r}} \geq \mathtt{r}(\mathtt{M}), \qquad \widehat{\mathtt{C}}_1 \leq \mathtt{C}_1(\mathtt{M}), \qquad \widehat{\mathtt{C}}_2 \leq \mathtt{C}_2(\mathtt{M}), \qquad \widehat{\mathtt{C}}_3 \leq \mathtt{C}_3(\mathtt{M}).$$

$\square$

9. Gevrey stability. Proof of Theorem 1.1

Actually we prove of Theorem 1.1 for the slightly longer stability time $|t| \leq \frac{2^4 e \delta_{\mathtt{G}}^2}{\gamma \delta^2} e^{\left(\ln \frac{\delta_{\mathtt{G}}}{\delta}\right)^{1+\theta/4}}$, where $\delta_{\mathtt{G}} > \boldsymbol{\delta}_{\mathtt{G}}$ (recall Appendix A). We set

$$r := 2\delta$$

and choose

$$(9.1) \qquad \mathtt{N}(r) := \left[\left(2 \ln \frac{2\delta_{\mathtt{G}}}{r}\right)^{\theta/4}\right] = \left[\left(2 \ln \frac{\delta_{\mathtt{G}}}{\delta}\right)^{\theta/4}\right].$$

Recalling (8.5) by Corollary 5.2 solutions of the PDE (1.1) in the space $\mathtt{h}_{p,s,a}$, correspond, by Fourier identification (1.19), to orbits of the Hamiltonian System (1.21) in the space

$$\mathtt{h}_{\mathtt{w}} \qquad \text{with} \qquad \mathtt{w}_j = e^{a|j| + s\langle j \rangle^\theta} \langle j \rangle^p.$$

An initial datum u_0 satisfying $|u_0|_{p,s,a} \leq \delta$ corresponds to[23] $u_0 \in \mathtt{h}_{\mathtt{w}}$ with $|u_0|_{\mathtt{w}} \leq \delta$.

We claim that $r \leq 2\delta_{\mathtt{G}}$ implies

$$(9.2) \qquad \frac{r \mathtt{N} e^{\mathcal{C}_1 \left(\frac{\mathtt{N}}{\eta_{\mathtt{G}}}\right)^{\frac{3}{\theta}}}}{2\delta_{\mathtt{G}}} \leq 1.$$

Indeed we have

$$\mathtt{N}(r) \geq \mathtt{N}_{\mathtt{G}} := \max\left\{\frac{16(4\mathcal{C}_1)^\theta}{\eta_{\mathtt{G}}^3}, 2^{\frac{2\theta+4}{4-\theta}}\right\}$$

and by (9.1) $r \leq 2\delta_{\mathtt{G}} e^{-\frac{1}{2}(\mathtt{N}(r)/2)^{4/\theta}}$ and (9.2) follows if we show that the function

$$\mathtt{N} \to e^{-\frac{1}{2}(\mathtt{N}/2)^{4/\theta}} \mathtt{N} e^{\mathcal{C}_1 \left(\frac{\mathtt{N}}{\eta_{\mathtt{G}}}\right)^{\frac{3}{\theta}}}$$

is ≤ 1 for $\mathtt{N} \geq \mathtt{N}_{\mathtt{G}}$. This is true since the function is decreasing for $\mathtt{N} \geq \mathtt{N}_{\mathtt{G}}$ and is ≤ 1 for $\mathtt{N} = \mathtt{N}_{\mathtt{G}}$. This proves the claim (9.2).

Then we apply Theorem 8.1 in the case $\mathtt{G}$. Recalling (8.1), by (8.4) and (9.2)

$$\mathtt{C}_3(\mathtt{G}) r^{2(\mathtt{N}+1)} \leq \frac{\gamma r^2}{2^9 e \delta_{\mathtt{G}}^2} \left(\frac{r}{2\delta_{\mathtt{G}}}\right)^{\mathtt{N}(r)} = \frac{\gamma \delta^2}{2^7 e \delta_{\mathtt{G}}^2} \left(\frac{\delta}{\delta_{\mathtt{G}}}\right)^{\mathtt{N}(r)} \leq \frac{\gamma \delta^2}{2^7 e \delta_{\mathtt{G}}^2} e^{-\left(\ln \frac{\delta_{\mathtt{G}}}{\delta}\right)^{1+\theta/4}},$$

since $\mathtt{N}(r) \geq \left(\ln \frac{2\delta_{\mathtt{G}}}{r}\right)^{\theta/4} = \left(\ln \frac{\delta_{\mathtt{G}}}{\delta}\right)^{\theta/4}$. We deduce the stability time by applying Lemma 5.1.

---

[23]We still denote it by u_0.

10. Sobolev stability

Before proving Proposition 1.1 we add a comment on the optimality of condition (1.10).

Remark 10.1. We construct a finite dimensional Hamiltonian, which is a reduction of (1.1) to a finite number of Fourier indices and which exhibits fast drift in a time of order 1. For instance, consider

$$H(u_1, u_j) := (1 + V_1)|u_1|^2 + (j^2 + V_j)|u_j|^2 + e^{-\mathtt{a}j} \operatorname{Re}(|u_1|^2 u_1 \bar{u}_j),$$

which is a finite dimensional model for (1.1) with $f(x, |u|^2) = e^{-\mathtt{a}j} \cos((j-1)x)|u|^2$. Consider now the initial datum $u(0) = (u_1(0), u_j(0)) = (\delta/4, |j|^{-p}\delta/4)$, which clearly has H_p norm $< \delta$. A direct computation shows that in a time T of order 1, the Sobolev norm of $u(T)$ is of order $\delta^3 e^{-\mathtt{a}j} j^p$ hence greater than 4δ if $\delta^2 e^{-\mathtt{a}j} j^p$ is large. Maximizing on j we get a constraint of the form $\delta^2 e^{-p}(\mathtt{a}^{-1}p)^p < 1$.

Of course this pathological "fast diffusion" phenomenon comes from the fact that f is NOT traslation invariant (and hence H does not preserve momentum). Actually, restricting to translation invariant Hamiltonians would not result in signficantly weaker constraints on the smallness of δ w.r.t. p. This can be seen in the following example. Consider the familiy of Hamiltonians (in three degrees of freedom)

$$K^{(j)} := V_0|u_0|^2 + (1 + V_1)|u_1|^2 + (j^2 + V_j)|u_j|^2 + \operatorname{Re}(\bar{u}_0^{j-1} u_1^j \bar{u}_j)$$

with the constants of motion

$$L = |u_0|^2 + |u_1|^2 + |u_j|^2, \quad M = |u_1|^2 + j|u_j|^2.$$

Following the same approach as in the previous example one shows that $|u_j|^2$ can have a drift of order $j^{-p}\delta^{2j}$ in a time T of order 1. This means that the Sobolev norm of $u(T)$ is of order $\delta^{2j} j^p$. Maximizing on j we get a constraint of the form $\delta e^{p^{1^-}} < 1$.

Proof of Proposition 1.1. As before we set $r := 2\delta$. An initial datum u_0 satisfying $|u_0|_{L^2} + |\partial_x^p u_0|_{L^2} \le \delta$ corresponds to[24] $u_0 \in \mathtt{h}_{\mathtt{w}(p)}$ with $|u_0|_p \le \delta$ by (5.28). We apply the Birkhoff Normal Form Theorem 8.1 in the case $\mathtt{S}$ (recall that $\mathtt{N} = \left[\frac{p-1}{\tau_{\mathtt{S}}}\right]$). Recalling the definition of $\mathtt{r}(\mathtt{S})$ in (8.2), we verify that, for any $\mathtt{N} \ge 1$

$$(10.1) \qquad \delta_{\mathtt{S}}(\mathtt{k}_{\mathtt{S}}p)^{-3p} \le \frac{\mathtt{d}_{\mathtt{S}}}{2\sqrt{\mathtt{N}\mathcal{C}_2(4\mathtt{N}, \mathtt{a}/2\mathtt{N}, \mathtt{N}\tau)}}.$$

Indeed

$$\frac{\mathtt{d}_{\mathtt{S}}}{2\sqrt{\mathtt{N}\mathcal{C}_2(4\mathtt{N}, \mathtt{a}/2\mathtt{N}, \mathtt{N}\tau)}} \overset{(8.9)}{=} \frac{\mathtt{d}_{\mathtt{S}}}{2\sqrt{2}e^{27(2+q)/2}} \frac{1}{\sqrt{\mathtt{N}}} \frac{1}{(\sqrt{12\tau}\mathtt{N} \max\{2, \mathtt{a}^{-1/2}\})^{3\mathtt{N}\tau}}$$

$$\ge \frac{\mathtt{d}_{\mathtt{S}}\sqrt{\tau}}{2\sqrt{2}e^{27(2+q)/2}} \left(\sqrt{\frac{12}{\tau}} \max\{2, \mathtt{a}^{-1/2}\}\right)^{-3(p-1)} (p-1)^{-3(p-1)-1/2}$$

$$= \delta_{\mathtt{S}}(\mathtt{k}_{\mathtt{S}})^{-3p}(p-1)^{-3(p-1)-1/2}$$

setting

$$\delta_{\mathtt{S}} = \frac{\mathtt{d}_{\mathtt{S}}\sqrt{\tau}}{2\sqrt{2}e^{27(2+q)/2}} \mathtt{k}_{\mathtt{S}}^3,$$

---

[24] We still denote it by u_0.

(10.1) follows by verifying that $\boldsymbol{\delta}_{\mathtt{S}} \le \delta_{\mathtt{S}}$ and noting that $p^{-3p} < (p-1)^{-3(p-1)-1/2}$ for $p > 1$.
By (8.11) and (8.9)

$$
\begin{aligned}
&\mathtt{C}_3(\mathtt{S})(2\delta)^{2(\mathtt{N}+1)} \\
&= \delta^2 \frac{2^{10} C_{\mathtt{Nem}}(p - \mathtt{N}\tau, 0, \mathtt{a}/2)|f|_{\mathtt{a},R}}{eR} \left(\frac{\mathtt{N} C_{\mathtt{mon}}(4\mathtt{N}, \mathtt{a}/2\mathtt{N}, \tau)\mathcal{C}_2(4\mathtt{N}, \mathtt{a}/2\mathtt{N}, \tau)\delta^2}{\mathtt{d}_{\mathtt{S}}{}^2} \right)^{\mathtt{N}} \\
&= \delta^2 \frac{2^{10} C_{\mathtt{Nem}}(p - \mathtt{N}\tau, 0, \mathtt{a}/2)|f|_{\mathtt{a},R}}{eR} \left(\frac{4e^{27(2+q)}3^{3\tau}(4\max\{4, (1/\mathtt{a})\})^{4\tau}\delta^2}{\tau \mathtt{d}_{\mathtt{S}}{}^2} \right)^{\mathtt{N}} (\tau\mathtt{N})^{\frac{4\tau+1}{\tau}(\tau\mathtt{N})} \\
&\le \delta^2 \frac{2^{10} C_{\mathtt{Nem}}(p - \mathtt{N}\tau, 0, \mathtt{a}/2)|f|_{\mathtt{a},R}}{eR} \left(\frac{4e^{27(2+q)}3^{3\tau}(4\max\{4, (1/\mathtt{a})\})^{4\tau}\delta^2}{\tau \mathtt{d}_{\mathtt{S}}{}^2} \right)^{\frac{p-1}{\tau}-1} (p-1)^{\frac{4\tau+1}{\tau}(p-1)}
\end{aligned}
$$

(remember that $\mathtt{N} = [(p-1)/\tau]$). Then, noting that $(p-1)^{\frac{4\tau+1}{\tau}(p-1)} < p^{5p}$ for $p > 1$ (recall that $\tau \ge 15$), we get

$$
\mathtt{C}_3(\mathtt{S})(2\delta)^{2(\mathtt{N}+1)} < \frac{1}{8\mathtt{T}_{\mathtt{S}}} p^{5p} \left(\frac{\delta}{\boldsymbol{\delta}_{\mathtt{S}}} \right)^{2\frac{p-1}{\tau}} .
$$

We conclude by applying Lemma 5.1 and (5.28)

$$
|u(t)|_{L^2} + |\partial_x^p u(t)|_{L^2} \le 2|u(t)|_p \le 4\delta, \qquad \forall |t| \le \mathtt{T}_{\mathtt{S}} p^{-5p} \left(\frac{\boldsymbol{\delta}_{\mathtt{S}}}{\delta} \right)^{\frac{2(p-1)}{\tau_{\mathtt{S}}}} ,
$$

proving (1.11).

Proof of Theorem 1.2. It is similar to the previous case but now we consider

$$
\mathtt{h}_{\mathtt{w}} \qquad \text{with} \qquad \mathtt{w}_j = \lfloor j \rfloor^p \quad (\text{and } |\cdot|_{\mathtt{w}} = \|\cdot\|_p) .
$$

We set $r = 4\delta$, an initial datum u_0 satisfying $2^p |u_0|_{L^2}, |u_0|_{L^2} + |\partial_x^p u_0|_{L^2} \le \delta$ corresponds to $u_0 \in \mathtt{h}_{\mathtt{w}}$ with $\|u_0\|_p \le 2\delta$ by (5.30). Now we can apply the Birkhoff Normal Form Theorem 8.1 with $\mathtt{N} = [\frac{p-1}{\tau_1}]$

$$
(10.2) \qquad 4\delta \le 4\frac{\boldsymbol{\delta}_{\mathtt{M}}}{\sqrt{p}} \le \mathtt{r}(\mathtt{M}) = \min\left\{ \frac{\boldsymbol{\delta}_{\mathtt{M}}}{\sqrt{\mathtt{N}}}, \frac{\sqrt{R}}{\sqrt{2^{\tau_{\mathtt{M}}+4}}} \right\} .
$$

Proceeding as in the case $\mathtt{S}$ and noting that now

$$
8\mathtt{C}_3(\mathtt{M})(4\delta)^{2(\mathtt{N}+1)} = 2^{\tau_{\mathtt{M}}+12}\frac{|f|_R}{R}\left(\frac{\mathtt{N}\delta^2}{2\delta_{\mathtt{M}}^2} \right)^{\mathtt{N}}\delta^2 \le 2^{\tau_{\mathtt{M}}+13}\delta_{\mathtt{M}}^2\frac{|f|_R}{R}(\frac{p-1}{\tau_1})^{\frac{p-1}{\tau_1}}(\frac{\delta^2}{2\delta_{\mathtt{M}}^2})^{\frac{p-1}{\tau_1}} \le \frac{1}{\mathtt{T}_{\mathtt{M}}}\left(\frac{(p-1)\delta^2}{8\boldsymbol{\delta}_{\mathtt{M}}{}^2} \right)^{\frac{p-1}{\tau_{\mathtt{M}}}} .
$$

Finally by Corollary 5.1 and (5.30) we get

$$
|u(t)|_{L^2} + |\partial_x^p u(t)|_{L^2} \le 2\|u(t)\|_p \le 8\delta, \qquad \forall |t| \le \mathtt{T}_{\mathtt{M}}\left(\frac{8\boldsymbol{\delta}_{\mathtt{M}}{}^2}{(p-1)\delta^2} \right)^{\frac{p-1}{\tau_1}} ,
$$

proving (1.13).

Proof of Corollary 1.1 In case $\mathtt{S}$ we start by noticing that for $3p\ln(\mathtt{k}_{\mathtt{S}}p) \le \ln(\boldsymbol{\delta}_{\mathtt{S}}/\delta)$ the function $\frac{\mathtt{T}_{\mathtt{S}}}{\delta^2}(p)^{-5p}\left(\frac{\boldsymbol{\delta}_{\mathtt{S}}}{\delta} \right)^{\frac{2(p-1)}{\tau_{\mathtt{S}}}}$ is increasing in p.

Let us check that $p(\delta)$ defined in (1.14) satisfies (1.10) and is $\ge 3\tau_{\mathtt{S}} + 1$ namely, passing to the logarithms and setting $y := \ln(\boldsymbol{\delta}_{\mathtt{S}}/\delta)$, we have to check that $\frac{y}{\ln(y)} > 6\tau_{\mathtt{S}}$ and $3p\ln(\mathtt{k}_{\mathtt{S}}p) \le y$. The first bound follows from the definition of $\bar{\delta}_{\mathtt{S}}$. For the second, we have

$$
3p\ln(\mathtt{k}_{\mathtt{S}}p) \le 3\left(1 + \frac{1}{6}\frac{y}{\ln(y)} \right)\left(\ln(\mathtt{k}_{\mathtt{S}}) + \ln(1 + \frac{1}{6}\frac{y}{\ln(y)}) \right) \le y
$$

provided that[25]

$$y \geq \max\{\mathtt{k_S}, 40\}.$$

Now we have to show that

$$\mathtt{T_S} e^{\frac{\ln^2(\delta_S/\delta)}{4\tau \ln\ln(\delta_S/\delta)}} \leq \mathtt{T_S}(p)^{-5p} \left(\frac{\delta_S}{\delta}\right)^{\frac{2(p-1)}{\tau}}$$

wich amounts to

$$e^{\frac{y^2}{4\tau \ln y}} (p)^{5p} e^{-\frac{2(p-1)}{\tau}y} \leq 1$$

or equivalently

$$\frac{y^2}{4\tau \ \ln y} + 5p \ln(p) - \frac{2(p-1)}{\tau} y \leq 0 \,.$$

Assuming $\frac{y}{\ln y} > 6$, we have $1 + \frac{y}{6\ln y} - \frac{\tau}{6} < p < \frac{y}{3\ln y}$ we get

$$\frac{y^2}{4\tau \ln y} + 5p \ln(\mathtt{K_S}p) - \frac{2(p-1)}{\tau} y \leq \frac{y^2}{4\tau \ln y} + \frac{5}{3}\frac{y}{\ln(y)} \ln\left(\frac{y}{3\ln(y)}\right) - 2y\left(\frac{1}{6\tau}\frac{y}{\ln(y)} - \frac{1}{6}\right)$$

$$\leq -\frac{y^2}{12\tau \ln y} + 2y < 0$$

if $\frac{y}{\ln(y)} > 24\tau > 6$. Note that the last inequality holds if $y \geq 24\tau^2$ (recall that $\tau \geq 15$). Recollecting the condition that y has to satisfy is

$$y \geq \max\{\mathtt{k_S}, \ 24\tau^2\} \,,$$

namely $\delta \leq \bar{\delta}_\mathtt{S}$.

M) Since we are assuming $\delta \leq \bar{\delta}_\mathtt{M}$ we have that p defined in (1.16) satisfies $p > 1 + 3\tau_\mathtt{M}$. Moreover by (1.16), the bound (1.12) holds. Then Theorem 1.2 applies and (1.17) follows directly by (1.13).

Part 3. Appendices

APPENDIX A. CONSTANTS.

In this subsection are listed all the constants appearing along the paper. We first introduce some auxiliary constants. Given $t, \sigma, \zeta > 0$, $p > 1/2$, $0 < \theta < 1$, $s, q \geq 0$, we set[26]

$$C_{\mathtt{alg}}(p) := 2^p \left(\sum_{i\in\mathbb{Z}} \langle i \rangle^{-2p}\right)^{1/2}, \qquad C_{\mathtt{alg},\mathtt{M}}(p) := \sqrt{2}\sqrt{2 + \frac{2p+1}{2p-1}},$$

$$C_{\mathtt{Nem}}(p,s,t) := C_{\mathtt{alg}}(p)\left(e^s + \sup_{x\geq 1} x^p e^{-tx+sx^\theta}\right), \quad C_{\mathtt{mon}}(t,\sigma,p) := 2^{p+1} p^p \max\left\{(2t)^p, \sigma^{-p}, 1\right\},$$

$$\mathcal{C}_1 := 28\,\theta^{-1}(q+3)\left(\frac{2^3 \cdot 13(q+3)}{\theta(1-\theta)}\right)^{\frac{2}{\theta}} \left(\ln\left(\frac{2^3 \cdot 13(q+3)}{\theta(1-\theta)}\right)\right)^{\frac{2}{\theta}+1}, \quad C_* = 13/(1-\theta),$$

$$\mathcal{C}_2(t,\sigma,\zeta) := e^{27(2+q)} C_{\mathtt{mon}}(t,\sigma,3\zeta),$$

$$\tau := \tau_0(2+q), \qquad \tau_0 := 15/2, \qquad \tau_1 := 2\left(\tau_0 + \frac{3}{2\ln 2}\right)(2+q)$$

---

[25] Note that the function

$$y \mapsto y - 3\left(1 + \frac{1}{6}\frac{y}{\ln(y)}\right)\left(\ln y + \ln(1 + \frac{1}{6}\frac{y}{\ln(y)})\right)$$

is positive for $y \geq 40$.

[26] Regarding $C_{\mathtt{Nem}}$ note that

$$\sup_{x\geq 1} x^p e^{-tx+sx^\theta} \leq \exp\left((1-\theta)\left(\frac{s}{t^\theta}\right)^{\frac{1}{1-\theta}}\right) \max\left\{\frac{p}{e(1-\theta)t}, e^{-\frac{t(1-\theta)}{p}}\right\}^p \,.$$

Here are the constants appearing in Theorem 1.1:

$$\boldsymbol{\delta}_{\mathsf{G}} \quad := \quad \min\left\{\frac{\sqrt{R}}{4C_{\mathtt{alg}}(p)}\,,\, \delta_{\mathsf{G}} e^{-\left(\max\left\{16(4\mathcal{C}_1)^\theta \eta_{\mathsf{G}}^{-3}, 2^{\frac{2\theta+4}{4-\theta}}\right\}\right)^{4/\theta}}\right\}\,,\quad \mathtt{T}_{\mathsf{G}} := \frac{2^4 e \boldsymbol{\delta}_{\mathsf{G}}^{\,2}}{\gamma}\,, \qquad \text{where}$$

$$\delta_{\mathsf{G}} \quad := \quad \frac{\sqrt{\gamma R}}{C_{\mathtt{alg}}(p)\sqrt{2^{11}eC_{\mathtt{Nem}}(p, s-\eta_{\mathsf{G}}, \mathtt{a}-a-\eta_{\mathsf{G}})|f|_{\mathtt{a},R}}}\,,\qquad \eta_{\mathsf{G}} := \min\left\{\frac{\mathtt{a}-a}{2}, s\right\}\,.$$

Here are the constants appearing in Proposition 1.1

$$\tau_{\mathsf{S}} := \tau = \frac{15}{2}(2+q)\,,\qquad \boldsymbol{\delta}_{\mathsf{S}} = \min\left\{\frac{\sqrt{3^\tau \tau \gamma R}}{2^{10}e^{2\tau}(\mathtt{k}_{\mathsf{S}}\tau)^{4\tau}\sqrt{C_{\mathtt{Nem}}(p-\mathtt{N}\tau_{\mathsf{S}}, 0, \mathtt{a}/2)|f|_{\mathtt{a},R}}}, \frac{\sqrt{R}}{2^{\tau+5}}\right\}$$

$$\mathtt{k}_{\mathsf{S}} := \sqrt{\frac{12}{\tau}}\max\left\{2, \mathtt{a}^{-1/2}\right\}\,,\qquad \mathtt{T}_{\mathsf{S}} = \frac{2^6 e^{2\tau}3^{3\tau}(4\max\{4, (1/\mathtt{a})\})^{4\tau}}{\tau\gamma}\,.$$

Here are the constants in Theorem 1.2

$$\tau_{\mathsf{M}} := \tau_1 = \left(15+\frac{3}{\ln 2}\right)(2+q)\,,\qquad \boldsymbol{\delta}_{\mathsf{M}} := \min\left\{\frac{\sqrt{\tau_1}\delta_{\mathsf{M}}}{4}, \sqrt{\frac{R}{2^{\tau_1+8}}}\right\}\,,\qquad \mathtt{T}_{\mathsf{M}} := \frac{R}{2^{\tau_1+13}\delta_{\mathsf{M}}^2|f|_R} = \frac{2^3 e 6^{\tau_1}(4^6 e^{27})^{2+q}}{\gamma}\,;$$

where, recalling 8.2,

$$\delta_{\mathsf{M}} = \frac{\sqrt{\gamma R}}{\sqrt{2^{17}e12^{\tau_1}(4^6 e^{27})^{2+q}|f|_R}}\,.$$

Here are the constants appearing in Corollary 1.1:

$$\bar{\delta}_{\mathsf{S}} \quad := \quad \boldsymbol{\delta}_{\mathsf{S}} e^{-\max\{\mathtt{k}_{\mathsf{S}}, 24\tau^2\}}\,,\qquad \bar{\delta}_{\mathsf{M}} := \frac{\delta_{\mathsf{M}}}{4\tau_1}\,.$$

Appendix B. Proofs of the main properties of the norms

Lemma B.1. *Let $0 < r_1 < r$. Let E be a Banach space endowed with the norm $|\cdot|_E$. Let $X : B_r \to E$ a vector field satisfying*

$$\sup_{B_r}|X|_E \le \delta_0\,.$$

Then the flow $\Phi(u, t)$ of the vector field[27] is well defined for every

$$|t| \le T := \frac{r-r_1}{\delta_0}$$

and $u \in B_{r_1}$ with estimate

$$|\Phi(u, t) - u|_E \le \delta_0|t|\,,\qquad \forall\, |t| \le T\,.$$

---

[27]Namely the solution of the equation $\partial_t \Phi(u, t) = X(\Phi(u, t))$ with initial datum $\Phi(u, 0) = u$.

Proof. Fix $u \in B_{r_1}$. Let us first prove that $\Phi(u,t)$ exists $\forall |t| \leq T$. Otherwise there exists a time[28] $0 < t_0 < T$ such that $|\Phi(u,t)|_E < r$ for every $0 \leq t < t_0$ but $|\Phi(u,t_0)|_E = r$. Then, by the fundamental theorem of calculus

$$(B.1) \qquad \Phi(u,t_0) - u = \int_0^{t_0} X(\Phi(u,\tau))d\tau \,.$$

Therefore

$$
\begin{aligned}
r - r_1 &\leq& |\Phi(u,t_0)|_E - |u|_E \leq |\Phi(u,t_0) - u|_E \leq \int_0^{t_0} |X(\Phi(u,\tau))|_E d\tau \leq \delta_0 t_0 \\
&<& \delta_0 T = r - r_1 \,,
\end{aligned}
$$

which is a contradiction Finally, for every $|t| \leq T$,

$$|\Phi(u,t) - u|_E \leq \left| \int_0^t |X(\Phi(u,\tau))|_E d\tau \right| \leq \delta_0 |t| \,.$$

$\square$

Proof of Lemma 2.1. For brevity we set, for every $r' > 0$

$$|\cdot|_{r'} := |\cdot|_{r',\eta,\mathbf{w}} \,.$$

We use Lemma B.1, with $E \to \mathbf{h}_{\mathbf{w}}$, $X \to X_S$, $\delta_0 \to (r+\rho)|S|_{r+\rho}$, $r \to r+\rho$, $r_1 \to r$, $T \to 8e$. Then the fact that the time 1-Hamiltonian flow $\Phi_S^1 : B_r(\mathbf{h}_{\mathbf{w}}) \to B_{r+\rho}(\mathbf{h}_{\mathbf{w}})$ is well defined, analytic, symplectic follows, since for any $\eta \geq 0$

$$\sup_{u \in B_{r+\rho}(\mathbf{h}_{\mathbf{w}})} |X_S|_{\mathbf{h}_{\mathbf{w}}} \leq (r+\rho)|S|_{r+\rho} < \frac{\rho}{8e} \,.$$

Regarding the estimate (2.3), again by Lemma B.1 (choosing $t = 1$), we get

$$\sup_{u \in B_r(\mathbf{h}_{\mathbf{w}})} \left| \Phi_S^1(u) - u \right|_{\mathbf{h}_{\mathbf{w}}} \leq (r+\rho)|S|_{r+\rho} < \frac{\rho}{8e} \,.$$

Estimates (2.4),(2.5),(2.6) directly follow by (2.7) with $h = 0,1,2$, respectively and $c_k = 1/k!$, recalling that by Lie series

$$H \circ \Phi_S^1 = e^{\mathrm{ad}_S} H = \sum_{k=0}^{\infty} \frac{\mathrm{ad}_S^k H}{k!} = \sum_{k=0}^{\infty} \frac{H^{(k)}}{k!} \,,$$

where $H^{(i)} := \mathrm{ad}_S^i(H) = \mathrm{ad}_S(H^{(i-1)})$, $H^{(0)} := H$.

Let us prove (2.7). Fix $k \in \mathbb{N}$, $k > 0$ and set

$$r_i := r + \rho(1 - \frac{i}{k}) \,, \qquad i = 0,\dots,k \,.$$

Note that, by the immersion properties of the norm (recall Remark 2.1)

$$(B.2) \qquad\qquad |S|_{r_i} \leq |S|_{r+\rho} \,, \qquad \forall i = 0,\dots,k \,.$$

Noting that

$$(B.3) \qquad\qquad 1 + \frac{k r_i}{\rho} \leq k\left(1 + \frac{r}{\rho}\right) \,, \qquad \forall i = 0,\dots,k \,,$$

by using k times (2.1) we have

$$
\begin{aligned}
|H^{(k)}|_r &=& |\{S, H^{(k-1)}\}|_r \leq 4(1 + \frac{k r}{\rho})|H^{(k-1)}|_{r_{k-1}}|S|_{r_{k-1}} \\
&\overset{(B.2)}{\leq}& |H|_{r+\rho}|S|_{r+\rho}^k 4^k \prod_{i=1}^{k}(1 + \frac{k r_i}{\rho}) \overset{(B.3)}{\leq} |H|_{r+\rho}\left(4k\left(1 + \frac{r}{\rho}\right)|S|_{r+\rho}\right)^k \,.
\end{aligned}
$$

---

[28]We assume t_0 positive, the negative case is analogous.

Then, using $k^k \le e^k k!$, we get

$$
\left| \sum_{k \ge h} c_k H^{(k)} \right|_r \le \sum_{k \ge h} |c_k| |H^{(k)}|_r \le |H|_{r+\rho} \sum_{k \ge h} \left(4e \left(1 + \frac{r}{\rho} \right) |S|_{r+\rho} \right)^k
$$

$$
= |H|_{r+\rho} \sum_{k \ge h} (|S|_{r+\rho}/2\delta)^k \overset{(2.2)}{\le} 2|H|_{r+\rho}(|S|_{r+\rho}/2\delta)^h .
$$

Finally, if S and H satisfy mass conservation so does each $\mathrm{ad}_S^k H$, $k \ge 1$, hence $H \circ \Phi_S^1$ too. $\qquad\square$

Proof of Lemma 3.1 . We first prove (i). It is easily seen that:

$$
X_{\underline{H}_\eta}^{(j)}(u) = \mathrm{i} \sum_{\boldsymbol{\alpha},\boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}} |H_{\boldsymbol{\alpha},\boldsymbol{\beta}}| \boldsymbol{\beta}_j e^{\eta|\pi(\boldsymbol{\alpha}-\boldsymbol{\beta})|} u^{\boldsymbol{\alpha}} \bar{u}^{\boldsymbol{\beta}-e_j} .
$$

Now

$$
|X_{\underline{H}_\eta}(u)|_{\mathtt{w}} \le |X_{\underline{H}_\eta}(\underline{u})|_{\mathtt{w}} , \quad \underline{u} = (|u_j|)_{j \in \mathbb{Z}}
$$

hence, in evaluating the supremum of $|X_{\underline{H}_\eta}|_{\mathtt{w}}$ over $|u|_{\mathtt{w}} \le r$ we ca restrict to the case in which $u = (u_j)_{j \in \mathbb{Z}}$ has all real positive components. Hence

$$
|H|_{r,\eta,\mathtt{w}} = r^{-1} \sup_{|u|_{\mathtt{w}} \le r} \left| \left(\sum^* |H_{\boldsymbol{\alpha},\boldsymbol{\beta}}| \boldsymbol{\beta}_j e^{\eta|\pi(\boldsymbol{\alpha}-\boldsymbol{\beta})|} |u|^{\boldsymbol{\alpha}+\boldsymbol{\beta}-e_j} \right)_{j \in \mathbb{Z}} \right|_{\mathtt{w}} .
$$

Then

(B.4)
$$
|H|_{r,\eta,\mathtt{w}} = \frac{1}{2r} \sup_{|u|_{\mathtt{w}} \le r} \left| \left(W_\eta^{(j)}(u) \right)_{j \in \mathbb{Z}} \right|_{\mathtt{w}}
$$

where

$$
W_\eta^{(j)}(u) = \sum^* |H_{\boldsymbol{\alpha},\boldsymbol{\beta}}| \left(\boldsymbol{\alpha}_j + \boldsymbol{\beta}_j \right) e^{\eta|\pi(\boldsymbol{\alpha}-\boldsymbol{\beta})|} u^{\boldsymbol{\alpha}+\boldsymbol{\beta}-e_j} ,
$$

since, by the reality condition 1.23, we have

$$
\sum^* |H_{\boldsymbol{\alpha},\boldsymbol{\beta}}| \boldsymbol{\beta}_j e^{\eta|\pi(\boldsymbol{\alpha}-\boldsymbol{\beta})|} u^{\boldsymbol{\alpha}+\boldsymbol{\beta}-e_j} = \sum^* |H_{\boldsymbol{\alpha},\boldsymbol{\beta}}| \boldsymbol{\alpha}_j e^{\eta|\pi(\boldsymbol{\alpha}-\boldsymbol{\beta})|} u^{\boldsymbol{\alpha}+\boldsymbol{\beta}-e_j} = \frac{1}{2} W_\eta^{(j)}(u).
$$

By the linear map

$$
L_{r,\mathtt{w}} : \ell^2 \to \mathtt{h}_{\mathtt{w}} , \qquad y_j \mapsto \frac{r}{\mathtt{w}_j} y_j = u_j ,
$$

the ball of radius 1 in ℓ^2 is isomorphic to the the ball of radius r in $\mathtt{h}_{\mathtt{w}}$, namely $L_{r,\mathtt{w}}(B_1(\ell^2)) = B_r(\mathtt{h}_{\mathtt{w}})$. We have

$$
Y_H^{(j)}(y; r, \eta, \mathtt{w}) = \frac{1}{2} W_\eta^{(j)}(L_{r,\mathtt{w}} y) .
$$

Then (i) follows.

In order to prove item (ii) we rely on the fact that, since we are using the η-majorant norm, the supremum over y in the norm is achieved on the real positive cone. Moreover, given $u, v \in \ell^2$, if

$$
|u_j| \le |v_j| , \quad \forall j \in \mathbb{Z}
$$

then $|u|_{\ell^2} \le |v|_{\ell^2}$. $\qquad\square$

Proof of Lemma 5.4.   Let us look at the time evolution of $|v(t)|_{\mathtt{w}}^2$. By construction and Cauchy-Schwarz inequality

$$2|v(t)|_{\mathtt{w}} \left| \frac{d}{dt}|v(t)|_{\mathtt{w}} \right| \quad = \quad \left| \frac{d}{dt}|v(t)|_{\mathtt{w}}^2 \right| = 2|\operatorname{Re}(v, \dot{v})_{\mathtt{h_w}}| = 2|\operatorname{Re}(v, X_R)_{\mathtt{h_w}}|$$
$$\leq \quad 2|v(t)|_{\mathtt{w}}|X_{\underline{R}}|_{\mathtt{w}} \leq 2r|v(t)|_{\mathtt{w}}|R|_{r,\eta,\mathtt{w}}$$

as long as $|v(t)|_{\mathtt{w}} \leq r$; namely

(B.5)
$$\left| \frac{d}{dt}|v(t)|_{\mathtt{w}} \right| \leq r|R|_{r,\eta,\mathtt{w}}$$

as long as $|v(t)|_{\mathtt{w}} \leq r$.

Assume by contradiction that there exists a time[29]

$$0 < T_0 < \frac{1}{8|R|_{r,\eta,\mathtt{w}}}$$

such that

(B.6)
$$\left| |v(t)|_{\mathtt{w}} - |v_0|_{\mathtt{w}} \right| < \frac{r}{8}, \quad \forall\, 0 \leq t < T_0, \qquad \text{but} \quad \left| |v(T_0)|_{\mathtt{w}} - |v_0|_{\mathtt{w}} \right| = \frac{r}{8}.$$

Then

$$|v(t)|_{\mathtt{w}} \leq |v_0|_{\mathtt{w}} + \frac{r}{8} < r \qquad \forall\, 0 \leq t \leq T_0.$$

By (B.5) we get

$$\left| |v(T_0)|_{\mathtt{w}} - |v_0|_{\mathtt{w}} \right| \leq r|R|_{r,\eta,\mathtt{w}}T_0 < \frac{r}{8},$$

which contradicts (B.6), proving (5.26).

Proof of Lemma 5.5.   We first note that (see, e.g. Lemma 17 of [BDG10]) for $p > 1/2$ and every sequence $\{x_i\}_{i\in\mathbb{Z}}$, $x_i \geq 0$,

$$\left(\sum_{i\in\mathbb{Z}} x_i \right)^2 \leq c \sum_{i\in\mathbb{Z}} \left(\frac{\langle i \rangle^p \langle j - i \rangle^p}{\langle j \rangle^p} x_i \right)^2,$$

with $c := 4^p \sum_{i\in\mathbb{Z}} \langle i \rangle^{-2p} = (C_{\mathtt{alg}}(p))^2$. Then

$$|f \star g|_{p,s,a}^2 \quad \leq \quad \sum_j e^{2s\langle j \rangle^\theta} e^{2a|j|} \langle j \rangle^{2p} \left(\sum_i |f_i||g_{j-i}| \right)^2$$
$$\leq \quad c \sum_j e^{2s\langle j \rangle^\theta} e^{2a|j|} \sum_i \langle i \rangle^{2p} \langle j - i \rangle^{2p} |f_i|^2 |g_{j-i}|^2$$
$$= \quad c \sum_i e^{2s\langle i \rangle^\theta} e^{2a|i|} \langle i \rangle^{2p} |f_i|^2 \sum_j \langle j - i \rangle^{2p} e^{2s\langle j-i \rangle^\theta} e^{2a|j-i|} |g_{j-i}|^2$$
$$= \quad c|f|_{p,s,a}^2 |g|_{p,s,a}^2$$

Regarding the second estimate, we set

$$\phi(i,j) := \frac{\lfloor j \rfloor}{\lfloor i \rfloor \lfloor j - i \rfloor}, \qquad \forall\, i, j \in \mathbb{Z}.$$

Note that

(B.7)
$$\phi(i,j) = \phi(j,i) = \phi(-i,-j).$$

We claim that

(B.8)
$$\phi(i,j) \leq 1.$$

---

[29]The case $T_0 < 0$ is analogous.

Indeed by (B.7) we can consider only the case $j \geq 0$. Since $\phi(-|i|, j) \leq \phi(|i|, j)$ we can consider only the case $i \geq 0$. Again by (B.7) we can assume $j \geq i$. In particular we can take $j > i > 0$, (B.8) being trivial in the cases $j = i$, $i = 0$. We have

$$\phi(i+1, i) = \frac{i+1}{2\lfloor i \rfloor} \leq \frac{3}{4}, \qquad \phi(j, 1) = \frac{j}{2(j-1)} \leq 1 \,.$$

Then it remains also to discuss the case $j - 2 \geq i \geq 2$; we have

$$\phi(i, j) = \frac{j}{i(j-i)} = \frac{1}{i} + \frac{1}{j-i} \leq 1 \,,$$

proving (B.8).

For $q \geq 0$ set

(B.9)
$$c_q := \sup_{j \in \mathbb{Z}} \sum_{i \in \mathbb{Z}} (\phi(i, j))^q = \sup_{j \geq 0} \sum_{i \in \mathbb{Z}} (\phi(i, j))^q \,.$$

We claim that

(B.10)
$$c_q \leq 4 + 2\frac{q+1}{q-1} < \infty, \qquad \forall\, q > 1 \,.$$

Indeed, since $\lfloor j \rfloor / \lfloor j+1 \rfloor \leq 1$ and $\lfloor j \rfloor / \lfloor j-1 \rfloor \leq 3/2$ for $j \geq 0$, we have[30]

$$
\begin{aligned}
c_q \;=\;& \sup_{j \geq 0} \left(\frac{\lfloor j \rfloor^q}{2^{q-1} \lfloor j+1 \rfloor^q} + \frac{1}{2^{q-1}} + \frac{\lfloor j \rfloor^q}{2^{q-1} \lfloor j-1 \rfloor^q} + \sum_{i \leq -2,\, 2 \leq i \leq j-2,\, i \geq j+2} (\phi(i, j))^q \right) \\
\leq\;& 2^{3-q} + \sup_{j \geq 0} \left(\sum_{i \geq 2} \frac{\lfloor j \rfloor^q}{i^q (j+i)^q} + \sum_{2 \leq i \leq j-2} \left(\frac{1}{i} + \frac{1}{(j-i)} \right)^q + \sum_{i \geq j+2} \frac{\lfloor j \rfloor^q}{i^q (i-j)^q} \right) \\
\leq\;& 2^{3-q} + \sum_{i \geq 2} \frac{1}{i^q} + 2^{q-1} \sum_{2 \leq i \leq j-2} \left(\frac{1}{i^q} + \frac{1}{(j-i)^q} \right) + \sum_{i \geq j+2} \frac{1}{(i-j)^q} \\
\leq\;& 4 + 2\frac{q+1}{q-1} \,,
\end{aligned}
$$

using that $(x + y)^q \leq 2^{q-1}(x^q + y^q)$ for $x, y \geq 0$ and that[31]

$$\sum_{i \geq 2} i^{-q} \leq \frac{q+1}{2^q(q-1)} \,.$$

Note that for every $q, q_0 \geq 0$ we have

(B.11)
$$c_{q_0+q} \leq c_{q_0}$$

since

$$c_{q_0+q} := \sup_{j \in \mathbb{Z}} \sum_{i \in \mathbb{Z}} (\phi(i, j))^{q_0} (\phi(i, j))^q \overset{\text{(B.8)}}{\leq} \sup_{j \in \mathbb{Z}} \sum_{i \in \mathbb{Z}} (\phi(i, j))^{q_0} = c_{q_0} \,.$$

We now note that for $p > 1/2$, $j \in \mathbb{Z}$ and every sequence $\{x_i\}_{i \in \mathbb{Z}}$, $x_i \geq 0$, we have by Cauchy-Schwarz inequality

$$\left(\sum_{i \in \mathbb{Z}} x_i \right)^2 = \left(\sum_{i \in \mathbb{Z}} (\phi(i, j))^p (\phi(i, j))^{-p} x_i \right)^2 \leq c_{2p} \sum_{i \in \mathbb{Z}} \left((\phi(i, j))^{-p} x_i \right)^2 \,,$$

---

[30]Note that the term $\left(\frac{1}{i} + \frac{1}{(j-i)} \right)^q$ for $j = 4$ and $i = 2$ is 1 for every q.

[31]$\sum_{i \geq 2} 1^{-q} \leq 2^{-q} + \int_2^\infty x^{-q} dx$

with c_{2p} defined in (B.9). Using the above inequality we get

$$
\begin{aligned}
\|f \star g\|_p^2 &\leq \sum_j \lfloor j \rfloor^{2p} \Big(\sum_i |f_i| |g_{j-i}| \Big)^2 \\
&\leq c_{2p} \sum_j \sum_i \lfloor i \rfloor^{2p} \lfloor j-i \rfloor^{2p} |f_i|^2 |g_{j-i}|^2 \\
&= c_{2p} \sum_i \lfloor i \rfloor^{2p} |f_i|^2 \sum_j \lfloor j-i \rfloor^{2p} |g_{j-i}|^2 \\
&= c_{2p} \|f\|_p^2 \|g\|_p^2 .
\end{aligned}
$$

The proof ends recalling (B.10). □

Lemma B.2 (Nemitskii operators). *Let $p > 1/2$. (i) Fix $s \geq 0, a_0 \geq 0$. Consider a sequence $F^{(d)} = \left(F_j^{(d)} \right)_{j \in \mathbb{Z}} \in h_{p,s,a_0}$, $d \geq 1$, such that*

$$
\tag{B.12} \sum_{d=1}^{\infty} d |F^{(d)}|_{p,s,a_0} R^d < \infty
$$

for some $R > 0$.
For $u = (u_j)_{j \in \mathbb{Z}}$ let $\bar{u} = (\overline{u_{-j}})_{j \in \mathbb{Z}}$ and consider the Hamiltonian

$$
H(u) = \sum_{d=1}^{\infty} \left(F^{(d)} \star \underbrace{u \star \cdots \star u}_{d \text{ times}} \star \underbrace{\bar{u} \star \cdots \star \bar{u}}_{d \text{ times}} \right)_0 .
$$

For all (η, a, r) such that $\eta + a \leq a_0$ and $(C_{\mathtt{alg}}(p)r)^2 \leq R$, we have that $H \in \mathcal{H}_{r,\eta}(h_{p,s,a})$ and

$$
|H|_{r,\eta,\mathtt{w}(p,s,a)} \leq r^{-1} \sum_{d=1}^{\infty} d |F^{(d)}|_{p,s,a_0} (C_{\mathtt{alg}}(p)r)^{2d-1} < \infty.
$$

(ii) Analogously if $F^{(d)}$ are constants satisfying

$$
\tag{B.13} \sum_{d=1}^{\infty} d |F^{(d)}| R^d < \infty
$$

and $(C_{\mathtt{alg,M}}(p)r)^2 \leq R$, then $H \in \mathcal{H}^{r,p}$ with

$$
\tag{B.14} \|H\|_{r,p} \leq 2^p r^{-1} \sum_{d=1}^{\infty} d |F^{(d)}| (C_{\mathtt{alg,M}}(p)r)^{2d-1} < \infty.
$$

Proof. (i) By definition the η-majorant Hamiltonian is

$$
\underline{H}_\eta = \sum_d \sum_{\substack{j_0, j_1, \ldots, j_{2d} \\ j_0 + \sum_{i=1}^{2d} (-1)^i j_i = 0}} e^{\eta |\pi_{j_1, \ldots, j_{2d}}|} |F_{j_0}^{(d)}| \overline{u_{j_1}} u_{j_2} \overline{u_{j_3}} \ldots u_{j_{2d}}
$$

where

$$
\pi_{j_1, \ldots, j_{2d}} = \sum_{i=1}^{2d} (-1)^i j_i = -j_0 ,
$$

hence

$$
\underline{H}_\eta = \sum_d \left(\underline{F}_\eta^{(d)} \star \underbrace{u \star \cdots \star u}_{d \text{ times}} \star \underbrace{\bar{u} \star \cdots \star \bar{u}}_{d \text{ times}} \right)_0 , \quad \underline{F}_\eta^{(d)} := \left(e^{\eta |j|} |F_j^{(d)}| \right)_{j \in \mathbb{Z}} .
$$

consequently

$$X_{\underline{H}_\eta}^{(j)} = \sum_d d\left(\underline{F}_\eta^{(d)} \star \underbrace{u \star \cdots \star u}_{d \text{ times}} \star \underbrace{\bar{u} \star \cdots \star \bar{u}}_{d-1 \text{ times}}\right)_j .$$

Moreover

$$|X_{\underline{H}_\eta}|_{p,s,a} \le \sum_d d(C_{\texttt{alg}}(p))^{2d-1}|\underline{F}_\eta^{(d)}|_{p,s,a}(|u|_{p,s,a})^{2d-1} .$$

Since

$$|\underline{F}_\eta^{(d)}|_{p,s,a} = |F^{(d)}|_{s,a+\eta,p} \le |F^{(d)}|_{p,s,a_0}$$

we get

$$|X_{\underline{H}_\eta}|_{p,s,a} \le \sum_d d(C_{\texttt{alg}}(p))^{2d-1}|F^{(d)}|_{p,s,a_0}(|u|_{p,s,a})^{2d-1} .$$

Therefore

$$|H|_{r,p,\eta}^{(p,s,a,0)} = r^{-1}\left(\sup_{|u|_{p,s,a}<r}\left|X_{\underline{H}_\eta}\right|_{p,s,a}\right) \le r^{-1}\sum_d d|F^{(d)}|_{p,s,a_0}(C_{\texttt{alg}}(p)r)^{2d-1} < \infty.$$

(ii) The proof is analogous to point (i). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of Proposition 6.2 . We start by Taylor expanding H in homogeneous components. The majorant analiticity implies that for a homogeneous component of degree d one has

$$|H^{(d)}|_{r,\eta,\mathbf{w}(p,s,a)} \le |H|_{r,\eta,,\mathbf{w}(p,s,a)}$$

Now let us consider the polinomial map (homogeneous of degree $d-1$) $X_{H^{(d)}} : \mathbf{h}_{p,s,a} \to \mathbf{h}_{p,s,a}$; as is habitual we identify the polynomial map with the corresponding symmetric multilinear operator $M^{(d-1)} : \mathbf{h}_{p,s,a}^{d-1} \to \mathbf{h}_{p,s,a}$. Since we are in a Hilbert space, one has that

$$|\underline{M}|_{p,s,a}^{\text{op}} := \sup_{\substack{u_1,\ldots u_{d-1}\in\mathbf{h}_{p,s,a} \\ |u_i|_{p,s,a}\le 1}} |\underline{M}^{(d-1)}(u_1,\ldots,u_{d-1})|_{p,s,a} = \sup_{|u|_{p,s,a}\le 1} |\underline{M}^{(d-1)}(u,\ldots,u)|_{p,s,a}$$

$$= \sup_{|u|_{p,s,a}\le 1} |X_{\underline{H}^{(d)}}|_{p,s,a} \le r^{-d+2}|H|_{r,\eta,\mathbf{w}(p,s,a)}$$

for all $\eta \ge 0$. Now let us compute the tame norm on a homogeneous component, i.e.

$$\sup_{|u|_{p_0,s,a}\le r-\rho} \frac{|\underline{M}^{(d-1)}(u^{d-1})|_{p,s,a}}{|u|_{p,s,a}} = \sup_{|u|_{p_0,s,a}\le r-\rho} \frac{|\underline{N_p}^{(d-1)}(u^{d-1})|_{p_0,s,a}}{|u|_{p,s,a}}$$

where

$$\underline{N_p}^{(d-1,j)}(u^{d-1}) = \langle j\rangle^{p-p_0} \sum_{j_1,\ldots,j_{d-1}} |M_{j_1,\ldots j_{d-1}}^{(d-1,j)}|u_{j_1}\ldots u_{j_{d-1}}$$

now setting $\pi = \sum_i j_i - j$ we have

$$
\underline{N_p}^{(d-1)}(u_1, \ldots, u_{d-1})
$$
$$
\leq (d-1)\langle j\rangle^{p-p_0} \sum_{\substack{j_1,\ldots,j_{d-1}: \\ |j_1| \geq |j_i|}} |M^{(d-1,j)}_{j_1,\ldots j_{d-1}}| u_{j_1} \ldots u_{j_{d-1}}
$$
$$
\leq (d-1) \sum_{\substack{j_1,\ldots,j_{d-1}: \\ |j_1| \geq |j_i|}} \left(\sum_i \langle j_i\rangle + |\pi|\right)^{p-p_0} |M^{(d-1,j)}_{j_1,\ldots j_{d-1}}| u_{j_1} \ldots u_{j_{d-1}}
$$
$$
\leq (d-1)2^{p-p_0} C(\eta,p) \sum_{j_1,\ldots,j_{d-1}} e^{\eta|\pi|} |M^{(d-1,j)}_{j_1,\ldots j_{d-1}}| u_{j_1} \ldots u_{j_{d-1}}
$$
$$
+ (d-1)2^{p-p_0}(d-1)^{p-p_0} \sum_{j_1,\ldots,j_{d-1}} |M^{(d-1,j)}_{j_1,\ldots j_{d-1}}| \langle j_1\rangle^{p-p_0} u_{j_1} \ldots u_{j_{d-1}}
$$

which means that for any $|u|_{p_0,s,a} \leq r - \rho$ one has

$$
|\underline{N_p}^{(d-1)}(u^{d-1})|_{p_0,s,a}
$$
$$
\leq (d-1)2^{p-p_0} C(\eta,p) |H^{(d)}|_{r-\rho,\eta,\mathtt{w}(p_0,s,a)} |u|_{p_0,s,a} + 2^{p-p_0}(d-1)^{p-p_0+1} |\underline{M}|^{\mathrm{op}}_{p_0,s,a}(r-\rho)^{d-2} |u|_{p,s,a}
$$
$$
\leq (d-1)2^{p-p_0}(C(\eta,p) + (d-1)^{p-p_0})(1-\tfrac{\rho}{r})^{d-2} |H|_{r,\eta,\mathtt{w}(p_0,s,a)} |u|_{p,s,a}
$$

We conclude that

$$
\sup_{|u|_{p_0,s,a} \leq r} \frac{|X_{\underline{H}}|_{p,s,a}}{|u|_{p,s,a}} \leq 2^{p-p_0} |H|_{r,\eta,\mathtt{w}(p_0,s,a)} \sum_{d\geq 2}(d-1)\big(C(\eta,p) + (d-1)^{p-p_0}\big)(1-\tfrac{\rho}{r})^{d-2}
$$

and the thesis follows since the right hand side is convergent. $\square$

APPENDIX C. SMALL DIVISOR ESTIMATES

Let us start with two techincal lemmata.

Lemma C.1. *For $p, \beta > 0$ and $x_0 \geq 0$ we have that*

$$
\max_{x \geq x_0} x^p e^{-\beta x} = \begin{cases} (p/\beta)^p e^{-p} & \text{if } x_0 \leq p/\beta, \\ x_0^p e^{-\beta x_0} & \text{if } x_0 > p/\beta. \end{cases}
$$

Lemma C.2. *Let $0 < a < 1$ and $x_1 \geq x_2 \geq \ldots \geq x_N \geq 2$. Then*

$$
\frac{\sum_{1\leq \ell \leq N} x_\ell}{\prod_{1\leq \ell \leq N} x_\ell^a} \leq x_1^{1-a} + \frac{2}{a x_1^a}.
$$

Proof. By induction over N. It is obviously true for $N = 1$. Assume that it hols for N and prove it for $N+1$. $\square$

Proof of Lemma 6.1. The fact that this (6.5) holds true when $\pi = 0$ is proven in [Bou96b] and [CLSY]. The bound (6.5) is equivalent to proving

$$(C.1) \qquad \sum_{l \geq 1} \widehat{n}_l^\theta - 2\widehat{n}_1^\theta + \theta|\pi| - (2 - 2^\theta) \sum_{l \geq 3} \widehat{n}_l^\theta \geq 0.$$

i.e.

$$(C.2) \qquad \sum_{l \geq 2} \widehat{n}_l^\theta - \widehat{n}_1^\theta + \theta|\pi| - (2 - 2^\theta) \sum_{l \geq 3} \widehat{n}_l^\theta \geq 0.$$

Inequality (C.2) then follows from

$$(C.3) \qquad f(|\pi|) := \sum_{l \geq 2} \widehat{n}_l^\theta - \left(|\pi| + \sum_{l \geq 2} \widehat{n}_l\right)^\theta + \theta|\pi| - (2 - 2^\theta) \sum_{l \geq 3} \widehat{n}_l^\theta \geq 0,$$

which we are now going to prove. We shall show that the function $f(x)$ is increasing in $x \geq 0$; then the result follows by showing $f(0) \geq 0$, which is what was proven by Yuan and Bourgain.
We now verify that $f'(x) \geq 0$. By direct computation we see that

$$f'(x) = -\theta\left(x + \sum_{l \geq 2} \widehat{n}_l\right)^{\theta-1} + \theta,$$

so it suffices to prove that

$$(C.4) \qquad 1 \leq \left(x + \sum_{l \geq 2} \widehat{n}_l\right)^{1-\theta},$$

which is indeed true, since $\sum_{i \geq 2} \widehat{n}_i \geq \widehat{n}_2 \geq 1$ holds, by mass conservation. $\qquad\square$

Proof of Lemma 7.1. In this subsection we will take

$$(C.5) \qquad \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}} \quad \text{with} \quad 1 \leq |\boldsymbol{\alpha}| = |\boldsymbol{\beta}| < \infty.$$

Given $u \in \mathbb{Z}^{\mathbb{Z}}$, with $|u| < \infty$, consider the set

$$\{j \neq 0, \quad \text{repeated} \quad |u_j| \text{ times}\},$$

where $D < \infty$ is its cardinality. Define the vector $m = m(u)$ as the reordering of the elements of the set above such that $|m_1| \geq |m_2| \geq \cdots \geq |m_D| \geq 1$. Given $\boldsymbol{\alpha} \neq \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}$, with $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}| < \infty$ we consider $m = m(\boldsymbol{\alpha} - \boldsymbol{\beta})$ and $\widehat{n} = \widehat{n}(\boldsymbol{\alpha} + \boldsymbol{\beta})$. If we denote by D the cardinality of m and N the one of $\widehat{n}$ we have

$$(C.6) \qquad D + \boldsymbol{\alpha}_0 + \boldsymbol{\beta}_0 \leq N$$

and

$$(C.7) \qquad (|m_1|, \ldots, |m_D|, \underbrace{1, \ldots, 1}_{N-D \text{ times}}) \leq (\widehat{n}_1, \ldots \widehat{n}_N).$$

Set

$$\sigma_l = \text{sign}(\boldsymbol{\alpha}_{m_l} - \boldsymbol{\beta}_{m_l}).$$

For every function g defined on $\mathbb{Z}$ we have that

$$\sum_{i \in \mathbb{Z}} g(i)|\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| = g(0)|\boldsymbol{\alpha}_0 - \boldsymbol{\beta}_0| + \sum_{l \geq 1} g(m_l),$$

$$(C.8) \qquad \sum_{i \in \mathbb{Z}} g(i)(\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i) = g(0)(\boldsymbol{\alpha}_0 - \boldsymbol{\beta}_0) + \sum_{l \geq 1} \sigma_l g(m_l).$$

Lemma C.3. *Assume that g defined on $\mathbb{Z}$ is non negative, even and not decreasing on $\mathbb{N}$. Then, if $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$,*

$$\text{(C.9)} \qquad \sum_{i \in \mathbb{Z}} g(i)|\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \leq 2g(m_1) + \sum_{l \geq 3} g(\widehat{n}_l) \,.$$

Proof. By (C.8)

$$
\begin{aligned}
\sum_{i \in \mathbb{Z}} g(i)|\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \;&=\; g(0)|\boldsymbol{\alpha}_0 - \boldsymbol{\beta}_0| + \sum_{l \geq 1} g(m_l) \\
&\leq\; g(1)(\boldsymbol{\alpha}_0 + \boldsymbol{\beta}_0) + 2g(m_1) + \sum_{l \geq 3} g(m_l)
\end{aligned}
$$

and (C.9) follows by (C.6) and (C.7). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We denote as before the momentum by π so by (C.8)

$$\text{(C.10)} \qquad \pi = \sum_{i \in \mathbb{Z}} (\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i)i = \sum_{l} \sigma_l m_l$$

and

$$\text{(C.11)} \qquad \sum_{i} (\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i)i^2 = \sum_{l} \sigma_l m_l^2 \,.$$

Analogously

$$\text{(C.12)} \qquad \sum_{i} |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| = D + |\boldsymbol{\alpha}_0 - \boldsymbol{\beta}_0| \overset{\text{(C.6)}}{\leq} N \,.$$

Finally note that

$$\text{(C.13)} \qquad \sigma_l \sigma_{l'} = -1 \qquad \Longrightarrow \qquad m_l \neq m_{l'} \,.$$

Note that

$$\text{(C.14)} \qquad \boldsymbol{\alpha} \neq \boldsymbol{\beta} \quad \Longrightarrow \quad N \geq 3 \ \text{ or } \ \pi \neq 0 \,,$$

indeed, by mass conservation, $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}| = 1$ therefore if $N = 2$ we get $\boldsymbol{\alpha} - \boldsymbol{\beta} = e_{j_1} - e_{j_2}$ so if $\pi = 0$ we have $\boldsymbol{\alpha} = \boldsymbol{\beta}$. Note also that

$$\text{(C.15)} \qquad \boldsymbol{\alpha} \neq \boldsymbol{\beta} \quad \Longrightarrow \quad D \geq 1 \,,$$

indeed, if $D = 0$ then $\boldsymbol{\alpha}_l - \boldsymbol{\beta}_l = 0$ for every $|l| \geq 1$ and, by mass conservation $\boldsymbol{\alpha}_0 = \boldsymbol{\beta}_0$, contradicting $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$.

Lemma C.4. *Given $\boldsymbol{\alpha} \neq \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}$, with $1 \leq |\boldsymbol{\alpha}| = |\boldsymbol{\beta}| < \infty$ and satisfying (7.1), we have*[32]

$$\text{(C.16)} \qquad |m_1| \leq 20|\pi| + 31 \sum_{l \geq 3} \widehat{n}_l^2 \,.$$

Proof. In the case $D = 1$ by (C.10) $|\pi| = |m_1|$ and (C.16) follows. Let us now consider the case $D = 2$, i.e.

$$\boldsymbol{\alpha} - \boldsymbol{\beta} = \sigma_1 e_{m_1} + \sigma_2 e_{m_2} + (\boldsymbol{\alpha}_0 - \boldsymbol{\beta}_0)e_0 \,.$$

Let us start with the case $\sigma_1 \sigma_2 = 1$. By mass conservation $|\sigma_1 + \sigma_2| = |\boldsymbol{\beta}_0 - \boldsymbol{\alpha}_0| = 2$. By (C.12) $N \geq 4$. Then conditions (7.1) and (C.12) imply that

$$m_1^2 + m_2^2 \leq 20 + 10|\boldsymbol{\alpha}_0 - \boldsymbol{\beta}_0| = 40 \,.$$

---

[32]Note that by (C.14) the r.h.s. of (C.16) is at least 20.

Then

$$|m_1| \leq \sqrt{40} \leq \frac{\sqrt{40}}{2} \sum_{\ell=3}^{N} \widehat{n}_\ell^2$$

since $N \geq 4$ and $\widehat{n}_\ell \geq 1$. When $\sigma_1 \sigma_2 = -1$ we have $m_1 \neq m_2$, $|\pi| = |m_1 - m_2| \geq 1$ and by mass conservation $\boldsymbol{\alpha}_0 - \boldsymbol{\beta}_0 = 0$. Then

$$(|m_1| + |m_2|)(|m_1| - |m_2|) = m_1^2 - m_2^2 \leq 20 \,.$$

If $|m_1| > |m_2|$ then

(C.17) $$|m_1| \leq 20 \leq 20|\pi| \,.$$

Otherwise $m_1 = -m_2$ and, therefore, $|\pi| = 2|m_1|$, completing the proof in the case $D = 2$.

Let us now consider the case $D \geq 3$. By (7.1),(C.11) and (C.12)

$$\begin{aligned}
m_1^2 + \sigma_1 \sigma_2 m_2^2 &\leq& 10N + \sum_{l=3}^{D} m_l^2 \leq 10N + \sum_{l=3}^{N} \widehat{n}_l^2 \\
&=& 20 + \sum_{l=3}^{N} (10 + \widehat{n}_l^2) \leq 20 + 11 \sum_{l=3}^{N} \widehat{n}_l^2 \overset{N \geq 3}{\leq} 31 \sum_{l=3}^{N} \widehat{n}_l^2 \,.
\end{aligned}$$

If $\sigma_1 \sigma_2 = 1$ then

$$|m_1|, |m_2| \leq \sqrt{31 \sum_{l \geq 3} \widehat{n}_l^2}.$$

If $\sigma_1 \sigma_2 = -1$

$$(|m_1| + |m_2|)(|m_1| - |m_2|) = m_1^2 - m_2^2 \leq 31 \sum_{l \geq 3} \widehat{n}_l^2.$$

Now, if $|m_1| \neq |m_2|$ then

$$|m_1| + |m_2| \leq 31 \sum_{l \geq 3} \widehat{n}_l^2.$$

Conversely, if $|m_1| = |m_2|$, by (C.13), $m_1 \neq m_2$, hence $m_1 = -m_2$. By substituting this relation into (C.10), we have

$$2|m_1| \leq |\pi| + \sum_{l \geq 3} |m_l| \leq |\pi| + \sum_{l \geq 3} \widehat{n}_l^2 \,,$$

concluding the proof. □

Conclusion of the proof of Lemma 7.1. As above, given $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{\mathbb{Z}}$, with $1 \leq |\boldsymbol{\alpha}| = |\boldsymbol{\beta}| < \infty$ we consider $m = m(\boldsymbol{\alpha} - \boldsymbol{\beta})$ and $\widehat{n} = \widehat{n}(\boldsymbol{\alpha} + \boldsymbol{\beta})$. Note that $N := |\boldsymbol{\alpha} + \boldsymbol{\beta}| \geq 2$.

We have[33]

$$\begin{aligned}
\sum_i |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \langle i \rangle^{\theta/2} &\overset{(C.9)}{\leq}& 2|m_1|^{\frac{\theta}{2}} + \sum_{l \geq 3} \widehat{n}_l^{\frac{\theta}{2}} \\
&\overset{(C.16)}{\leq}& 2 \left(20|\pi| + 31 \sum_{l \geq 3} \widehat{n}_l^2 \right)^{\frac{\theta}{2}} + \sum_{l \geq 3} \widehat{n}_l^{\frac{\theta}{2}} \\
&\leq& 2(20|\pi|)^{\frac{\theta}{2}} + 2(31)^{\frac{\theta}{2}} \sum_{l \geq 3} \widehat{n}_l^{\theta} + \sum_{l \geq 3} \widehat{n}_l^{\frac{\theta}{2}} \\
(C.18) &\leq& \frac{13}{1 - \theta} \left((1 - \theta)|\pi| + (2 - 2^{\theta}) \left(\sum_{l \geq 3} \widehat{n}_l^{\theta} \right) \right),
\end{aligned}$$

---

[33]Using that for $x, y \geq 0$ and $0 \leq c \leq 1$ we get $(x + y)^c \leq x^c + y^c$.

using that $1 - \theta \leq 2 - 2^\theta$ for $0 \leq \theta \leq 1$. Then by Lemma 6.1 and (C.18) we get

$$\sum_i |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \langle i \rangle^{\theta/2} \leq \frac{13}{1 - \theta} \left((1 - \theta)|\pi| + \sum_i \langle i \rangle^\theta (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) + \theta|\pi| - 2\widehat{n}_1^\theta \right)$$

$$\leq \frac{13}{1 - \theta} \left[\sum_i \langle i \rangle^\theta (\boldsymbol{\alpha}_i + \boldsymbol{\beta}_i) + |\pi| - 2\langle j \rangle^\theta \right],$$

proving (7.2).

Let us now prove (7.3) passing to the logarithm. We have

(C.19)
$$\sum_i \ln(1 + |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \langle i \rangle)$$
$$= \sum_{|i| \leq 1} \ln(1 + |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i|) + \sum_{|i| \geq 2} \ln(1 + |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i||i|)$$
$$\leq 3\ln(1 + N) + \sum_{|i| \geq 2} \ln(1 + |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i||i|)$$
$$\leq 3\ln 2 + 3\ln N + \frac{3}{2} \sum_{|i| \geq 2} |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \ln|i| ,$$

using that $1 + cx \leq \frac{3}{2}x^c$ for $c \geq 1$, $x \geq 2$. If $\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i = 0$ for every $|i| \geq 2$ then (7.3) follows. Assume now that $\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i \neq 0$ for some $|i| \geq 2$. By (C.14) we have

(C.20)
$$N \geq 3 \quad \text{or} \quad \pi \neq 0 .$$

We claim that, when $N \geq 3$,

(C.21)
$$\ln \left(\sum_{l=3}^N \widehat{n}_l^2 \right) \leq \ln N + \sum_{l=3}^N \ln \widehat{n}_l^2 .$$

Let $\mathcal{S} := \{3 \leq l \leq N, \text{ s.t. } \widehat{n}_l \geq 2\}$. If $\mathcal{S} = \emptyset$ we have the equality in (C.21). Otherwise $\sum_{l \in \mathcal{S}} \widehat{n}_l^2 \geq 4$ and[34]

$$\ln \left(\sum_{l=3}^N \widehat{n}_l^2 \right) \leq \ln \left(N + \sum_{l \in \mathcal{S}} \widehat{n}_l^2 \right) \leq \ln N + \sum_{l \in \mathcal{S}} \ln \widehat{n}_l^2 ,$$

proving (C.21).
We claim that

(C.22)
$$\ln \left(20|\pi| + 31 \sum_{l=3}^N \widehat{n}_l^2 \right) \leq \ln(1 + |\pi|) + \ln N + \sum_{l=3}^N \ln \widehat{n}_l^2 + \ln 20 + \ln 31 .$$

Indeed consider first the case $\pi = 0$, then $N \geq 3$ by (C.20) and (C.22) follows by (C.21). Consider now the case $|\pi| \geq 1$. If $N < 3$ (C.22) follows (there is no sum). If $N \geq 3$ we have[35]

$$\ln \left(20|\pi| + 31 \sum_{l=3}^N \widehat{n}_l^2 \right) \leq \ln(20|\pi|) + \ln \left(31 \sum_{l=3}^N \widehat{n}_l^2 \right)$$

$$\leq \ln(|\pi|) + \ln \left(\sum_{l=3}^N \widehat{n}_l^2 \right) + \ln 20 + \ln 31 .$$

Recalling (C.21) this complete the proof of (C.22).

---

[34] Use that $\ln(x + y) \leq \ln x + \ln y$ if $x, y \geq 2$.
[35] Recall footnote 34.

Let us continue the proof of (7.3). Set $g(i) := 0$ if $|i| \leq 1$ and $g(i) := \ln|i|$ if $|i| \geq 2$ and apply (C.9) to (C.19); we get

$$\sum_{|i| \geq 2} |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i| \ln|i| \leq 2\ln|m_1| + \sum_{l \geq 3} \ln|\widehat{n}_l|$$

$$\overset{(C.16)}{\leq} 2\ln\left(20|\pi| + 31\sum_{l \geq 3} \widehat{n}_l^2\right) + \sum_{l \geq 3} \ln\widehat{n}_l$$

$$\overset{(C.22)}{\leq} 2\ln(1 + |\pi|) + 2\ln N + 5\sum_{l=3}^{N} \ln\widehat{n}_l + 16\,.$$

Inserting in (C.19) we obtain

$$\sum_i \ln(1 + |\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i|\langle i\rangle) \leq 3\ln(1 + |\pi|) + 6\ln N + \frac{15}{2}\sum_{l=3}^{N} \ln\widehat{n}_l + 27\,.$$

concluding the proof of (7.3). $\square$

Proof of Lemma 7.2. First of all we note that

$$\sum_i f_i(|\ell_i|) = \sum_{i \text{ s.t. } \ell_i \neq 0} f_i(|\ell_i|)$$

since $f_i(0) = 0$. We have that[36]

$$f_i(x) \leq -\frac{\sigma}{C_*}\langle i\rangle^{\frac{\theta}{2}}x + 2\ln(x) + (2 + q)\ln\langle i\rangle + 1\,, \qquad \forall\, x \geq 1\,.$$

We have that

$$\max_{x \geq 1}\left(-\frac{\sigma}{C_*}\langle i\rangle^{\frac{\theta}{2}}x + 2\ln(x)\right) = \begin{cases} -\dfrac{\sigma}{C_*}\langle i\rangle^{\frac{\theta}{2}} & \text{if } \langle i\rangle \geq i_0\,, \\[2ex] -2 + 2\ln\dfrac{2C_*}{\sigma} - \theta\ln\langle i\rangle & \text{if } \langle i\rangle < i_0\,, \end{cases}$$

where

$$i_0 := \left(\frac{2C_*}{\sigma}\right)^{\frac{2}{\theta}}\,,$$

since the maximum is achieved for $x = 1$ if $\langle i\rangle \geq i_0$ and $x = \frac{2C_*}{\sigma\langle i\rangle^{\theta/2}}$ if $\langle i\rangle < i_0$. Note that $i_0 \geq e$. Then we get

$$\sum_i f_i(|\ell_i|) = \sum_{i \text{ s.t. } \ell_i \neq 0} f_i(|\ell_i|) \leq$$

$$\sum_{\langle i\rangle \geq i_0 \text{ s.t. } \ell_i \neq 0}\left((2 + q)\ln\langle i\rangle + 1 - \frac{\sigma}{C_*}\langle i\rangle^{\frac{\theta}{2}}\right) + \sum_{\langle i\rangle < i_0}\left(2\ln\frac{2C_*}{\sigma} + \left(2 + q - \theta\right)\ln\langle i\rangle\right)\,.$$

We immediately have that

$$\sum_{\langle i\rangle < i_0}\left(2\ln\frac{2C_*}{\sigma} + \left(2 + q - \theta\right)\ln\langle i\rangle\right) \leq 3i_0\left(2\ln\frac{2C_*}{\sigma} + (2 + q)\ln i_0\right)$$

$$= 3\left(2 + \frac{2}{\theta}(2 + q)\right)\left(\frac{2C_*}{\sigma}\right)^{\frac{2}{\theta}}\ln\frac{2C_*}{\sigma}\,.$$

---

[36]Using that $\ln(1 + y) \leq 1 + \ln y$ for every $y \geq 1$.

Moreover, in the case $\langle i \rangle \geq i_0 \geq e$,

$$(2+q)\ln\langle i\rangle + 1 - \frac{\sigma}{C_*}\langle i\rangle^{\frac{\theta}{2}} \leq (2+q+1)\ln\langle i\rangle - \frac{\sigma}{C_*}\langle i\rangle^{\frac{\theta}{2}} = \frac{2}{\theta}(2+q+1)\left(\ln\langle i\rangle^{\frac{\theta}{2}} - 2\mathfrak{C}\langle i\rangle^{\frac{\theta}{2}}\right)$$

where

$$\mathfrak{C} := \frac{\sigma\theta}{4C_*(2+q+1)} < 1.$$

Therefore

$$S_* := \sum_{\langle i\rangle \geq i_0 \text{ s.t. } \ell_i \neq 0}\left((2+q)\ln\langle i\rangle + 1 - \frac{\sigma}{C_*}\langle i\rangle^{\frac{\theta}{2}}\right)$$

satisfies

$$S_* \leq \sum_{\langle i\rangle \geq i_0 \text{ s.t. } \ell_i \neq 0}\frac{2}{\theta}(2+q+1)\left(\ln\langle i\rangle^{\frac{\theta}{2}} - 2\mathfrak{C}\langle i\rangle^{\frac{\theta}{2}}\right).$$

We have that[37]

$$\ln\langle i\rangle^{\frac{\theta}{2}} - 2\mathfrak{C}\langle i\rangle^{\frac{\theta}{2}} \leq -\mathfrak{C}\langle i\rangle^{\frac{\theta}{2}}, \qquad \text{when} \qquad \langle i\rangle \geq i_* := \left(\frac{2}{\mathfrak{C}}\ln\frac{1}{\mathfrak{C}}\right)^{\frac{2}{\theta}}.$$

Note that

$$i_\sharp \geq \max\{i_0, i_*\}.$$

Therefore

$$
\begin{aligned}
S_* &\leq \frac{2}{\theta}(2+q+1)\left(\sum_{\langle i\rangle < i_\sharp}\ln\langle i\rangle^{\frac{\theta}{2}} - \sum_{\langle i\rangle \geq i_\sharp \text{ s.t. } \ell_i \neq 0}\left(\mathfrak{C}\langle i\rangle^{\frac{\theta}{2}}\right)\right) \\
&\leq (2+q+1)\left(3i_\sharp\ln i_\sharp - \frac{2\mathfrak{C}}{\theta}M_\ell^{\frac{\theta}{2}}\right).
\end{aligned}
$$

where

$$M_\ell := \max\{|i| \geq i_\sharp, \text{ s.t. } \ell_i \neq 0\}$$

and $M_\ell := 0$ if $|\ell_i| = 0$ for every $|i| \geq i_\sharp$. In conclusion we get

$$
\begin{aligned}
\sum_i f_i(|\ell_i|) &\leq 3\left(2+\frac{2}{\theta}(2+q)\right)\left(\frac{2C_*}{\sigma}\right)^{\frac{2}{\theta}}\ln\frac{2C_*}{\sigma} + (2+q+1)\left(3i_\sharp\ln i_\sharp - \frac{2\mathfrak{C}}{\theta}M_\ell^{\frac{\theta}{2}}\right) \\
&\leq 6(q+3)i_\sharp\ln i_\sharp - \frac{\sigma}{2C_*}M_\ell^{\frac{\theta}{2}} \\
&\leq 7(q+3)i_\sharp\ln i_\sharp - \frac{\sigma}{2C_*}\left(\widehat{n}_1(\ell)\right)^{\frac{\theta}{2}},
\end{aligned}
$$

noting that $\widehat{n}_1(\ell) = M_\ell$ if $M_\ell \neq 0$, otherwise $\widehat{n}_1(\ell) < i_\sharp$, and, therefore,

$$\frac{\sigma}{2C_*}\left(\widehat{n}_1(\ell)\right)^{\frac{\theta}{2}} < \frac{\sigma}{2C_*}i_\sharp^{\frac{\theta}{2}} \leq (q+3)i_\sharp\ln i_\sharp$$

$\square$

---

[37]Using that, for every fixed $0 < \mathfrak{C} \leq 1$, we have $\mathfrak{C}x \geq \ln x$ for every $x \geq \frac{2}{\mathfrak{C}}\ln\frac{1}{\mathfrak{C}}$.

Proof of Lemma 4.1 . For $\ell \in \mathbb{Z}^{\mathbb{Z}}$ with $0 < |\ell| < \infty$ we define

$$\mathcal{R}_\ell := \left\{ \omega \in \Omega_q \; : \; |\omega \cdot \ell| \leq \frac{\gamma}{1 + |\ell_0|^{\mu_1}} \prod_{n \neq 0} \frac{1}{(1 + |\ell_n|^{\mu_1}|n|^{\mu_2+q})} \right\}$$

$-$ if ℓ is such that $\ell_n = 0 \; \forall n \neq 0$ then

$$\mu(\mathcal{R}_\ell) = \frac{\gamma}{1 + |\ell_0|^{\mu_1}}$$

$-$ Otherwise: let $s = s(\ell) > 0$ be the smallest positive index i such that $|\ell_i| + |\ell_{-i}| \neq 0$ and $S = S(\ell)$ be the biggest. Then we have[38]

$$\mu(\mathcal{R}_\ell) \leq \frac{\gamma s^q}{(1 + |\ell_0|^{\mu_1})} \prod_{n \neq 0} \frac{1}{(1 + |\ell_n|^{\mu_1}|n|^{\mu_2+q})}.$$

Let us write

$$\frac{1}{1 + |\ell_0|^{\mu_1}} \prod_{n \neq 0} \frac{1}{(1 + |\ell_n|^{\mu_1}|n|^{\mu_2+q})} = \frac{1}{1 + |\ell_0|^{\mu_1}} \prod_{n > 0} \frac{1}{(1 + |\ell_n|^{\mu_1}|n|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-n}|^{\mu_1}|n|^{\mu_2+q})}$$

$$= \frac{1}{1 + |\ell_0|^{\mu_1}} \prod_{s(\ell) \leq n \leq S(\ell)} \frac{1}{(1 + |\ell_n|^{\mu_1}|n|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-n}|^{\mu_1}|n|^{\mu_2+q})}$$

Now

$$(C.23) \qquad \mu(\Omega_q \setminus \mathsf{D}_{\gamma,q}) \leq \sum_\ell \mu(\mathcal{R}_\ell) = \sum_{\ell_0} \frac{\gamma}{1 + |\ell_0|^{\mu_1}}$$

$$(C.24) \qquad + \sum_{s > 0} \sum_{\ell:s(\ell)=S(\ell)=s} \frac{1}{1 + |\ell_0|^{\mu_1}} \frac{\gamma s^q}{|\ell_s|(1 + |\ell_s|^{\mu_1}|s|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-s}|^{\mu_1}|s|^{\mu_2+q})}$$

$$(C.25) \qquad + \sum_{0 < s < S} \sum_{\substack{\ell:s(\ell)=s, \\ S(\ell)=S}} \frac{\gamma s^q}{1 + |\ell_0|^{\mu_1}} \prod_{s \leq n \leq S} \frac{1}{(1 + |\ell_n|^{\mu_1}|n|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-n}|^{\mu_1}|n|^{\mu_2+q})}.$$

Let us estimate (C.24)

$$\sum_{s > 0} \sum_{\ell_0 \in \mathbb{Z}} \frac{1}{1 + |\ell_0|^{\mu_1}} \sum_{\substack{\ell_s, \ell_{-s} \in \mathbb{Z} \\ |\ell_s| + |\ell_{-s}| > 0}} \frac{\gamma s^q}{(1 + |\ell_s|^{\mu_1}|s|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-s}|^{\mu_1}|s|^{\mu_2+q})}$$

$$\leq c(\mu_1)\gamma \sum_{s > 0} s^q \sum_{\substack{\ell_s, \ell_{-s} \in \mathbb{Z} \\ |\ell_s| + |\ell_{-s}| > 0}} \frac{1}{(1 + |\ell_s|^{\mu_1}|s|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-s}|^{\mu_1}|s|^{\mu_2+q})}$$

Now since

$$\sum_{h=1}^{\infty} \frac{1}{(1 + h^{\mu_1}|n|^{\mu_2+q})} \leq \sum_{h=1}^{\infty} \frac{1}{h^{\mu_1}|n|^{\mu_2+q}} \leq \frac{c(\mu_1)}{|n|^{\mu_2+q}}$$

we have

$$\sum_{h \in \mathbb{Z}} \frac{1}{(1 + |h|^{\mu_1}|n|^{\mu_2+q})} \leq 1 + \frac{2c(\mu_1)}{|n|^{\mu_2+q}}.$$

Then we have

$$\sum_{\substack{\ell_s, \ell_{-s} \in \mathbb{Z} \\ |\ell_s| + |\ell_{-s}| > 0}} \frac{1}{(1 + |\ell_s|^{\mu_1}|s|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-s}|^{\mu_1}|s|^{\mu_2+q})} \leq \frac{c_1(\mu_1)}{|s|^{\mu_2+q}}$$

---

[38]Assume, e.g. that $\ell_s \neq 0$, then $|\partial_{\xi_s} \omega \cdot \ell| \geq s^{-q}$.

and consequently (C.24) is bounded by

$$c_2(\mu_1)\gamma \sum_{s>0} |s|^b \leq c_3(\mu_1, \mu_2)\gamma.$$

Regarding the third line in (C.23), we note that for all n we have

$$\sum_{\ell_n, \ell_{-n} \in \mathbb{Z}} \frac{1}{(1 + |\ell_n|^{\mu_1}|n|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-n}|^{\mu_1}|n|^{\mu_2+q})} \leq \left(1 + 2\frac{c(\mu_1)}{|n|^{\mu_2+q}}\right)^2.$$

Hence

$$\sum_{\substack{\ell:s(\ell)=s, \\ S(\ell)=S}} \frac{1}{1 + |\ell_0|^{\mu_1}} \prod_{s \leq n \leq S} \frac{1}{(1 + |\ell_n|^{\mu_1}|n|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-n}|^{\mu_1}|n|^{\mu_2+q})}$$

$$= \sum_{\ell_0 \in \mathbb{Z}} \frac{1}{1 + |\ell_0|^{\mu_1}} \times \sum_{\substack{\ell_s, \ell_{-s} \in \mathbb{Z} \\ |\ell_s|+|\ell_{-s}|>0}} \frac{1}{(1 + |\ell_s|^{\mu_1}|s|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-s}|^{\mu_1}|s|^{\mu_2+q})}$$

$$\times \sum_{\substack{\ell_S, \ell_{-S} \in \mathbb{Z} \\ |\ell_S|+|\ell_{-S}|>0}} \frac{1}{(1 + |\ell_S|^{\mu_1}|S|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-S}|^{\mu_1}|S|^{\mu_2+q})}$$

$$\times \prod_{s<n<S} \sum_{\ell_n, \ell_{-n} \in \mathbb{Z}} \frac{1}{(1 + |\ell_n|^{\mu_1}|n|^{\mu_2+q})} \frac{1}{(1 + |\ell_{-n}|^{\mu_1}|n|^{\mu_2+q})}$$

$$\leq \frac{c_4(\mu_1)}{s^{\mu_2+q}S^{\mu_2+q}} \prod_{s<n<S} \left(1 + 2\frac{c(\mu_1)}{|n|^{\mu_2+q}}\right)^2$$

$$\leq \frac{c_4(\mu_1)}{s^{\mu_2+q}S^{\mu_2+q}} \exp\left(\sum_{n\geq 1} \ln\left(1 + 2\frac{c(\mu_1)}{|n|^{\mu_2+q}}\right)^2\right)$$

$$\leq \frac{c_5(\mu_1, \mu_2)}{s^{\mu_2+q}S^{\mu_2+q}}.$$

Then, multiplying by γs^q and taking the $\sum_{0<s<S}$, we have that also (C.25) is bounded by some constant $C_{\mathtt{meas}}(\mu_1, \mu_2)\gamma$. $\hfill\square$

References

[Bam99a] D. Bambusi. Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. Math. Z., 230(2):345–387, 1999.

[Bam99b] D. Bambusi. On long time stability in Hamiltonian perturbations of nonresonant linear PDE's. Nonlinearity, 12:823–850, 1999.

[Bam03] D. Bambusi. Birkhoff normal form for some nonlinear PDEs. Comm. Math. Phys., 234(2):253–285, 2003.

[BBP13] M. Berti, L. Biasco, and M. Procesi. KAM theory for the Hamiltonian derivative wave equation. Annales Scientifiques de l'ENS, 46(2):299–371, 2013.

[BD18] M. Berti and J.M. Delort. Almost global existence of solutions for capillarity-gravity water waves equations with periodic spatial boundary conditions. Springer, 2018.

[BDG10] Luca Biasco and Laura Di Gregorio. A Birkhoff-Lewis type theorem for the nonlinear wave equation. Arch. Ration. Mech. Anal., 196(1):303–362, 2010.

[BDGS07] D. Bambusi, J.-M. Delort, B. Grébert, and J. Szeftel. Almost global existence for Hamiltonian semilinear Klein-Gordon equations with small Cauchy data on Zoll manifolds. Comm. Pure Appl. Math., 60(11):1665–1690, 2007.

[BFG88] G. Benettin, J. Fröhlich, and A. Giorgilli. A Nekhoroshev-type theorem for Hamiltonian systems with infinitely many degrees of freedom. Comm. Math. Phys., 119(1):95–108, 1988.

[BFG18] J. Bernier, E. Faou, and B. Grébert. Rational normal forms and stability of small solutions to nonlinear schrödinger equations, 2018. arXiv:1812.11414.

[BFN15]   A. Bounemoura, B. Fayad, and L. Niederman. Double exponential stability for generic real-analytic elliptic equilibrium points. 2015. Preprint ArXiv : arxiv.org/abs/1509.00285.

[BG03]   D. Bambusi and B. Grébert. Forme normale pour NLS en dimension quelconque. C. R. Math. Acad. Sci. Paris, 337(6):409–414, 2003.

[BG06]   D. Bambusi and B. Grébert. Birkhoff normal form for partial differential equations with tame modulus. Duke Math. J., 135(3):507–567, 2006.

[Bou96a] J. Bourgain. Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations. Geom. Funct. Anal., 6(2):201–230, 1996.

[Bou96b] J. Bourgain. On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE. Internat. Math. Res. Notices, (6):277–304, 1996.

[Bou05]   J. Bourgain. On invariant tori of full dimension for 1D periodic NLS. J. Funct. Anal., 229(1):62–94, 2005.

[CKS$^+$10] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation. Invent. Math., 181(1):39–113, 2010.

[CLSY]   H. Cong, J. Liu, Y. Shi, and X. Yuan. The stability of full dimensional kam tori for nonlinear schrödinger equation. preprint 2017, arXiv:1705.01658.

[CMW]   H. Cong, L. Mi, and P. Wang. A Nekhoroshev type theorem for the derivative nonlinear Schrödinger equation. preprint 2018.

[Del12]   J. M. Delort. A quasi-linear birkhoff normal forms method. application to the quasi-linear klein-gordon equation on S^1. Astérisque, 341, 2012.

[DS04]   J.-M. Delort and J. Szeftel. Long-time existence for small data nonlinear Klein-Gordon equations on tori and spheres. Int. Math. Res. Not., (37):1897–1966, 2004.

[DS06]   J.-M. Delort and J. Szeftel. Bounded almost global solutions for non Hamiltonian semi-linear Klein-Gordon equations with radial data on compact revolution hypersurfaces. Ann. Inst. Fourier (Grenoble), 56(5):1419–1456, 2006.

[FG13]   E. Faou and B. Grébert. A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus. Anal. PDE, 6(6):1243–1262, 2013.

[FGL13]   E. Faou, L. Gauckler, and C. Lubich. Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus. Comm. Partial Differential Equations, 38(7):1123–1140, 2013.

[FI]   R. Feola and F. Iandoli. Long time existence for fully nonlinear nls with small cauchy data on the circle. preprint 2018, arXiv:1806.03437.

[GHP16]   M. Guardia, E. Haus, and M. Procesi. Growth of Sobolev norms for the analytic NLS on $\mathbb{T}^2$. Adv. Math., 301:615–692, 2016.

[GK15]   M. Guardia and V. Kaloshin. Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation. J. Eur. Math. Soc. (JEMS), 17(1):71–149, 2015.

[GT12]   Benoî t Grébert and Laurent Thomann. Resonant dynamics for the quintic nonlinear Schrödinger equation. Ann. Inst. H. Poincaré Anal. Non Linéaire, 29(3):455–477, 2012.

[Gua14]   Marcel Guardia. Growth of Sobolev norms in the cubic nonlinear Schrödinger equation with a convolution potential. Comm. Math. Phys., 329(1):405–434, 2014.

[Han14]   Zaher Hani. Long-time instability and unbounded Sobolev orbits for some periodic nonlinear Schrödinger equations. Arch. Ration. Mech. Anal., 211(3):929–964, 2014.

[HP17]   E. Haus and M. Procesi. KAM for beating solutions of the quintic NLS. Comm. Math. Phys., 354(3):1101–1132, 2017.

[KP10]   S. Kuksin and G. Perelman. Vey theorem in infinite dimensions and its application to kdv. Discrete Contin. Dyn. Syst., 27:1–24, 2010.

[MG95]   A. Morbidelli and A. Giorgilli. Superexponential stability of KAM tori. J. Statist. Phys., 78(5-6):1607–1617, 1995.

[MSW18]   L. Mi, Y. Sun, and P Wang. Long time stability of plane wave solutions to the cubic NLS on torus, 2018. preprint.

[Nik86]   N.V. Nikolenko. The method of Poincaré normal forms in problems of integrability of equations of evolution type. Russian Math. Surveys, 41:5:63–114, 1986.

[YZ14]   Xiaoping Yuan and Jing Zhang. Long time stability of Hamiltonian partial differential equations. SIAM J. Math. Anal., 46(5):3176–3222, 2014.

Università degli Studi Roma Tre
E-mail address: `biasco@mat.uniroma3.it`

Centro di Ricerca Matematica E. De Giorgi, Scuola Normale Superiore di Pisa
E-mail address: `jessica.massetti@sns.it`

Università degli Studi Roma Tre
E-mail address: `procesi@mat.uniroma3.it`