



# About Linearization of Infinite-Dimensional Hamiltonian Systems

Michela Procesi<sup>1</sup>, Laurent Stolovitch<sup>2</sup> 

<sup>1</sup> Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, L.go Murialdo, Rome, Italy  
E-mail: [procesi@mat.uniroma3.it](mailto:procesi@mat.uniroma3.it)

<sup>2</sup> CNRS and Laboratoire J.-A. Dieudonné U.M.R. 7351, Université Côte d'Azur, Parc Valrose, 06108 Nice Cedex 02, France. E-mail: [stolo@unice.fr](mailto:stolo@unice.fr)

Received: 25 August 2021 / Accepted: 2 April 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

**Abstract:** This article is concerned with analytic Hamiltonian dynamical systems in infinite dimension in a neighborhood of an elliptic fixed point. Given a quadratic Hamiltonian, we consider the set of its analytic higher order perturbations. We first define the subset of elements which are formally symplectically conjugated to a (formal) Birkhoff normal form. We prove that if the quadratic Hamiltonian satisfies a Diophantine-like condition and if such a perturbation is formally symplectically conjugated to the quadratic Hamiltonian, then it is also analytically symplectically conjugated to it. Of course what is an analytic symplectic change of variables depends strongly on the choice of the phase space. Here we work on periodic functions with Gevrey regularity.

## 1. Introduction

In finite dimension, studying the behavior of the orbits of a vector field (or of diffeomorphism) nearby a fixed point is a fundamental and classical problem. The very first natural step into this understanding is to compare the dynamical system with its linearization at the fixed point. This is done by trying to transform the dynamical system into its linear part by a change of coordinates. There are formal obstructions to do so, called resonances. Hence, in general, one can merely expect the dynamical system to be transformed into a *normal form*, that is supposed to capture effect the very nonlinearities, through a formal change of coordinates. It was understood by the end of the 19th century that if the convex hull of the eigenvalues of the linear part does not contain the origin (one says then that the linear part is in the “Poincaré domain”), and if an higher order analytic perturbation is formally conjugate to the linear part, then it is also analytically so. When the linear part does not satisfy this property, then one has so-called “small divisors” that may forbid the transformation to be analytic. It was a major step forward

---

Research of L. Stolovitch was supported by the French government, through the UCAJEDI Investments in the Future project managed by the National Research Agency (ANR) with the Reference Number ANR-15-IDEX-01.

made by C.L. Siegel [Sie42], followed by H. Rüssmann [Rue77] (for diffeomorphisms) and by A.D. Brjuno [Bru72] (for vector fields) who devised a sufficient “small divisors condition” ensuring the analyticity of a linearizing transformation as soon as there exists a formal one. Linearizing (resp. Normalizing) problems for diffeomorphisms were devised by J. Pöschel [Pös86] and for commuting families by the second author [Sto15] (resp. [Sto00]). By the end of the 70’s, it became clear to few people that some PDE’s problems could be translated into an infinite dimensional dynamical systems to which one would have tried to apply methods of finite dimension. In particular, we mention the work by E. Zehnder [Zeh77] and V. Nikolenko [Nik86] who gave results similar to finite dimensional ones. It happens that the “small divisors condition” they required are too strong and are rarely satisfied. Furthermore, in general, the notion of formal normal form and formal change of variables should be clarified (for instance if one defines formal polynomials and formal power series it is not in general true that this space has a Poisson algebra structure). Nevertheless, in some very peculiar situation, this problem can be handled [BS20].

Starting from the mid 80’, there has been a lot of interest in studying long time behavior of solutions of PDEs. For those PDEs which can be considered as Hamiltonian (infinite dimensional) dynamical systems related to a symplectic structure, one natural way to proceed is to prove the existence of finite dimensional invariant tori in the phase space. This usually implies the existence of quasi-periodic solutions, which are defined for all time. Lot of progresses has been done on the problem of extending KAM theory to PDEs. This circle of problems are very related, though distinct, to the ones solved in this article. Indeed, here one considers a dynamical system close to an elliptic fixed point with the purpose of conjugating it to its most simple *normal form* : its linear part at the fixed point. On the other hand, in KAM theory, one looks for the existence of a *finite dimensional invariant flat torus* on which the dynamics is the linear translation by a diophantine frequency. There is by now a wide literature dealing the subject related to semi-linear PDEs, starting from [K88,Pös90,KP96,Way90,CW93], (for instance, see [EK10,GYX,PP16,BKM18,Y21] for more recent treatments). It has been early understood that these results might be seen through elaborated versions of “Nash–Moser” theorem see for instance [Bou98,BB15,BCP,CM18]. We finally mention [FGPr,BBHM,BM21,FG] for the case of fully-nonlinear PDEs. See also [BMP21,CY21] and references therein for infinite-dimensional tori.

Birkhoff normal form (BNF) methods have been used in order to prove long time existence results and control of Sobolev norms for many classes of evolution PDEs close to an elliptic fixed point. Loosely speaking the point is to canonically transform  $H$  into a Hamiltonian Normal form which depends only on the actions plus a remainder term whose the Taylor polynomial, at the origin is of degree  $N + 1$ . If one achieves this then initial data which are  $\delta$ -small (with respect to the norm on the phase space) stay small (in the same norm) for times of order  $\delta^{-N}$ . A more precise formulation is given in the *Strategy section* below. Of course in the infinite dimensional setting this stability time depends strongly on the choice of the phase space as well as on the nature of the non-linear terms. A further problem is that in general it is not obvious that one can perform even one step of this procedure, indeed the generating function of the desired change of variables is a *formal* polynomial which in infinite dimension is not necessarily analytic. This is a particularly difficult problem in the case of PDEs with derivatives in the nonlinearity.

Let us briefly describe some of the literature. Regarding applications to PDEs (and particularly the NLS) the first results were given in [Bou96a] by Bourgain, who proved

that for any  $\mathbb{N}$  there exists  $p = p(\mathbb{N})$  such that small initial data 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 super-analytic initial data stay small in analytic norm for sub-exponentially long times. Following the strategy proposed in [Bam03] for the Klein–Gordon equation Bambusi and Grébert in [BG03] first considered NLS equations on  $\mathbb{T}^d$  and then, in [BG06], proved polynomial bounds for a class of *tame-modulus* PDEs. 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 sub-exponential bounds on the stability time for classes of NLS equations in  $\mathbb{T}^d$ . In [BMP18] the first author with Biasco and Massetti studied an abstract Birkhoff normal form on *sequence spaces* proving sub-exponential stability times for Gevrey regular initial data. A similar result was proved in [CMW]. An interesting feature of the last three papers is that instead on relying on tameness properties they use the fact that the equations they study have some symmetries, namely they are gauge and translation invariant (actually in [BMP18] the translation invariance condition is weakened).

All the preceding results regard semi-linear PDEs. Regarding equations with derivatives in the nonlinearity, the first results were in [YZ14] for the semi-linear case. Then we mention [Del12, D15] for the Klein–Gordon equation, [BD18] for the water waves and [FI18] for the reversible NLS equation. Recently, Feola and Iandoli, [FI20] prove polynomial lower bounds for the stability times of Hamiltonian NLS equations with two derivatives in the nonlinearity. In the context of infinite chains with a finite range coupling, similar considerations can be done and we mention [BFG88].

*1.1. Statements.* We study Hamiltonians on infinite dimensional sequence spaces, which are higher order (*M-regular*) analytic perturbations of quadratic Hamiltonians nearby an elliptic fixed point (i.e a zero) and satisfying the *Momentum conservation* property, namely they are formally translation invariant, see Definition 9.

We first show that the space  $\mathcal{F}$  of formal Hamiltonians in infinite variables  $u = (u_j)_{j \in \mathbb{Z}}$  satisfying this *Momentum conservation property* is well defined and closed w.r.t Poisson brackets, then we define a scaling degree (which is the homogeneity degree minus two, see Definition 2.3, so that the degree of the Poisson bracket of two functions is the sum of the respective degrees) so that  $\mathcal{F}$  has a natural filtered Lie algebra structure. Thus  $\mathcal{F}$  is decomposed in homogeneous components  $\mathcal{F}^d$  and we define  $\widehat{\mathcal{F}}^{\geq d} := \widehat{\bigoplus}_{h \geq d} \mathcal{F}^h$ .

Given a rationally independent  $\omega \in \mathbb{R}^{\mathbb{Z}}$ , namely such that all non-trivial finite rational combinations of  $\omega$  are non zero, we consider the affine space  $D_\omega + \widehat{\mathcal{F}}^{\geq 1}$  of formal Hamiltonians of the form

$$H = D_\omega + P, \quad D_\omega = \sum_{j \in \mathbb{Z}} \omega_j |u_j|^2, \quad P = O(u^3), \quad (1)$$

and acting on this space we define the group of formal symplectic (i.e canonical) transformations  $e^{\{\widehat{\mathcal{F}}^{\geq 1}, \cdot\}}$ . Finally we define the space of normal forms as those formal Hamiltonians which Poisson commute with  $D_\omega$ . We prove the following

**Theorem.** *All Hamiltonians  $H$  as above are formally symplectically conjugated to normal form. Moreover the normal form Hamiltonian associated to  $H$  is unique.*

Having properly developed the formal framework, we consider the question of formal vs. analytic linearization in the infinite dimensional setting on the phase space of Gevrey regular functions.

In order to keep technical difficulties to a minimum, we work on Nonlinear Schrödinger like Hamiltonians with the standard symplectic structure on  $\ell_2 = \ell_2(\mathbb{Z}, \mathbb{C})$ . As phase space we consider the sequences of Gevrey regularity, namely we consider the weighted space

$$\mathfrak{h}_{s,p,\theta} := \left\{ u \in \ell^2(\mathbb{Z}, \mathbb{C}) : |u|_s^2 := \sum_{j \in \mathbb{Z}} \langle j \rangle^{2p} e^{2s\langle j \rangle^\theta} |u_j|^2 < \infty \right\} \quad (2)$$

where  $\langle j \rangle := \max(|j|, 1)$ ,  $s > 0$ ,  $p \geq \frac{1}{2}$  and  $0 < \theta < 1$ . Then, given  $r > 0$ , we consider the space of  $M$ -regular Hamiltonians  $P \in \mathcal{H}_r(\mathfrak{h}_{s,p,\theta})$ , such that the Cauchy majorant of the map  $u \rightarrow X_P(u)$  is analytic from the ball  $B_r(\mathfrak{h}_{s,p,\theta})$ , centered at the origin and of radius  $r$  into  $\mathfrak{h}_{s,p,\theta}$ .

Now we consider a Hamiltonian as in (1), with the additional condition that  $P \in \mathcal{H}_{r_0}(\mathfrak{h}_{s_0,p,\theta})$  and the frequency  $\omega$  is ‘‘Diophantine’’ in the following sense introduced by Bourgain [Bou05]. We set

$$\Omega := \left\{ \omega = (\omega_j)_{j \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}, \quad \sup_j |\omega_j - j^2| < 1/2 \right\} \quad (3)$$

**Definition 1.1.** Given  $\gamma > 0$ , we denote by  $\mathcal{D}_\gamma$  the set of Diophantine frequencies

$$\mathcal{D}_\gamma := \left\{ \omega \in \Omega : |\omega \cdot \ell| > \gamma \prod_{n \in \mathbb{Z}} \frac{1}{(1 + |\ell_n|^2 \langle n \rangle^2)}, \quad \forall \ell \in \mathbb{Z}_f^{\mathbb{Z}} \setminus \{0\} \right\}. \quad (4)$$

The map  $(\omega_j)_{j \in \mathbb{Z}} \rightarrow (j^2 - \omega_j)_{j \in \mathbb{Z}}$  identifies  $\Omega$  with  $[-1/2, 1/2]^{\mathbb{Z}}$ . Hence we endow  $\Omega$  with the product topology and with the corresponding probability measure. With respect to such measure Diophantine frequencies are typical,<sup>1</sup> namely  $\Omega \setminus \mathcal{D}_\gamma$  has measure proportionally bounded by  $\gamma$  (see for instance [BMP18] [Lemma 4.1]).

Then we prove:

**Theorem.** *If  $H = D_\omega + P$  with  $P \in \mathcal{H}_{r_0}(\mathfrak{h}_{s_0,p,\theta})$  is formally conjugated to  $D_\omega$ , then there exists  $r_1 < r_0$ ,  $s_1 > s_0$  and a close to identity analytic symplectic change of variables  $\Psi : B_{r_1}(\mathfrak{h}_{s_1,p,\theta}) \rightarrow \mathfrak{h}_{s_1,p,\theta}$  such that  $H \circ \Psi = \sum_{j \in \mathbb{Z}} \omega_j |u_j|^2$ .*

*Remark 1.2.* We have formulated our result for frequencies close to  $j^2$ , however the proof holds, essentially verbatim, in the more general case of Hamiltonian perturbations of  $D_\omega$  when the frequency  $\omega \sim j^\alpha$  with  $\alpha > 1$ . More precisely our main result holds for all Hamiltonians  $H = D_\omega + P$  with  $P \in \mathcal{H}_{r_0}(\mathfrak{h}_{s_0,p,\theta})$  provided that

$$\mathcal{D}_{\gamma,\alpha} := \left\{ \omega = (\omega_j)_{j \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}} : \sup_j |\omega_j - j^\alpha| < 1/2, \quad |\omega \cdot \ell| > \gamma \prod_{n \in \mathbb{Z}} \frac{1}{(1 + |\ell_n|^2 \langle n \rangle^2)}, \quad \forall \ell \in \mathbb{Z}_f^{\mathbb{Z}} \setminus \{0\} \right\}. \quad (5)$$

Note that the measure of  $\mathcal{D}_{\gamma,\alpha}$  may be estimated uniformly in  $\alpha$ , however in the proof of Lemma 3.8 (precisely in Lemma A.10) we need the condition  $\alpha > 1$ , and one gets  $r_1 \rightarrow 0$  as  $\alpha \rightarrow 1^+$ .

<sup>1</sup> If we fix any point  $\omega_0 \in \mathbb{R}^{\mathbb{Z}}$ , then the set of Diophantine frequencies  $\omega$  such that  $|\omega - \omega_0|_\infty < 1/2$  is again typical.

*Remark 1.3.* The requirement  $\theta < 1$  in the definition of the phase space comes from dependence of the constant  $C_1$  in Lemma 3.8. Indeed (see Lemma A.10) one can verify that  $r_1 \rightarrow 0$  (or equivalently  $s_1 \rightarrow \infty$ ) as  $\theta \rightarrow 1^-$ . Similarly the requirement  $s_1 > s_0 > 0$  comes from Lemma 3.8 and again  $r_1 \rightarrow 0$  in the limit  $s_1, s_0 \rightarrow 0^+$ .

As we have explained above our result does not cover the Sobolev case, because if  $s = 0$  we are not able to control the small divisors sufficiently well. If we assume that  $D_\omega$  has no small divisors, that is

$$\inf_{\ell \in \mathbb{Z}^d: |\omega \cdot \ell| \neq 0} |\omega \cdot \ell| \geq \gamma > 0.$$

then we have the following result.

**Theorem.** *Assume that  $s = 0$  (Sobolev case) and that  $D_\omega$  has no small divisors. If  $H$  is formally conjugated to  $D_\omega$ , then there exists  $r_1 < r_0$ , and a close to identity analytic symplectic change of variables  $\Psi : B_{r_1}(\mathfrak{h}_{0,p,\theta}) \rightarrow \mathfrak{h}_{0,p,\theta}$  such that  $H \circ \Psi = \sum_{j \in \mathbb{Z}} \omega_j |u_j|^2$ .*

*1.1.1. Examples* Although we have stated and proved our results in the context of weighted sequence spaces, our statements are written with the PDE context in mind. Indeed it is not difficult to produce (non-local) Pseudo-Differential Equations which satisfy the hypotheses of our main result. The simplest way of course is to start from a linear equation and then perform a non-linear change of variables which puts it in the form  $D_\omega + P$  discussed above. For instance one can start with a linear Schrödinger equation with a convolution potential on the circle

$$i w_t - w_{xx} + V \star w = 0, \quad V \star \sum_{j \in \mathbb{Z}} w_j e^{ijx} = \sum_{j \in \mathbb{Z}} V_j w_j e^{ijx},$$

it is well known that this is a Hamiltonian PDE whose Hamiltonian in the Fourier basis is  $D_\omega$  with  $\omega_j = j^2 + V_j$ . If we take  $V = (V_j)_{j \in \mathbb{Z}} \in \ell_\infty(\mathbb{R})$  as a set of parameters, then for typical  $V$ 's the corresponding  $\omega$  is Diophantine. On the other hand if  $V = 0$  then we are in the case without small divisors.

Now consider the change of variables generated by the time one Hamiltonian flow  $\Psi_K$  of

$$\begin{aligned} K &:= \int_{\mathbb{T}} dx |\langle \partial_x \rangle^{-2} u|^4 \\ &= \sum_{j_1 - j_2 + j_3 - j_4 = 0} \frac{u_{j_1} \bar{u}_{j_2} u_{j_3} \bar{u}_{j_4}}{\langle j_1 \rangle^2 \langle j_2 \rangle^2 \langle j_3 \rangle^2 \langle j_4 \rangle^2}, \quad \text{here } \langle \partial_x \rangle e^{ijx} = \langle j \rangle e^{ijx}. \end{aligned}$$

One can easily verify that (here  $\Pi_0$  is the projection on zero-mean functions)

$$\begin{aligned} D_\omega \circ \Psi_K &= e^{\{K, \cdot\}} \int_{\mathbb{T}} dx |\partial_x u|^2 \\ &= \int_{\mathbb{T}} dx |\partial_x u|^2 + 2 \operatorname{Im} \int_{\mathbb{T}} \bar{u} \Pi_0 (|\langle \partial_x \rangle^{-2} u|^2 \langle \partial_x \rangle^{-2} u) + \text{h.o.t.}, \end{aligned}$$

satisfies all the hypotheses of our theorem.

*1.2. Strategy.* In order to describe our strategy consider 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

$$H = \sum_{j=1}^n \omega_j |u_j|^2 + O(u^3), \quad \text{where } \omega_j \in \mathbb{R} \text{ are the linear frequencies.} \quad (6)$$

Here if the frequencies  $\omega$  are rationally independent, then one can perform the so-called Birkhoff normal form procedure: for  $N \geq 1$  Hamiltonian (6) is transformed into

$$\sum_{j=1}^n \omega_j |u_j|^2 + Z + R, \quad (7)$$

where  $Z$  depends only on the actions  $(|u_i|^2)_{i=1}^n$  while  $R = O(|u|^{N+3})$  has a zero of order at least  $N+3$  in  $|u|$ . At each step, the generating function of the change of variables is a polynomial, so it is analytic and generates a flow in a sufficiently small ball  $B_\delta$  around the origin. It is well known that this procedure generically diverges in  $N$ , but assuming that  $\omega$  is appropriately non resonant, say diophantine<sup>2</sup> one can control  $R$  and hence find  $N = N(\delta)$  which minimizes the size of the remainder  $R$ . It can be shown that it is bounded by an exponentially flat function of  $\delta$ , of order related to  $\tau$  (for a general treatment, see instance, [IoL05,LS10]). This phenomenon is also related to Nekhoroshev kind of result [Pös99,BGG85,N77,Ni04,BCG].

If  $H$  in (6) is “formally linearizable”, namely there exists a formal symplectic change of variables which conjugates  $H$  to  $\sum_{j=1}^n \omega_j |u_j|^2$ , and  $\omega$  is Diophantine, then at each step of the procedure described above, we find  $Z = 0$  and one can prove convergence. In order to apply this general scheme in the infinite dimensional setting we first discuss the BNF procedure at the level of formal power series. Here the fundamental difference w.r.t. the finite dimensional case is that even polynomials can be just formal power series, so it is not a priori obvious that the space of formal power series is well defined and has a Poisson algebra structure (which coincides with the usual one on finite dimensional subspaces). As a simple example consider the formal power series  $H = \sum_j u_j$ , then

$$\{H, \bar{H}\} = \sum_i \sum_j \{u_j, \bar{u}_i\} = \infty.$$

We show that for translation invariant formal Hamiltonians the Poisson brackets are well defined (see also [FGP]), and that formal Hamiltonians are a filtered Lie algebra with respect to a *scaling degree*. Then we define a group of formal symplectic changes of variables, and prove our BNF result. In order to define our changes of variables and prove the group structure we strongly rely on the properties of the *scaling degree* as well as on the Baker Campbell Hausdorff formula.

Then we restrict to functions on the sequence space  $\mathfrak{h}_{s,p,\theta}$ , introduce the space of regular Hamiltonians and state the main relevant properties. All properties were proved in [BMP18] in the more restrictive case of Gauge invariant Hamiltonians, generalizing the results to our case requires some tedious case analysis but follows very closely the approach of [BMP18]; for completeness we give all the proofs in the appendix. Once we

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

have all the basic properties needed to perform Birkhoff Normal Form, proving that formal linearizability implies analytic linearizability becomes a relatively straightforward induction.

## 2. Formal Birkhoff Normal on Sequence Spaces

As usual given a vector  $k \in \mathbb{Z}^{\mathbb{Z}}$ ,  $|k| := \sum_{j \in \mathbb{Z}} |k_j|$ . We denote  $\mathbb{N}_f^{\mathbb{Z}}$  to be the set of finitely supported sequences of non negative integers, similarly for  $\mathbb{Z}_f^{\mathbb{Z}}$ . If  $j \in \mathbb{Z}$  then  $\mathbf{e}_j \in \mathbb{Z}_f^{\mathbb{Z}}$  denotes the vector the  $j$ -coordinate of which is 1, while the others are zero.

**Definition 2.1** (*Formal power series*). We consider the space  $\mathcal{F}$  of formal power series expansions in  $u \in \mathbb{C}^{\mathbb{Z}}$ :

$$H(u) = \sum_{\alpha, \beta \in \mathbb{N}_f^{\mathbb{Z}}} H_{\alpha, \beta} u^\alpha \bar{u}^\beta, \quad u \in \mathbb{C}^{\mathbb{Z}}, \quad u^\alpha := \prod_{j \in \mathbb{Z}} u_j^{\alpha_j} \quad |v| := \sum_i |v_i|$$

with the following properties:

- (1)  $H_{0,0} = 0$ ,  $H_{\mathbf{e}_0,0} = H_{0,\mathbf{e}_0} = 0$
- (2) Reality condition:

$$H_{\alpha, \beta} = \overline{H_{\beta, \alpha}}; \tag{8}$$

- (3) Momentum conservation:

$$H_{\alpha, \beta} = 0 \quad \text{if} \quad \pi(\alpha, \beta) := \sum_{j \in \mathbb{Z}} j(\alpha_j - \beta_j) \neq 0 \tag{9}$$

*Remark 2.2.* The condition (3) means that the formal Hamiltonian is invariant w.r.t. the symmetry  $u_j \rightarrow e^{ij\tau} u_j$ ,  $\tau \in \mathbb{R}$ .

We shall denote

$$\mathcal{M} := \{(\alpha, \beta) \in \mathbb{N}_f^{\mathbb{Z}} : \pi(\alpha, \beta) = 0\}$$

so that  $H \in \mathcal{F}$  can be written as

$$\sum_{(\alpha, \beta) \in \mathcal{M}} H_{\alpha, \beta} u^\alpha \bar{u}^\beta$$

Finally we define

$$\begin{aligned} \mathcal{K} &:= \left\{ Z \in \mathcal{F} : Z(u) = \sum_{\alpha \in \mathbb{N}_f^{\mathbb{Z}}} Z_{\alpha, \alpha} |u|^{2\alpha} \right\}, \\ \mathcal{R} &:= \left\{ R \in \mathcal{F} : R(u) = \sum_{\alpha, \beta \in \mathcal{M}: \alpha \neq \beta} R_{\alpha, \beta} u^\alpha \bar{u}^\beta \right\} \end{aligned} \tag{10}$$

and we can decompose  $\mathcal{F} = \mathcal{K} \oplus \mathcal{R}$  as each element of  $\mathcal{F}$  can uniquely be expressed in term of monomials the coefficients of which is either zero or not zero.

**Definition 2.3** (*Scaling degree*). For  $d \in \mathbb{N}$ , we denote by  $\mathcal{F}^d \subset \mathcal{F}$  the vector space of homogeneous formal polynomials of degree  $d + 2$ , and define

$$\mathcal{F}^{\leq d} = \bigoplus_{h \leq d} \mathcal{F}^h, \quad \mathcal{F}^{> d} := \widehat{\bigoplus}_{h > d} \mathcal{F}^h, \quad \mathcal{F}^{\geq d} := \mathcal{F}^{> d} \oplus \mathcal{F}^d, \quad \mathcal{F} = \mathcal{F}^{\leq d} \oplus \mathcal{F}^{> d}, \dots$$

We define the projections associated to these direct sum decompositions

$$\Pi^{(\leq d)} H = \sum_{|\alpha|+|\beta|=d+2} H_{\alpha,\beta} u^\alpha \bar{u}^\beta, \quad \Pi^{(> d)} H = \sum_{|\alpha|+|\beta|>d+2} H_{\alpha,\beta} u^\alpha \bar{u}^\beta, \dots$$

Elements of  $\mathcal{F}^{\geq d}$  (resp.  $\mathcal{F}^{> d}$ ) are said to be of *scaling order*  $\geq d + 2$  (resp.  $> d + 2$ ). In the sequel, for simplicity, we shall just say that an element of  $f \in \mathcal{F}^{\geq d}$  is of “order  $d$ ” and we shall say that  $f$  is “exactly of order  $d$ ” if it has a non vanishing component  $\Pi^{(d)} f$  in  $\mathcal{F}^d$ . Finally we define

$$\Pi^{\mathcal{K}} H = \sum_{\alpha} H_{\alpha,\alpha} |u|^{2\alpha}, \quad \Pi^{\mathcal{R}} H = \sum_{\alpha \neq \beta} H_{\alpha,\beta} u^\alpha \bar{u}^\beta.$$

We denote by  $\mathcal{K}^d := \mathcal{F}^d \cap \mathcal{K}$  and similarly for  $\mathcal{R}$  and  $\geq d, \leq d$ . Note that  $\mathcal{F} = \widehat{\bigoplus}_d \mathcal{F}^d$ .

*Remark 2.4.* Of course, since we are in infinite dimension, even if the  $\mathcal{F}^d$  are homogeneous they are only formal polynomials. However if we restrict to monomials  $u^\alpha \bar{u}^\beta$  with  $|\alpha_j| + |\beta_j| = 0$  for all  $j > N$  we are working on the usual space of polynomials on which we have the standard symplectic structure  $i \sum_{j \leq N} du_j \wedge d\bar{u}_j$ . We now show that such structure extends to  $\mathcal{F}$ .

**Proposition 2.5.** *The following Formula (11) is well defined and endows  $\mathcal{F}$  with a Poisson algebra structure which is a filtered Lie algebra w.r.t. the  $\mathcal{F}^{\geq d}$ 's.*

$$\{F, G\} := i \sum_{(\alpha^{(i)}, \beta^{(i)}) \in \mathcal{M}} F_{\alpha^{(1)}, \beta^{(1)}} G_{\alpha^{(2)}, \beta^{(2)}} \sum_j \left( \alpha_j^{(1)} \beta_j^{(2)} - \beta_j^{(1)} \alpha_j^{(2)} \right) u^{\alpha^{(1)} + \alpha^{(2)}} \bar{u}^{\beta^{(1)} + \beta^{(2)}} e_j \quad (11)$$

Before proving our assertion we need a technical lemma. Let  $\mathbf{e}_j \in \mathbb{N}_f^{\mathbb{Z}}$  be the  $j$ th vector of the standard basis.

**Lemma 2.6.** (1) *Given  $\alpha \in \mathbb{N}_f^{\mathbb{Z}}$  there is only a finite number of pairs  $\alpha^{(1)}, \alpha^{(2)} \in \mathbb{N}_f^{\mathbb{Z}}$  with  $\alpha = \alpha^{(1)} + \alpha^{(2)}$ . (2) *Given  $(\alpha, \beta) \in \mathcal{M}$  there is only a finite number of pairs  $(\alpha^{(1)}, \beta^{(1)}), (\alpha^{(2)}, \beta^{(2)}) \in \mathcal{M}$  and indices  $j \in \mathbb{Z}$  such that:**

- (i)  $(\alpha, \beta) = (\alpha^{(1)}, \beta^{(1)}) + (\alpha^{(2)}, \beta^{(2)}) - (\mathbf{e}_j, \mathbf{e}_j)$
- (ii) *one has  $\alpha_j^{(1)} \beta_j^{(2)} + \alpha_j^{(2)} \beta_j^{(1)} \neq 0$ .*

*Proof.* (1) is clear since for all  $j$  one has  $0 \leq (\alpha_1)_j \leq \alpha_j$ .

(2) By item (1) we may divide  $(\alpha, \beta) = (a^{(1)}, b^{(1)}) + (a^{(2)}, b^{(2)})$  in a finite number of ways. Then the pairs  $(\alpha^{(1)}, \beta^{(1)}), (\alpha^{(2)}, \beta^{(2)})$  can only have one of the following forms (up to exchanging the indices)

- (A)  $(\alpha^{(1)}, \beta^{(1)}) = (a^{(1)}, b^{(1)}) + (\mathbf{e}_j, \mathbf{e}_j), \quad (\alpha^{(2)}, \beta^{(2)}) = (a^{(2)}, b^{(2)})$
- (B)  $(\alpha^{(1)}, \beta^{(1)}) = (a^{(1)}, b^{(1)}) + (\mathbf{e}_j, 0), \quad (\alpha^{(2)}, \beta^{(2)}) = (a^{(2)}, b^{(2)}) + (0, \mathbf{e}_j),$



for some index  $j \in \mathbb{Z}$ .

If we are in case (A) then by condition (ii) we have  $j \in \text{Supp}(a^{(2)}+b^{(2)})$ , which restricts to a finite number of possible  $j$ 's. Otherwise in case (B) by momentum conservation we have  $j = -\pi(a^{(1)}, b^{(1)}) = \pi(a^{(2)}, b^{(2)})$  and again  $j$  is restricted to a finite number of possible choices.  $\square$

*Proof of Proposition 2.5.* The fact that the Poisson bracket is well defined follows immediately from the previous Lemma. Indeed by construction

$$\{F, G\} = \sum_{\alpha, \beta} P_{\alpha, \beta} u^\alpha \bar{u}^\beta \in \mathcal{F}$$

where  $P_{\alpha, \beta} = 0$  if  $\pi(\alpha, \beta) \neq 0$  and otherwise

$$P_{\alpha, \beta} = i \sum_j \sum_{\substack{\alpha^{(i)}, \beta^{(i)} \in \mathbb{N}_j^{\mathbb{Z}}: \pi(\alpha^{(i)}, \beta^{(i)})=0 \\ \alpha = \alpha^{(1)} + \alpha^{(2)} - \mathbf{e}_j, \beta = \beta^{(1)} + \beta^{(2)} - \mathbf{e}_j}} F_{\alpha^{(1)}, \beta^{(1)}} G_{\alpha^{(2)}, \beta^{(2)}} \left( \alpha_j^{(1)} \beta_j^{(2)} - \beta_j^{(1)} \alpha_j^{(2)} \right). \quad (12)$$

Then item 2 of the previous Lemma implies that  $P_{\alpha, \beta}$  above is given by a finite sum.

The fact that it endows  $\mathcal{F}$  with a Poisson algebra structure follows from the fact that the infinitely many identities defining such a structure involve only a finite number of elements  $u_i, \bar{u}_i$  and then we are in the canonical Poisson algebra.

The filtered Lie algebra property comes from the fact that in (12) we get  $|\alpha| + |\beta| = |\alpha^{(1)}| + |\alpha^{(2)}| + |\beta^{(1)}| + |\beta^{(2)}| - 2$ , this shows that if  $F \in \mathcal{F}^{\geq d_1}$ , and  $G \in \mathcal{F}^{\geq d_2}$  then

$$|\alpha| + |\beta| \geq d_1 + 2 + d_2 + 2 - 2 = d_1 + d_2 + 2. \quad (13)$$

so  $\{F, G\} \in \mathcal{F}^{\geq d_1 + d_2}$ .  $\square$

*Remark 2.7.* Let  $H_i \in \mathcal{F}^{\geq d_i}$  be a sequence of formal Hamiltonians with  $d_{i+1} \geq d_i$  for all  $i \geq 1$ . Then the series

$$H = \sum_{i=1}^{\infty} H_i \in \mathcal{F}^{\geq d_1}$$

is well defined since for any  $d \geq d_0$  the projection

$$\Pi^{(d)} H = \Pi^{(d)} \sum_{i: d_i \leq d} H_i$$

is a finite sum.

We say that a linear operator  $L : \mathcal{F} \rightarrow \mathcal{F}$  is of order (or increase the order by)  $d$  if for all  $h$

$$L : \mathcal{F}^{\geq h} \rightarrow \mathcal{F}^{\geq h+d}.$$

**Lemma 2.8.** *Let  $L_n$  be a sequence of linear operators on  $\mathcal{F}$  and let  $\mathfrak{d}_n$  be the order of  $L_n$ . If the sequence  $\mathfrak{d}_n$  increases to infinity then*

$$L := \sum_{n=1}^{\infty} L_n, \quad T = \prod_{n=1}^{\infty} (\text{id} + L_n) - \text{id}$$

are linear operators on  $\mathcal{F}$  of order  $\mathfrak{d}_1$ .

*Proof.* For the first statement, for all  $\mathfrak{d} \in \mathbb{N}$  let  $N(\mathfrak{d})$  be the largest  $N$  such that  $\mathfrak{d}_N \leq \mathfrak{d}$ . By construction  $\Pi^{(\leq \mathfrak{d})} L_n K = 0$  for all  $n > N(\mathfrak{d})$  and for any  $K \in \mathcal{F}$ . Then for all  $K \in \mathcal{F}$  and  $N > N(\mathfrak{d})$  one has

$$\Pi^{(\leq \mathfrak{d})} \sum_{n=1}^N L_n K = \Pi^{(\leq \mathfrak{d})} \sum_{n=1}^{N(\mathfrak{d})} L_n K,$$

and the claim follows.

Regarding the second statement we proceed similarly

$$\prod_{n=1}^N (\text{id} + L_n) = \prod_{n=1}^{N-1} (\text{id} + L_n) + L_N \prod_{n=1}^{N-1} (\text{id} + L_n),$$

hence, for all  $\mathfrak{d} \geq 0$  and all  $N > N(\mathfrak{d})$

$$\Pi^{(\leq \mathfrak{d})} \prod_{n=1}^N (\text{id} + L_n) = \Pi^{(\leq \mathfrak{d})} \prod_{n=1}^{N(\mathfrak{d})} (\text{id} + L_n).$$

□

As a direct consequence we have the following.

**Corollary 2.9.** *Given  $G \in \mathcal{F}^{\geq \mathfrak{d}}$ , with  $\mathfrak{d} \geq 1$  we define*

$$\text{ad}_G := \{G, \cdot\}, \quad \Phi_G := \exp(\{G, \cdot\}) = \sum_{k \geq 0} \frac{\text{ad}_G^k}{k!}, \quad (14)$$

then  $\text{ad}_G$  and  $\Phi_G - \text{id}$  are operators of order  $\mathfrak{d}$ , namely

$$\text{ad}_G, \Phi_G - \text{id} : \mathcal{F}^{\geq h} \rightarrow \mathcal{F}^{\geq h + \mathfrak{d}}.$$

Similarly for any sequence  $b_k$  one has that

$$\sum_{k \geq n} b_k \text{ad}_G^k : \mathcal{F}^{\geq h} \rightarrow \mathcal{F}^{\geq h + \mathfrak{d}n}.$$

**Definition 2.10.** Given  $G \in \mathcal{F}^{\geq 1}$  we call the operator  $\Phi_G$  defined in (14) a formal symplectic change of variables on  $\mathcal{F}$ .

The following Lemma ensures the group structure of the formal symplectic changes of variables

**Lemma 2.11** (Baker–Campbell–Hausdorff). *Given  $F \in \mathcal{F}^{\geq d_1}$  and  $G \in \mathcal{F}^{\geq d_2}$ , with  $d_i \geq 1$ , then there exists  $K \in \mathcal{F}^{\geq 1}$ , such that*

$$e^{\{G,\cdot\}} e^{\{F,\cdot\}} = e^{\{K,\cdot\}}, \quad K - F - G \in \mathcal{F}^{\geq d_1 + d_2}$$

*Proof.* By the Baker–Campbell–Hausdorff formula ([Se92] [p. 29]) one has

$$K := \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{r_i + s_i > 0} \frac{[G^{r_1} F^{s_1} \dots G^{r_n} F^{s_n}]}{(\sum_{i=1}^n (r_i + s_i)) \prod_{i=1}^n r_i! s_i!}$$

where

$$[G^{r_1} F^{s_1} \dots G^{r_n} F^{s_n}] := \begin{cases} \text{ad}_G^{r_1} \text{ad}_F^{s_1} \dots \text{ad}_G^{r_n} F & \text{if } s_n = 1 \\ \text{ad}_G^{r_1} \text{ad}_F^{s_1} \dots \text{ad}_F^{s_{n-1}} G & \text{if } s_n = 0, \text{ and } r_n = 1 \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

Recalling that  $F \in \mathcal{F}^{\geq d_1}$  and  $G \in \mathcal{F}^{\geq d_2}$ , each term  $\text{ad}_G^{r_1} \text{ad}_F^{s_1} \dots \text{ad}_G^{r_n} F$  (resp.  $\text{ad}_G^{r_1} \text{ad}_F^{s_1} \dots \text{ad}_F^{s_{n-1}} G$ )

is of order  $(\sum_{i=1}^n r_i) d_2 + (\sum_{i=1}^n s_i) d_1 \geq n \min(d_1, d_2)$ .

Hence setting  $N(d)$  to be the largest  $N$  such that  $N \min(d_1, d_2) \leq dn$

$$\Pi^{\leq d} K = \Pi^{\leq d} \sum_{n=1}^{N(d)} \frac{(-1)^{n-1}}{n} \sum_{r_i + s_i > 0} \frac{[G^{r_1} F^{s_1} \dots G^{r_n} F^{s_n}]}{(\sum_{i=1}^n (r_i + s_i)) \prod_{i=1}^n r_i! s_i!}$$

Moreover if  $n \geq 2$  then the Hamiltonian in (15) is of order  $\geq d_1 + d_2$ , so  $K - F - G \in \mathcal{F}^{\geq d_1 + d_2}$ .  $\square$

**Lemma 2.12.** *Given a sequence of generating functions  $G_i \in \mathcal{F}^{\geq d_i}$  with  $d_{i+1} > d_i \geq 1$  then there exists  $\mathcal{G} \in \mathcal{F}^{\geq d_1}$  such that the composition*

$$\prod_i e^{\{G_i,\cdot\}} = e^{\{\mathcal{G},\cdot\}}$$

*Proof.* By Lemma 2.8 with  $L_n = e^{\{G_n,\cdot\}} - \text{id}$  we know that  $\prod_i e^{\{G_i,\cdot\}}$  is a well defined operator of  $\mathcal{F}$ . Using Lemma 2.11 we can define  $F_k \in \mathcal{F}^{\geq 1}$  iteratively so that

$$e^{\{F_k,\cdot\}} = e^{\{G_k,\cdot\}} e^{\{F_{k-1},\cdot\}}$$

since  $e^{\{G_k,\cdot\}} - \text{id}$  is of order  $d_k$  there exists  $N(d)$  such that if  $k > N(d)$  then

$$\Pi^{(\leq d)} F_k = \Pi^{(\leq d)} F_{N(d)}$$

Then  $\mathcal{G} = \lim_{k \rightarrow \infty} F_k$  is well defined.  $\square$

For any vector  $\omega \in \mathbb{R}^{\mathbb{Z}}$  such that

$$\omega \cdot \ell \neq 0, \quad \forall \ell \in \mathbb{Z}_f^{\mathbb{Z}} \setminus \{0\},$$

we define the *non-resonant* quadratic Hamiltonian

$$D_\omega := \sum_j \omega_j |u_j|^2.$$

**Lemma 2.13.** *The operator  $\text{ad}_{D_\omega}$  is invertible on  $\mathcal{R}^{(\mathfrak{d})}$  for all  $\mathfrak{d}$ .*

*Proof.* Given  $F \in \mathcal{R}^{(\mathfrak{d})}$ , we have

$$\{D_\omega, G\} = i \sum_{\substack{\alpha^{(2)}, \beta^{(2)} \in \mathbb{N}^{\mathbb{Z}}, \\ |\alpha^{(2)}| + |\beta^{(2)}| < \infty, \pi(\alpha^{(2)}, \beta^{(2)}) = 0}} G_{\alpha^{(2)}, \beta^{(2)}} \left( \sum_j \omega_j (\beta_j^{(2)} - \alpha_j^{(2)}) \right) u^{\alpha^{(2)}} \bar{u}^{\beta^{(2)}} = F$$

Hence, we have  $G := \text{ad}_{D_\omega}^{-1}(F)$  with for all  $\alpha^{(2)}, \beta^{(2)} \in \mathbb{N}^{\mathbb{Z}}, \alpha^{(2)} \neq \beta^{(2)}$  with  $|\alpha^{(2)}| + |\beta^{(2)}| < \infty$  and  $\pi(\alpha^{(2)}, \beta^{(2)}) = 0$ ,

$$G_{\alpha^{(2)}, \beta^{(2)}} := F_{\alpha^{(2)}, \beta^{(2)}} \left( \sum_j i \omega_j (\beta_j^{(2)} - \alpha_j^{(2)}) \right)^{-1}, \quad G_{\alpha^{(2)}, \alpha^{(2)}} = 0.$$

□

**Proposition 2.14.** (Birkhoff Normal Form) *Given any formal Hamiltonian of the form*

$$H = D_\omega + Z + R \tag{16}$$

where  $Z \in \mathcal{K}^{\geq 2}$  and  $R \in \mathcal{F}^{\geq \mathfrak{d}}$  with  $\mathfrak{d} \geq 1$ , then

(1) **Formal Normal Form:** *there exists  $S \in \mathcal{F}^{\geq \mathfrak{d}}$  such that*

$$e^{\{S, \cdot\}} H = D_\omega + \tilde{Z}, \quad \tilde{Z} - Z \in \mathcal{K}^{\geq \mathfrak{d}}.$$

(2) **Uniqueness:** *if  $G \in \mathcal{F}^{\geq 1}$  is such that  $e^{\{G, \cdot\}} H \in \mathcal{K}$  then  $e^{\{G, \cdot\}} H = e^{\{S, \cdot\}} H$ . Hence to each  $H$  as above we can associate a unique  $Z_H \in \mathcal{K}^{\geq 2}$  such that  $e^{\{S, \cdot\}} H = D_\omega + Z_H$ .*

*Proof.* For item (1) Let us first consider the case  $\mathfrak{d} \geq 2$ . we start with a Hamiltonian  $H_0 \in \mathcal{F}$  of the form  $D_\omega + Z_0 + P_0$  with  $P_0 \in \mathcal{F}^{\geq \mathfrak{d}}$  and we iteratively construct a sequence of generating functions  $S_i \in \mathcal{R}^{\geq 2i+\mathfrak{d}}$  and Hamiltonians  $H_i$  by setting

$$\{D_\omega, S_i\} = \Pi^{\mathcal{R}} H_i, \quad H_{i+1} = e^{\{S_i, \cdot\}} H_i.$$

We now show inductively that for each  $i$

$$\Pi^{(<2i+\mathfrak{d})} \Pi^{\mathcal{R}} H_i = 0, \quad S_i \in \mathcal{R}^{\geq 2i+\mathfrak{d}}$$

so in other words

$$H_i = D_\omega + Z_i + P_i, \quad Z_i \in \mathcal{K} \cap \mathcal{F}^{\leq 2i+\mathfrak{d}-1}, \quad P_i \in \mathcal{F}^{\geq 2i+\mathfrak{d}}.$$

For  $i = 0$  we just set  $Z_0 = \Pi^{<\mathfrak{d}} Z$  and  $P_0 = R + \Pi^{\geq \mathfrak{d}} Z$ . By induction we assume that  $P_i \in \mathcal{F}^{\geq 2i+\mathfrak{d}}$ . Then by Lemma 2.13,  $S_i \in \mathcal{F}^{\geq 2i+\mathfrak{d}}$ .

$$\begin{aligned} e^{\{S_i, \cdot\}} H_i &= D_\omega + Z_i + P_i + \{S_i, D_\omega\} + \sum_{h=2}^{\infty} \frac{\text{ad}_{S_i}^{h-1}}{h!} \{S_i, D_\omega\} + \sum_{k=1}^{\infty} \frac{\text{ad}_{S_i}^k}{k!} (Z_i + P_i) \\ &= D_\omega + Z_i + \Pi^{\mathcal{K}} P_i - \sum_{k=1}^{\infty} \frac{\text{ad}_{S_i}^k}{(k+1)!} \Pi^{\mathcal{R}} P_i + \sum_{k=1}^{\infty} \frac{\text{ad}_{S_i}^k}{k!} (Z_i + P_i). \end{aligned}$$

So we may set

$$Z_{i+1} := Z_i + \Pi^{(<2i+d+2)} \Pi^{\mathcal{K}} P_i, \quad P_{i+1} = e^{\{S_i, \cdot\}} H_i - D_\omega - Z_{i+1}$$

and verify that  $P_{i+1} \in \mathcal{F}^{\geq 2i+d+2}$  by applying Proposition 2.5 and noticing that, since  $4i + 2\mathfrak{d} \geq 2i + \mathfrak{d} + 2$ , the term of lowest degree is  $\{S_i, Z\}$ .

Then we set

$$\tilde{Z} = \lim_{i \rightarrow \infty} Z_i = Z_0 + \sum_{i=0}^{\infty} \Pi^{\leq 2i+d+1} \Pi^{\mathcal{K}} P_i = Z_0 + \Pi^{\mathcal{K}} \sum_{i=0}^{\infty} \Pi^{(<2i+d+2)} \Pi^{(\geq 2i+d)} P_i,$$

which is well defined by Remark 2.7. Finally by Lemma 2.12 we can define  $S \in \mathcal{F}^{\geq \mathfrak{d}}$  so that

$$e^{\{S, \cdot\}} = \prod_{i=0}^{\infty} e^{\{S_i, \cdot\}}.$$

If  $\mathfrak{d} = 1$  we perform a preliminary step in order to increase the degree by one and then we start the procedure explained above. We start with  $H = D_\omega + P$ , with  $P := R + Z$ . As before we fix  $S \in \mathcal{R}^{\geq 1}$  so that  $\{D_\omega, S\} = \Pi^{\mathcal{R}} H$  we set

$$H_0 := e^{\{S, \cdot\}} H = D_\omega + \Pi^{\mathcal{K}} P - \sum_{k=1}^{\infty} \frac{\text{ad}_S^k}{(k+1)!} \Pi^{\mathcal{R}} P + \sum_{k=1}^{\infty} \frac{\text{ad}_S^k}{k!} P.$$

then fixing  $Z_0 := \Pi^{\leq 2} \Pi^{\mathcal{K}} P$  and  $P_0 := H_0 - D_\omega - Z_0$  we are in the setting of the previous case.

Regarding item (2) we remark that If  $e^{\{S_1, \cdot\}}$  transforms a normal form  $D_\omega + K_1$  into a normal form  $D_\omega + K_2$ , then

$$e^{\{S_1, \cdot\}}(D_\omega + K_1) = D_\omega + K_1 + \sum_{h=1}^{\infty} \frac{\text{ad}_{S_1}^{h-1}}{h!} \{S_1, D_\omega + K_1\} = D_\omega + K_2.$$

Since  $\mathcal{K} = \mathcal{K}^{\geq 2}$  and  $S \in \mathcal{H}^{\geq 1}$ , comparing homogeneous terms of degree 1 we get  $\{S_1, D_\omega\} = 0$  so we should have  $S_1^{(1)} \in \mathcal{K}$  which can only be possible if  $S_1^{(1)} = 0$ . Comparing homogeneous terms of degree 2, we obtain  $K_1^{(2)} - K_2^{(2)} + \{S_1^{(2)}, D_\omega\} = 0$ . Recalling that  $\{S_1^{(2)}, D_\omega\} \in \mathcal{R}$  we have  $K_1^{(2)} - K_2^{(2)} \in \mathcal{K} \cap \mathcal{R}$  is zero and  $S_1^{(2)} \in \mathcal{K}$ . Assuming that  $K_1^{(j)} = K_2^{(j)} \in \mathcal{K}$  and  $S_1^{(j)} \in \mathcal{K}$  for  $2 \leq j \leq m$ . Then we have

$$\begin{aligned} & K_1^{(m+1)} - K_2^{(m+1)} + \{S_1^{(m+1)}, D_\omega\} + \sum_{h=2}^{\infty} \frac{1}{h!} \left( \sum_{j_1 + \dots + j_h = m+1} \{S_1^{(j_1)}, \{S_1^{(j_2)}, \dots, \{S_1^{(j_h)}, D_\omega\}\}\} \right. \\ & \left. + \sum_{j_1 + \dots + j_h + j_{h+1} = m+1} \{S_1^{(j_1)}, \{S_1^{(j_2)}, \dots, \{S_1^{(j_h)}, K_1^{(j_{h+1})}\}\}\} \right) = 0 \end{aligned} \quad (17)$$

By induction and since  $D_\omega$  is non resonant, then both sums above are zero. Hence, we the same reasoning as above, we obtain  $K_1^{(m+1)} = K_2^{(m+1)} \in \mathcal{K}$  and  $S_1^{(m+1)} \in \mathcal{K}$ . The result follows from Proposition 2.12.  $\square$

**Corollary 2.15.** For any  $H$  as in (16), if for  $G \in \mathcal{F}^{\geq 1}$  one has  $e^{\{G, \cdot\}} H = D_\omega + Z + R$  with  $R \in \mathcal{F}^{\geq d_1}$  then  $Z - Z_H \in \mathcal{K}^{\geq d_1}$ .

*Proof.* By Proposition 2.15 (1) there exists  $S \in \mathcal{F}^{\geq d_1}$  which normalizes  $D_\omega + Z + R$  to  $D_\omega + \tilde{Z}$  with  $\tilde{Z} - Z \in \mathcal{F}^{\geq d_1}$ . By Lemma 2.12 there exists  $G_1 \in \mathcal{F}$  such that  $e^{\{G_1, \cdot\}} = e^{\{S, \cdot\}} e^{\{G, \cdot\}}$ . Since  $G_1$  puts  $H$  in normal form, by Proposition 2.15 (2),  $\tilde{Z} = Z_H$  and the result follows.  $\square$

**Definition 2.16.** We say that  $H$  is formally linearizable if  $Z_H = 0$ .

**Corollary 2.17.** If  $H$  is formally linearizable and there exists a formal symplectic change of variables with  $e^{\{S, \cdot\}} H = D_\omega + Z + R$  with  $R \in \mathcal{F}^{\geq d}$  and  $Z \in \mathcal{K}^{< d}$  (this last condition does not imply any loss of generality) then  $Z = 0$ .

*Proof.* This follows directly from Corollary 2.15.  $\square$

**2.1. Resonant Hamiltonians.** In this section we do not assume  $D_\omega$  to be non-resonant. Instead, we consider

$$\mathcal{K} := \left\{ Z \in \mathcal{F} : Z(u) = \sum_{\alpha, \beta \in \mathcal{M}: \omega \cdot (\alpha - \beta) = 0} Z_{\alpha, \beta} u^\alpha \bar{u}^\beta \right\},$$

$$\mathcal{R} := \left\{ R \in \mathcal{F} : R(u) = \sum_{\alpha, \beta \in \mathcal{M}: \omega \cdot (\alpha - \beta) \neq 0} R_{\alpha, \beta} u^\alpha \bar{u}^\beta \right\} \quad (18)$$

and we can decompose  $\mathcal{F} = \mathcal{K} \oplus \mathcal{R}$  as each element of  $\mathcal{F}$  can uniquely be expressed in term of monomials the coefficients of which is either zero or not zero.

**Proposition 2.18.** Given any formal Hamiltonian of the form  $H = D_\omega + Z + R$ , where  $Z \in \mathcal{K}^{\geq 1}$  and  $R \in \mathcal{F}^{\geq \bar{d}}$  with  $\bar{d} \geq 1$ . Here  $D_\omega$  is **not assumed to be non resonant**.

(1) **Formal Normal Form:** there exists  $S \in \mathcal{F}^{\geq \bar{d}}$  such that

$$e^{\{S, \cdot\}} H = D_\omega + \tilde{Z}, \quad \tilde{Z} - Z \in \mathcal{K}^{\geq \bar{d}}.$$

(2) **Transformation from normal form to normal form :** if  $H \in \mathcal{K}$  and if  $G \in \mathcal{F}^{\geq 1}$  is such that  $e^{\{G, \cdot\}} H \in \mathcal{K}$  then  $G \in \mathcal{K}$

(3) **All normal forms of a formally linearizable Hamiltonian are the same :** If  $H$  is formally linearizable, then all its normal form are equal to  $D_\omega$ .

*Proof.* The proof of the second point follows verbatim the proof of the second point of Proposition 2.15. Indeed, let us assume that

$$e^{\{S_1, \cdot\}} (D_\omega + K_1) = D_\omega + K_1 + \sum_{h=1}^{\infty} \frac{\text{ad}_{S_1}^{h-1}}{h!} \{S_1, D_\omega + K_1\} = D_\omega + K_2.$$

Therefore, we have  $S_1^{(1)} \in \mathcal{K}$ , and  $K_1^{(2)} - K_2^{(2)} + \{S_1^{(2)}, D_\omega\} + \{S_1^{(1)}, K_1^{(1)}\} = 0$ . Let  $K_1, K_2 \in \mathcal{K}^2$ . By Jacobi identity, we have  $\{D_\omega, \{K_1, K_2\}\} = -\{K_1, \{K_2, D_\omega\}\} - \{K_2, \{D_\omega, K_1\}\} = 0$ . Hence  $\{K_1, K_2\} \in \mathcal{K}$ . Hence,  $\{S_1^{(1)}, K_1^{(1)}\} \in \mathcal{K}$  so that  $\{S_1^{(2)}, D_\omega\} = 0$ , that is  $S_1^{(2)} \in \mathcal{K}$ . Assuming that  $S_1^{(j)} \in \mathcal{K}$  for  $j = 1, \dots, m$ , collecting terms of

degree  $m + 1$  leads to (17). By assumption, its first sum is zero. We show by induction, using Jacobi identity, that  $\{S_1^{(j_1)}, \{\dots\{S_1^{(j_h)}, K_1^{(j_{h+1})}\}\}\} \in \mathcal{K}$ . From (17), we obtain  $\{S_1^{(m+1)}, D_\omega\} = 0$  since it also belongs to  $\mathcal{K}$ . Hence,  $S_1^{(m+1)} \in \mathcal{K}$ .

As to the last point, if  $K_1 = 0$  and  $S_1 \in \mathcal{K}$ , then  $K_2^{(1)} = 0$  and the two sums in (17) are zero by induction. Therefore, from  $K_1^{(m+1)} - K_2^{(m+1)} + \{S_1^{(m+1)}, D_\omega\} = 0$  we obtain  $K_2^{(m+1)} = 0$  for all  $m \geq 1$ .  $\square$

If we know a priori that  $H$  is formally linearizable then we get a faster growth of the degree of  $P_i$ .

**Lemma 2.19.** . *If  $H_0 \in \mathcal{F}$  of the form  $D_\omega + P_0$  with  $P_0 \in \mathcal{F}^{\geq 1}$  is formally linearizable then the sequence of generating functions*

$$\{D_\omega, S_i\} = \Pi^{\mathcal{R}} H_i, \quad H_{i+1} = e^{\{S_i, \cdot\}} H_i.$$

satisfies

$$H_i = D_\omega + P_i, \quad P_i \in \mathcal{F}^{\geq 2^i}.$$

*Proof.* By induction we assume that  $P_i \in \mathcal{F}^{\geq 2^i}$ . Then by construction  $S_i \in \mathcal{F}^{\geq 2^i}$ .

$$\begin{aligned} e^{\{S_i, \cdot\}} H_i &= D_\omega + P_i + \{S_i, D_\omega\} + \sum_{h=2}^{\infty} \frac{\text{ad}_{S_i}^{h-1}}{h!} \{S_i, D_\omega\} + \sum_{k=1}^{\infty} \frac{\text{ad}_S^k}{k!} P_i \\ &= D_\omega + \Pi^{\mathcal{K}} P_i - \sum_{k=1}^{\infty} \frac{\text{ad}_{S_i}^k}{(k+1)!} \Pi^{\mathcal{R}} P_i + \sum_{k=1}^{\infty} \frac{\text{ad}_{S_i}^k}{k!} P_i \\ &=: D_\omega + \Pi^{<2^{i+1}} \Pi^{\mathcal{K}} P_i + P_{i+1}. \end{aligned}$$

By Proposition 2.5 the two series in the formula above are in  $\mathcal{F}^{2^{i+1}}$  so to prove our claim we only need to show  $\Pi^{\mathcal{K}} \Pi^{<2^{i+1}} P_i = 0$ . This is a consequence of Corollary 2.17.  $\square$

### 3. Regular Hamiltonians

We now revisit the formal Birkhoff normal form in the case of analytic Hamiltonians. We start by introducing an appropriate functional setting.

*3.1. Spaces of Hamiltonians.* Let us consider the weighted space

$$\mathfrak{h}_s = \mathfrak{h}_{s,p,\theta} := \left\{ u \in \ell^2(\mathbb{Z}, \mathbb{C}) : |u|_s^2 := \sum_{j \in \mathbb{Z}} \langle j \rangle^{2p} e^{2s \langle j \rangle^\theta} |u_j|^2 < \infty \right\}$$

where  $\langle j \rangle := \max(|j|, 1)$ ,  $p \geq \frac{1}{2}$  and  $0 < \theta \leq 1$ . The spaces  $\mathfrak{h}_{s,p,\theta}$  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<sup>3</sup> 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  $\Omega = \sum_j dy_j \wedge dx_j$ .

Given  $H \in \mathcal{F}$ , we define its majorant as

$$\underline{H}(u) = \sum_{\substack{\alpha, \beta \in \mathbb{N}^{\mathbb{Z}}, \\ |\alpha| + |\beta| < \infty}} |H_{\alpha, \beta}| u^\alpha \bar{u}^\beta. \quad (19)$$

**Definition 3.1** (*M-regular Hamiltonians.*) For  $r > 0$ , let  $\mathcal{H}_{r,s}$  be the subspace of  $\mathcal{F}$  of formal power series  $H$  such that  $\underline{H}$  is point-wise absolutely convergent on  $B_r(\mathfrak{h}_s)$ , the ball of radius  $r$  centered at the origin of  $\mathfrak{h}_s$ , and

$$|H|_{B_r(\mathfrak{h}_s)} \equiv \|H\|_{r,s} := r^{-1} \left( \sup_{|u|_{\mathfrak{h}_s} \leq r} |X_{\underline{H}}|_{\mathfrak{h}_s} \right) < \infty.$$

Note that in  $\mathcal{F}$  one has  $H(0) = 0$  so this is actually a norm.

We shall show in the next subsection that  $H \in \mathcal{H}_{r,s}$  guarantees that the Hamiltonian flow of  $H$  exists at least locally and generates a symplectic transformation on  $\mathfrak{h}_s$ , i.e.  $\mathfrak{h}_s$  is an invariant subspace for the dynamics.

**Theorem 3.2** (Main). *Consider a Hamiltonian of the form*

$$\sum_{j \in \mathbb{Z}} \omega_j |u_j|^2 + P_0, \quad P_0 \in \mathcal{H}_{\mathfrak{r}, s_0} \cap \mathcal{F}^{\geq 1}$$

where  $\omega \in \mathbb{D}_\gamma$ . Assume that there exists  $G \in \mathcal{F}^{\geq 1}$  such that

$$e^{\{G, \cdot\}} H = \sum_{j \in \mathbb{Z}} \omega_j |u_j|^2,$$

then there exists  $r_1 < \mathfrak{r}$ ,  $s_1 > s_0$  and a close to identity change of variables  $\Psi$

$$\Psi : B_{r_1}(\mathfrak{h}_{s_1}) \rightarrow \mathfrak{h}_{s_1}$$

such that  $H \circ \Psi = \sum_{j \in \mathbb{Z}} \omega_j |u_j|^2$ .

Let us state a result which is valid in Sobolev spaces.

<sup>3</sup> 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

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



**Theorem 3.3.** Consider a Hamiltonian of the form

$$\sum_{j \in \mathbb{Z}} \omega_j |u_j|^2 + P_0, \quad P_0 \in \mathcal{H}_{\mathfrak{r},0} \cap \mathcal{F}^{\geq 1}$$

where  $D_\omega$  might be resonant but without small divisors, namely

$$\inf_{\ell \in \mathbb{Z}_j^{\neq}: |\omega \cdot \ell| \neq 0} |\omega \cdot \ell| \geq \gamma > 0.$$

Assume that there exists  $G \in \mathcal{F}^{\geq 1}$  such that

$$e^{\{G, \cdot\}} H = \sum_{j \in \mathbb{Z}} \omega_j |u_j|^2,$$

then there exists  $r_1 < \mathfrak{r}$ , and a close to identity change of variables  $\Psi$

$$\Psi : B_{r_1}(\mathfrak{h}_0) \rightarrow \mathfrak{h}_0$$

such that  $H \circ \Psi = \sum_{j \in \mathbb{Z}} \omega_j |u_j|^2$ .

*Remark 3.4.* The main result of [BS20] concerns the simultaneous normalization of an infinite family of commuting vector fields in infinite dimension. One of its main feature is that the natural condition to consider is a *small divisors of the family* of linear parts and not of individual members of the family. It happens that families considered in the aforementioned article do not have small divisors. If such a family were formally linearizable despite resonances, it would be also holomorphically so in the same spirit as the previous theorem.

*3.2. Poisson structure and homological equation.* The following Lemmata are proved in [BMP18] under the extra assumption of mass conservation, we discuss the proof in our slightly more general setting in the appendix.

**Lemma 3.5.** If  $H \in \mathcal{H}_{r,s} \cap \mathcal{F}^{\geq d}$ , then for all  $r^* \leq r$  one has

$$\|H\|_{r^*,s} \leq \left(\frac{r^*}{r}\right)^d \|H\|_{r,s}.$$

**Lemma 3.6.** If  $H \in \mathcal{H}_{r,s}$ , then for all  $s_1 \geq s$  one has

$$\|H\|_{r,s_1} \leq \|H\|_{r,s}.$$

**Lemma 3.7.** (Poisson brackets and Hamiltonian flow) Let  $0 < \rho < r$ , and  $F, G \in \mathcal{H}_{r+\rho,\eta}(\mathfrak{h}_s)$ , then

$$\| \{F, G\} \|_{r,s} \leq 4 \left(1 + \frac{r}{\rho}\right) \|F\|_{r+\rho,s} \|G\|_{r+\rho,s}. \quad (20)$$

For  $S \in \mathcal{H}_{r+\rho,\eta}(\mathfrak{h}_s)$  with

$$\|S\|_{r+\rho,s} \leq \delta := \frac{\rho}{8e(r+\rho)}. \quad (21)$$

Then the time 1-Hamiltonian flow  $\Psi_S^1 : B_r(\mathfrak{h}_s) \rightarrow B_{r+\rho}(\mathfrak{h}_s)$  is well defined, analytic, symplectic with

$$\sup_{u \in B_r(\mathfrak{h}_s)} \left\| \Psi_S^1(u) - u \right\|_{\mathfrak{h}_s} \leq (r + \rho) \|S\|_{r+\rho, s} \leq \frac{\rho}{8e}. \quad (22)$$

For any  $H \in \mathcal{H}_{r+\rho, s}$  we have that  $H \circ \Psi_S^1 = e^{\{S, \cdot\}} H \in \mathcal{H}_{r, s}$  and

$$\left\| e^{\{S, \cdot\}} H \right\|_{r, s} \leq 2 \|H\|_{r+\rho, s}, \quad (23)$$

$$\left\| \left( e^{\{S, \cdot\}} - \text{id} \right) H \right\|_{r, s} \leq \delta^{-1} \|S\|_{r+\rho, s} \|H\|_{r+\rho, s}, \quad (24)$$

$$\left\| \left( e^{\{S, \cdot\}} - \text{id} - \{S, \cdot\} \right) H \right\|_{r, s} \leq \frac{1}{2} \delta^{-2} \|S\|_{r+\rho, s}^2 \|H\|_{r+\rho, s} \quad (25)$$

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

$$\left\| \sum_{k \geq h} c_k \text{ad}_S^k(H) \right\|_{r, s} \leq 2 \|H\|_{r+\rho, s} (\|S\|_{r+\rho, s} / 2\delta)^h, \quad (26)$$

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

**Lemma 3.8.** Fix  $s \geq 0$  and  $\sigma > 0$  and  $\omega \in \mathcal{D}_\gamma$ . For any  $R \in \mathcal{H}_{r, s}^d$  with  $d \geq 1$  and such that  $\Pi_{\mathcal{K}} R = 0$ , the Homological equation  $L_\omega S = R$  has a unique solution  $S = L_\omega^{-1} R \in \mathcal{H}_{r, s+\sigma}^d$  such that  $\Pi_{\mathcal{K}} S = 0$  and moreover

$$\left\| L_\omega^{-1} R \right\|_{r, s+\sigma} \leq \gamma^{-1} e^{\mathcal{C}_1 \sigma^{-\frac{3}{\theta}}} \|R\|_{r, s} \quad (27)$$

**3.3. Proof of the main Theorem.** The theorem follows by the following holomorphic version of Lemma 2.19. If  $H_0 \in \mathcal{F} \cap \mathcal{H}_{r, s_0}$  of the form  $D_\omega + P_0$  with  $P_0 \in \mathcal{F}^{\geq 1}$  is formally linearizable.

Fix  $0 < r_0 < r$  and  $s_0 > 0$  so that

$$\varepsilon_0 := \gamma^{-1} \|P_0\|_{r_0, s_0} \leq \gamma^{-1} \frac{r_0}{r} \|P_0\|_{r, s_0}$$

is appropriately small. More precisely, fix  $C = 1 + \pi^2/6$  and assume

$$\varepsilon_0^{-1} \geq \mathbb{K} \sup_n e^{\mathcal{C}_2(s_0)n^{\frac{6}{\theta}}} n^2 \max(e^{n-\chi^n}, e^{-(2-\chi)\chi^n}), \quad (28)$$

where  $\mathbb{K}$  is an appropriately large absolute constant while  $\mathcal{C}_2(s_0) = \mathcal{C}_1 C^{\frac{3}{\theta}} s_0^{-\frac{3}{\theta}}$ .

Let

$$r_i = r_{i-1} - \rho_{i-1}, \quad s_i = s_{i-1} + \sigma_{i-1}, \quad \mathfrak{d}_i = 2^i, \quad \rho_i = \frac{r_0}{2C \langle i \rangle^2}, \quad \sigma_i = \frac{s_0}{C \langle i \rangle^2}$$

so that  $r_i \rightarrow r_0/2$  and  $s_i \rightarrow 2s_0$ .

Fix  $1 < \chi < 2$  such that<sup>4</sup>

$$\sup_{n \geq 0} 2^{n+1} \ln\left(1 - \frac{1}{2Cn^2}\right) + \chi^n (\chi - 1) \leq -0.1 \quad (29)$$

<sup>4</sup> for example if  $\chi = 15/14$  the sup on the left hand side is smaller than  $-0.2$ .

**Lemma 3.9.** *The sequence of generating functions and Hamiltonians of Lemma 2.19.*

$$\{D_\omega, S_i\} = \Pi^{\mathcal{R}} H_i, \quad H_i = e^{\{S_{i-1}, \cdot\}} H_{i-1}.$$

satisfies

$$H_i = D_\omega + P_i, \quad P_i \in \mathcal{F}^{\geq d_i} \cap \mathcal{H}_{r_i, s_i}.$$

with the bounds

$$\|S_{i-1}\|_{r_{i-1}, s_i} \leq \gamma^{-1} e^{C_1 \sigma_{i-1}^{-\frac{3}{\theta}}} \|P_{i-1}\|_{r_{i-1}, s_{i-1}}, \quad \|P_i\|_{r_i, s_i} \leq \|P_0\|_{r_0, s_0} e^{-\chi^i}.$$

Moreover each  $S_{i-1}$  defines a symplectic analytic change of variables  $\Psi_{i-1} : B_{r_i}(\mathfrak{h}_s) \rightarrow B_{r_{i-1}}(\mathfrak{h}_s)$  for all  $s \geq s_i$  satisfying

$$\sup_{|u|_s \leq r_i} |\Psi_i(u) - u|_s \leq 2^{-i} r_0 \quad (30)$$

Finally setting

$$\Phi_i = \Psi_1 \circ \Psi_2 \circ \dots \circ \Psi_i$$

we have that  $\Phi_i \rightarrow \Phi_\infty$  where  $\Phi_\infty$  is an invertible symplectic map  $B_{r_0/2}(\mathfrak{h}_{2s_0}) \rightarrow B_{r_0}(\mathfrak{h}_{2s_0})$  such that

$$H_0 \circ \Phi_\infty = D_\omega$$

*Proof.* By induction. Let us denote  $\gamma^{-1} \|P_0\|_{r_0, s_0} := \varepsilon_0$ . Fix  $k \geq 0$  and assume that for all  $i \leq k$  the Lemma holds. By definition

$$S_k = \text{ad}_{D_\omega}^{-1} \Pi^{\mathcal{R}} P_k.$$

For all  $s \geq s_k + \sigma_k \equiv s_{k+1}$ , by Lemma 3.8 and (28)

$$\|S_k\|_{r_k, s} \leq \|S_k\|_{r_k, s_{k+1}} \leq \gamma^{-1} e^{C_1 \sigma_k^{-\frac{3}{\theta}}} \|P_k\|_{r_k, s_k} \leq \varepsilon_0 e^{C_2 (s_0) k \frac{6}{\theta}} e^{-\chi^k} \leq \frac{1}{16e2Ck^2} \leq \frac{\rho_k}{8er_k}$$

so, by Lemma 3.7 the time one flow  $\Psi_{S_k}^1 : B_{r_{k+1}}(\mathfrak{h}_s) \rightarrow B_{r_k}(\mathfrak{h}_s)$  is well defined analytic, symplectic and, by (22) satisfies

$$\sup_{u \in B_{r_{k+1}}(\mathfrak{h}_s)} \left| \Phi_{S_k}^1(u) - u \right|_{\mathfrak{h}_s} \leq r_k \|S_k\|_{r_k, s} \leq C \varepsilon_0 r_0 k^{-2} e^{C_2 (s_0) k \frac{6}{\theta}} e^{-\chi^k} \stackrel{(28)}{\leq} 2^{-k} r_0. \quad (31)$$

Recalling that

$$\begin{aligned} H_{k+1} &:= e^{\{S_k, \cdot\}} H_k = D_\omega + P_k + \{S_k, D_\omega\} + \sum_{h=2}^{\infty} \frac{\text{ad}_{S_k}^{h-1}}{h!} \{S_k, D_\omega\} + \sum_{h=1}^{\infty} \frac{\text{ad}_{S_k}^h}{h!} P_k \\ &= D_\omega + \Pi^{\mathcal{K}} P_k - \sum_{h=1}^{\infty} \frac{\text{ad}_{S_k}^h}{(h+1)!} \Pi^{\mathcal{R}} P_k + \sum_{h=1}^{\infty} \frac{\text{ad}_{S_k}^h}{h!} P_k \\ &=: D_\omega + \Pi^{\langle 2^{k+1} \rangle} \Pi^{\mathcal{K}} P_k + P_{k+1}. \end{aligned}$$

and that in Lemma 2.19 we have proved that  $\Pi^{<2^{k+1}} \Pi^{\mathcal{K}} P_k = 0$ , we get

$$P_{k+1} = \Pi^{\geq 2^{k+1}} \Pi^{\mathcal{K}} P_k - \sum_{h=1}^{\infty} \frac{\text{ad}_{S_k}^h}{(h+1)!} \Pi^{\mathcal{R}} P_k + \sum_{h=1}^{\infty} \frac{\text{ad}_{S_k}^h}{h!} P_k$$

Now

$$\begin{aligned} \|\Pi^{\geq 2^{k+1}} \Pi^{\mathcal{K}} P_k\|_{r_{k+1}, s_{k+1}} &\leq \left(\frac{r_{k+1}}{r_k}\right)^{d_{k+1}} \|P_k\|_{r_k, s_k} \leq \varepsilon_0 \left(1 - \frac{1}{2Ck^2}\right)^{2^{k+1}} e^{-\chi^k} \\ \left\| \sum_{h=1}^{\infty} \frac{\text{ad}_{S_k}^h}{(h+1)!} \Pi^{\mathcal{R}} P_k + \sum_{h=1}^{\infty} \frac{\text{ad}_{S_k}^h}{h!} P_k \right\|_{r_{k+1}, s_{k+1}} &\leq \frac{16er_k}{\rho_k} \|P_k\|_{r_k, s_k} \|S_k\|_{r_{k+1}, s_{k+1}} \\ &\leq C\varepsilon_0^2 e^{C_2(s_0)k^{\frac{6}{\theta}}} e^{-2\chi^k} k^2 \end{aligned}$$

The bound on  $P_{k+1}$  follows from (28) and (29) which imply

$$\varepsilon_0 \left(1 - \frac{1}{2Ck^2}\right)^{2^{k+1}} e^{-\chi^k} + C\varepsilon_0^2 e^{C_2(s_0)k^{\frac{6}{\theta}}} e^{-2\chi^k} k^2 \leq \varepsilon_0 e^{-\chi^{k+1}}$$

In order to prove the convergence we remark that all the  $\Psi_i$  map  $B_{r_i}(h_{2s_0}) \rightarrow B_{r_{i-1}}(h_{2s_0})$ , consequently  $\Phi_i$  maps  $B_{r_i}(h_{2s_0}) \rightarrow B_{r_0}(h_{2s_0})$  and, by (30), it is a Cauchy sequence.  $\square$

*Proof of Theorem 3.3.* Since there are no small divisors, in Lemma 3.8 one can take  $\sigma = 0$  and one has  $\|L_{\omega}^{-1} R\|_{r, s+\sigma} \leq \gamma^{-1} \|R\|_{r, s}$ . Then the proof is essentially identical to the previous one except that we take  $s_0 = 0$  and  $C_2 = 0$ ; one follows the procedure of Lemma 3.9 (where of course all the  $\sigma_i = 0$ ) with the only difference that in the estimates on the generating functions one has the better bound  $\|S_i\|_{r_i, 0} \leq \gamma^{-1} \|R_i\|_{r_i, 0}$ .  $\square$

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Appendix A. Technical Lemmata

In the following, we adapt material from [BMP18] to non mass conservation situation. There are no new conceptual difficulties w.r.to [BMP18] but the proofs require a bit more care and some further case analysis. For the readers convenience, we have written also the proofs of a couple of technical lemmata from [BMP18] on which we rely.

*A.1. Proof of Lemmata 3.5 and 3.6.* We follow here [BMP18] [Appendix B. Proof of lemma 3.1]. For any  $H \in \mathcal{H}_{r, s}$  (we recall that this space depends on two extra parameters  $p \geq \frac{1}{2}$  and  $0 < \theta \leq 1$ ) we define a map

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

by setting

$$Y_H^{(j)}(y; r, s) := \sum_{*} |H_{\alpha, \beta}| \frac{(\alpha_j + \beta_j)}{2} c_{r, s}^{(j)}(\alpha, \beta) y^{\alpha + \beta - e_j} \quad (32)$$

where  $e_j$  is the  $j$ -th basis vector in  $\mathbb{N}^{\mathbb{Z}}$ , while the coefficient

$$c_{r, s}^{(j)}(\alpha, \beta) = r^{|\alpha| + |\beta| - 2} \left( \frac{\langle j \rangle^2}{\prod_i \langle i \rangle^{\alpha_i + \beta_i}} \right)^p e^{-s(\sum_i \langle i \rangle^{\theta} (\alpha_i + \beta_i) - 2\langle j \rangle^{\theta})} \quad (33)$$

For brevity, we set

$$\sum_{*} := \sum_{\alpha, \beta: \pi(\alpha, \beta) = 0} .$$

The vector field  $Y_H$  is a majorant analytic function on  $\ell^2$  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 A.1.** *Let  $r, r^* > 0$ ,  $s, s' \geq 0$ . The following properties hold.*

(1) *The norm of  $H$  can be expressed as*

$$\|H\|_{r, s} = \sup_{|y|_{\ell^2} \leq 1} |Y_H(y; r, s)|_{\ell^2} \quad (34)$$

(2) *Given  $H^{(1)} \in \mathcal{H}_{r^*, s'}$  and  $H^{(2)} \in \mathcal{H}_{r, s}$ , such that for all  $\alpha, \beta \in \mathbb{N}_f^{\mathbb{Z}}$  and  $j \in \mathbb{Z}$  with  $\alpha_j + \beta_j \neq 0$  one has*

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

*for some  $c > 0$ , then*

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

*Proof of Lemma 3.5.* Recalling (33), we have

$$\frac{c_{r^*, s'}^{(j)}(\alpha, \beta)}{c_{r, s}^{(j)}(\alpha, \beta)} = \left( \frac{r^*}{r} \right)^{|\alpha| + |\beta| - 2} .$$

Since  $|\alpha| + |\beta| - 2 \geq \bar{d}$ , the inequality follows by Lemma A.1 with  $H^{(1)} = H^{(2)}$  and  $s = s'$ .  $\square$

In order to prove Lemma 3.6 we need some notations and results proven in [Bou05] and [CLS $\bar{Y}$ ].

**Definition A.2.** Given a vector  $v = (v_i)_{i \in \mathbb{Z}} \in \mathbb{N}_f^{\mathbb{Z}}$  with  $|v| \geq 2$  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

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

*Remark A.3.* A good way of envisioning this list is as follows. Given an infinite set of variables  $(x_i)_{i \in \mathbb{Z}}$  and a vector  $v = (v_i)_{i \in \mathbb{Z}} \in \mathbb{N}_f^{\mathbb{Z}}$  consider the monomial  $x^v := \prod_i x_i^{v_i}$ . We can write

$$x^v = \prod_i x_i^{v_i} = x_{j_1} x_{j_2} \cdots x_{j_{|v|}}, \quad \text{with } j_k \in \mathbb{Z}$$

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

*Example A.4.* Let us set

$$v_{-1} = 2, v_0 = 3, v_1 = 1, v_3 = 1, v_4 = 2.$$

Hence, 1 is repeated 6 times, 3 is repeated 1 time, and 4 is repeated 2 times :

$$\widehat{n}_1 = 4, \widehat{n}_2 = 4, \widehat{n}_3 = 3, \widehat{n}_4 = \cdots = \widehat{n}_9 = 1$$

Given  $\alpha, \beta \in \mathbb{N}_f^{\mathbb{Z}}$  with  $|\alpha| + |\beta| \geq 2$  from now on we define

$$\widehat{n} = \widehat{n}(\alpha + \beta) \quad \text{and set } N := |\alpha| + |\beta|$$

which is the cardinality of  $\widehat{n}$ . We observe that,  $N \geq 2$  and since

$$0 = \sum_{i \in \mathbb{Z}} i(\alpha_i - \beta_i) = \sum_{h>0} h(\alpha_h - \beta_h - \alpha_{-h} + \beta_{-h}), \quad (35)$$

there exists a choice of  $\sigma_i = \pm 1, 0$  such that<sup>5</sup>

$$\sum_l \sigma_l \widehat{n}_l = 0. \quad (36)$$

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

$$\widehat{n}_1 \leq \sum_{l \geq 2} \widehat{n}_l. \quad (37)$$

Indeed, if  $\sigma_1 = \pm 1$ , the inequality follows directly from (36); if  $\sigma_1 = 0$ , then  $\widehat{n}_1 = 1$  and consequently  $\widehat{n}_l = 1 \forall l$ . Since  $|\alpha| + |\beta| \geq 2$ , the list  $\widehat{n}$  has at least two elements, so the inequality is achieved.

**Lemma A.5.** *Given  $\alpha, \beta$  such that  $\sum_i i(\alpha_i - \beta_i) = 0$ , and  $|\alpha| + |\beta| \geq 2$ , we have that setting  $\widehat{n} = \widehat{n}(\alpha + \beta)$*

$$\sum_i \langle i \rangle^\theta (\alpha_i + \beta_i) = \sum_{l \geq 1} \widehat{n}_l^\theta \geq 2\widehat{n}_1^\theta + (2 - 2^\theta) \sum_{l \geq 3} \widehat{n}_l^\theta. \quad (38)$$

<sup>5</sup> A given  $h > 1$  appears  $\alpha_h + \beta_h + \alpha_{-h} + \beta_{-h}$  times in the list  $\widehat{n}$ . Thus in order to get the summand  $h(\alpha_h - \beta_h - \alpha_{-h} + \beta_{-h})$  we assign to the  $\widehat{n}_l$  with  $\widehat{n}_l = h$  the sign  $\sigma_l = +$ ,  $\alpha_h + \beta_{-h}$  times and the sign  $\sigma_l = -$ ,  $\alpha_{-h} + \beta_h$  times. Let us now consider the case  $h = 1$ . By construction, 1 appears  $\alpha^{(1)} + \beta^{(1)} + \alpha_{-1} + \beta_{-1} + \alpha_0 + \beta_0$  times in  $\widehat{n}$ . Thus in order to obtain the summand  $(\alpha^{(1)} - \beta^{(1)} - \alpha_{-1} + \beta_{-1})$  we assign to the  $\widehat{n}_l$  with  $\widehat{n}_l = 1$  the sign  $\sigma_l = +$ ,  $\alpha_1 + \beta_{-1}$  times, the sign  $\sigma_l = -$ ,  $\alpha_{-1} + \beta_1$  times and  $\sigma_l = 0$  the remaining  $\alpha_0 + \beta_0$  times.

*Proof.* The lemma above was proved in [Bou05] for  $\theta = \frac{1}{2}$  and for general  $0 < \theta < 1$  in [CLSY][Lemma 2.1], in the case of zero mass and momentum. Below we give a proof, using only momentum conservation.

We start by noticing that if  $|\alpha| + |\beta| = 2$  then  $\hat{n}$  has cardinality equal to two and (38) becomes  $\hat{n}_1 + \hat{n}_2 \geq 2\hat{n}_1$ . Now, by (37), momentum conservation implies that  $\hat{n}_1 = \hat{n}_2$  and hence (38).

If  $|\alpha| + |\beta| \geq 3$  we write

$$\sum_i \langle i \rangle^\theta (\alpha_i + \beta_i) - 2\hat{n}_1^\theta = \sum_{l \geq 2} \hat{n}_l^\theta - \hat{n}_1^\theta \geq \sum_{l \geq 2} \hat{n}_l^\theta - \left( \sum_{l \geq 2} \hat{n}_l \right)^\theta$$

since the cardinality of  $\hat{n}$  is at least three we may write

$$\sum_{l \geq 2} \hat{n}_l^\theta - \left( \sum_{l \geq 2} \hat{n}_l \right)^\theta = \hat{n}_2^\theta + \sum_{l \geq 3} \hat{n}_l^\theta - \left( \hat{n}_2 + \sum_{l \geq 3} \hat{n}_l \right)^\theta$$

Now setting, for  $x_i \geq 1, i = 2, \dots, N$ ,

$$f(x_2, \dots, x_N) := x_2^\theta + (2^\theta - 1) \sum_{l \geq 3} x_l^\theta - \left( x_2 + \sum_{l \geq 3} x_l \right)^\theta.$$

Hence, we have  $\partial_{x_2} f \geq 0$  for  $x_2 \geq x_3 \geq 1$ . Then

$$f(x_2, \dots, x_N) \geq f(x_3, x_3, x_4, \dots, x_N) =: f_3(x_3, \dots, x_N).$$

Now we set

$$\begin{aligned} f_n(x_n, \dots, x_N) &:= f(\underbrace{x_n, \dots, x_n}_{n-1}, x_{n+1}, \dots, x_N) \\ &= (1 + (2^\theta - 1)(n - 2))x_n^\theta + \sum_{\ell \geq n+1} x_\ell - ((n - 1)x_n + \sum_{\ell \geq n+1} x_\ell)^\theta \end{aligned}$$

so that  $f(x_2, \dots, x_N) \geq f_3(x_3, \dots, x_N)$ . Assume inductively that for some  $3 \leq n < N$ , one has  $f(x_2, \dots, x_N) \geq f_3(x_3, \dots, x_N) \geq \dots \geq f_n(x_n, \dots, x_N)$ . By direct computation<sup>6</sup>

$$\begin{aligned} \partial_{x_n} f_n &= \theta \left[ \frac{(1 + (2^\theta - 1)(n - 2))}{x_n^{1-\theta}} - \frac{n - 1}{((n - 1)x_n + \sum_{\ell \geq n+1} x_\ell)^{1-\theta}} \right] \\ &\geq \theta x_n^{\theta-1} \left[ (1 + (2^\theta - 1)(n + 2)) - (n - 1)^\theta \right] \geq 0, \end{aligned}$$

so that the minimum is attained in  $x_n = x_{n+1}$  and  $f(x_2, \dots, x_N) \geq f_{n+1}(x_{n+1}, \dots, x_N)$ . In conclusion

$$f(x_2, \dots, x_N) \geq f(x_N, \dots, x_N) \geq 0$$

where the last inequality follows by recalling that  $1 + (2^\theta - 1)k - (k + 1)^\theta \geq 0$  for  $k \geq 1$ .  $\square$

<sup>6</sup> recalling that the  $x_\ell > 0$  and that  $1 + (2^\theta - 1)k - (k + 1)^\theta \geq 0$ , with  $k = n + 2 > 1$ .

The Lemma proved above, is fundamental in discussing the properties of  $\mathcal{H}_r(\mathfrak{h}_{p,s,a})$  with  $s > 0$ , indeed it implies

$$\sum_i \langle i \rangle^\theta (\alpha_i + \beta_i) - 2\langle j \rangle^\theta \geq (2 - 2^\theta) \left( \sum_{l \geq 3} \widehat{n}_l^\theta \right) \geq 0 \quad (39)$$

for all  $\alpha, \beta$  such that  $\alpha_j + \beta_j \neq 0$ . Indeed, this follows from the fact that  $\langle j \rangle \leq \widehat{n}_1$ .

*Proof of Lemma 3.6.* In all that follows we shall use systematically the fact that our Hamiltonians are momentum preserving, are zero at the origin and have no linear term so that  $|\alpha| + |\beta| \geq 2$ .

We need to show that

$$\frac{c_{r,s+\sigma}^{(j)}(\alpha, \beta)}{c_{r,s}^{(j)}(\alpha, \beta)} = \exp\left(-\sigma \left( \sum_i \langle i \rangle^\theta (\alpha_i + \beta_i) - 2\langle j \rangle^\theta \right)\right) \leq 1. \quad (40)$$

The first identity comes from (33), while the last inequality follows by (39) of Lemma A.5  $\square$

*A.2. Proof of Lemma 3.7.* This is a rather classical result, the proof we give is taken from [BMP18], where it is stated under the extra hypothesis of mass conservation, which is not used in the proof.

**Lemma A.6.** *Let  $0 < r_1 < r$ . Let  $E$  be a Banach space endowed with the norm  $|\cdot|_E$ . Let  $X : B_r \rightarrow E$  a vector field satisfying*

$$\sup_{B_r} |X|_E \leq \delta_0.$$

*Then the flow  $\Phi(u, t)$  of the vector field<sup>7</sup> is well defined for every*

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

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

$$|\Phi(u, t) - u|_E \leq \delta_0 |t|, \quad \forall |t| \leq T.$$

*Proof of Lemma 3.7.* The estimate for the Poisson bracket is proven in [BBP13]. In order to prove the other estimates we use Lemma A.6, with  $E \rightarrow \mathfrak{h}_s$ ,  $X \rightarrow X_S$ ,  $\delta_0 \rightarrow (r + \rho)|S|_{r+\rho}$ ,  $r \rightarrow r + \rho$ ,  $r_1 \rightarrow r$ ,  $T \rightarrow 8e$ . finally we do not write the dependence on  $s$  which is fixed.

Then the fact that the time 1-Hamiltonian flow  $\Phi_S^1 : B_r(\mathfrak{h}_s) \rightarrow B_{r+\rho}(\mathfrak{h}_s)$  is well defined, analytic, symplectic follows, since

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

<sup>7</sup> Namely the solution of the equation  $\partial_t \Phi(u, t) = X(\Phi(u, t))$  with initial datum  $\Phi(u, 0) = u$ .



Regarding the estimate (22), again by Lemma A.6 (choosing  $t = 1$ ), we get

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

Estimates (23), (24), (25) directly follow by (26) with  $h = 0, 1, 2$ , respectively and  $c_k = 1/k!$ , recalling that by Lie series

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

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

Let us prove (26). Fix  $k \in \mathbb{N}$ ,  $k > 0$  and set

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

Note that, by the immersion properties of the norm in Lemma 3.5,

$$\|S\|_{r_i} \leq \|S\|_{r+\rho}, \quad \forall i = 0, \dots, k. \quad (41)$$

Noting that

$$1 + \frac{kr_i}{\rho} \leq k \left(1 + \frac{r}{\rho}\right), \quad \forall i = 0, \dots, k, \quad (42)$$

by using  $k$  times (20) we have

$$\begin{aligned} \|H^{(k)}\|_r &= \|\{S, H^{(k-1)}\}\|_r \leq 4 \left(1 + \frac{kr}{\rho}\right) \|H^{(k-1)}\|_{r_{k-1}} \|S\|_{r_{k-1}} \\ &\stackrel{(41)}{\leq} \|H\|_{r+\rho} \|S\|_{r+\rho}^k 4^k \prod_{i=1}^k \left(1 + \frac{kr_i}{\rho}\right) \stackrel{(42)}{\leq} \|H\|_{r+\rho} \left(4k \left(1 + \frac{r}{\rho}\right) \|S\|_{r+\rho}\right)^k. \end{aligned}$$

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

$$\begin{aligned} \left\| \sum_{k \geq h} c_k H^{(k)} \right\|_r &\leq \sum_{k \geq h} |c_k| \|H^{(k)}\|_r \leq \|H\|_{r+\rho} \sum_{k \geq h} \left(4e \left(1 + \frac{r}{\rho}\right) \|S\|_{r+\rho}\right)^k \\ &= \|H\|_{r+\rho} \sum_{k \geq h} (\|S\|_{r+\rho}/2\delta)^k \stackrel{(21)}{\leq} 2 \|H\|_{r+\rho} (\|S\|_{r+\rho}/2\delta)^h. \end{aligned}$$

Finally, if  $S$  and  $H$  satisfy momentum conservation so does each  $\text{ad}_S^k H$ ,  $k \geq 1$ , hence  $H \circ \Phi_S^1$  too.  $\square$

A.3. *Proof of lemma 3.8.* Here we strongly use the fact that  $\omega_j \sim j^2$ . As we said in the introduction, it would not be hard to modify the proof in order to deal with  $\omega_j \sim j^\alpha$  with  $\alpha > 1$ , by adapting the proof of Lemmata A.9–A.10.

By Lemma A.1 (2), we have

$$\left\| L_\omega^{-1} R \right\|_{r,s+\sigma} \leq \gamma^{-1} K \|R\|_{r,s}$$

where

$$K = \gamma \sup_{\substack{j:\alpha_j+\beta_j \neq 0 \\ \pi(\alpha,\beta)=0}} \frac{e^{-\sigma(\sum_i \langle i \rangle^\theta (\alpha_i+\beta_i) - 2\langle j \rangle^\theta)}}{|\omega \cdot (\alpha - \beta)|}.$$

Therefore proving (27) amounts to showing that

$$K \leq e^{C_1 \sigma^{-\frac{3}{\theta}}}. \quad (43)$$

We divide in two cases regarding whether the inequality

$$\left| \sum_i (\alpha_i - \beta_i) i^2 \right| \leq 2 \sum_i |\alpha_i - \beta_i|, \quad (44)$$

holds or not. We remark that

$$\left| \sum_i (\alpha_i - \beta_i) i^2 \right| \geq 2 \sum_i |\alpha_i - \beta_i| \implies |\omega \cdot (\alpha - \beta)| \geq 1, \quad (45)$$

indeed denoting  $\omega_j = j^2 + \xi_j$  with  $|\xi_j| \leq \frac{1}{2}$ ,

$$|\omega \cdot (\alpha - \beta)| \geq 2 \sum_j |\alpha_j - \beta_j| - \frac{1}{2} \sum_j |\alpha_j - \beta_j| \geq 1.$$

Of course if  $|\omega \cdot (\alpha - \beta)| \geq 1$ , by (39) and (40) we get

$$\gamma \frac{e^{-\sigma(\sum_i \langle i \rangle^\theta (\alpha_i+\beta_i) - 2\langle j \rangle^\theta)}}{|\omega \cdot (\alpha - \beta)|} \leq 1$$

and the bound (43) is trivially achieved.

Otherwise, to deal with the case in which (44) holds, we need some notation. Given  $u \in \mathbb{Z}_f^{\mathbb{Z}}$ , consider the set

$$M(u) := \{j \neq 0, \text{ repeated } |u_j| \text{ times}\},$$

where  $D(u) < \infty$  is its cardinality. Define the vector  $m = m(u)$  as the reordering of the elements of the set above such that  $|m_1| \geq |m_2| \geq \dots \geq |m_D| \geq 1$ .

Given  $\alpha \neq \beta \in \mathbb{N}_f^{\mathbb{Z}}$  with  $|\alpha| + |\beta| \geq 3$  we consider  $m = m(\alpha - \beta)$  and  $\hat{n} = \hat{n}(\alpha + \beta)$ . If we denote by  $D$  the cardinality of  $m$  and  $N$  the one of  $\hat{n}$  we have

$$D + \alpha_0 + \beta_0 \leq N \quad (46)$$

and

$$(|m_1|, \dots, |m_D|, \underbrace{1, \dots, 1}_{N-D \text{ times}}) \leq (\hat{n}_1, \dots, \hat{n}_N). \quad (47)$$

*Example A.7.* Let set  $v = \alpha + \beta$  and  $u = \alpha - \beta$  with

$$\begin{aligned} \alpha_{-5} &= 1, \alpha_{-2} = 2, \alpha_0 = 2, \alpha_4 = 1 \\ \beta_{-5} &= 1, \beta_{-3} = 2, \beta_0 = 3, \beta_6 = 1 \\ \pi(\alpha, \beta) &= (-5)(1 - 1) + (-3)(-2) + (-2)(2) + 4(1) + 6(-1) = 0 \\ v_{-5} &= 2, v_{-3} = 2, v_{-2} = 2, v_0 = 5, v_4 = 1, v_6 = 1 \\ u_{-5} &= 0, u_{-3} = -2, u_{-2} = 2, u_0 = -1, u_4 = 1, u_6 = -1 \\ \widehat{n}(v) &= (6, 5, 5, 4, 3, 3, 2, 2, 1, 1, 1, 1, 1), N = 13(= \text{Card}(\widehat{n})) \\ M(u) &= \{-3, -3, -2, -2, 4, 6\}, m(u) = \{6, 4, -3, -3, -2, -2\}, D(u) = 6. \end{aligned}$$

Therefore, we have  $D(u) + \alpha_0 + \beta_0 = 8 \leq 13 = N(\widehat{n}(v))$ . Hence, (46) holds. Furthermore,  $(6, 4, 3, 3, 2, 2, 1, 1, 1, 1, 1, 1, 1) \leq \widehat{n}(v)$ , that is (47).

**Lemma A.8** (Lemma C.3 of [BMP18]). *Assume that  $g$  defined on  $\mathbb{Z}$  is non negative, even and not decreasing on  $\mathbb{N}$ . Then, if  $\alpha \neq \beta$ ,*

$$\sum_{i \in \mathbb{Z}} g(i) |\alpha_i - \beta_i| \leq 2g(m_1) + \sum_{l \geq 3} g(\widehat{n}_l). \quad (48)$$

*Proof.* By definition of  $m(\alpha - \beta)$  and setting  $\sigma_l = \text{sign}(\alpha_{m_l} - \beta_{m_l})$ , we have

$$\sum_{i \in \mathbb{Z}} g(i)(\alpha_i - \beta_i) = g(0)(\alpha_0 - \beta_0) + \sum_{l \geq 1} \sigma_l g(m_l). \quad (49)$$

Hence

$$\begin{aligned} \sum_{i \in \mathbb{Z}} g(i) |\alpha_i - \beta_i| &= g(0) |\alpha_0 - \beta_0| + \sum_{l \geq 1} g(m_l) \\ &\leq g(1)(\alpha_0 + \beta_0) + 2g(m_1) + \sum_{l \geq 3} g(m_l) \end{aligned}$$

and (48) follows by (46) and (47). □

By (49)

$$0 = \sum_{i \in \mathbb{Z}} (\alpha_i - \beta_i) i = \sum_l \sigma_l m_l \quad (50)$$

and

$$\sum_i (\alpha_i - \beta_i) i^2 = \sum_l \sigma_l m_l^2. \quad (51)$$

Analogously

$$\sum_i |\alpha_i - \beta_i| = D + |\alpha_0 - \beta_0| \stackrel{(46)}{\leq} N. \quad (52)$$

Finally note that

$$\sigma_l \sigma_{l'} = -1 \quad \implies \quad m_l \neq m_{l'}. \quad (53)$$

**Lemma A.9.** Given  $\alpha \neq \beta \in \mathbb{N}_f^{\mathbb{Z}}$ , such that  $\pi(\alpha - \beta) = 0$ ,  $N \geq 3$ ,  $D \geq 1$  and satisfying (44), we have

$$|m_1| \leq 7 \sum_{l \geq 3} \widehat{n}_l^2. \quad (54)$$

*Proof.* The case  $D = 1$  is not compatible with momentum conservation. Let us now consider the case  $D = 2$ , i.e.

$$\alpha - \beta = \sigma_1 \mathbf{e}_{m_1} + \sigma_2 \mathbf{e}_{m_2} + (\alpha_0 - \beta_0) \mathbf{e}_0.$$

If  $\sigma_1 \sigma_2 = -1$ , momentum conservation imposes  $m_1 = m_2$  but this contradicts (53). In the case  $\sigma_1 \sigma_2 = 1$ , by momentum conservation we have  $m_1 = -m_2$ . Then conditions (44) and (52) imply that

$$m_1^2 + m_2^2 \leq 2(D + |\alpha_0 - \beta_0|) \stackrel{(52)}{\leq} 2N \leq 6(N - 2) \leq 6 \sum_{l=3}^N \widehat{n}_l^2$$

since  $\widehat{n}_l \geq 1$ .

Let us now consider the case  $D \geq 3$ . By (44), (51) and (52)

$$m_1^2 + \sigma_1 \sigma_2 m_2^2 \leq 2(D + |\alpha_0 - \beta_0|) + \sum_{l=3}^D m_l^2 \leq 2N + \sum_{l=3}^D m_l^2 \leq 2N + \sum_{l=3}^N \widehat{n}_l^2 \leq 7 \sum_{l=3}^N \widehat{n}_l^2.$$

since (recall  $N \geq 3$ )  $2N \leq 6(N - 2) \leq 6 \sum_{l=3}^N \widehat{n}_l^2$ .

If  $\sigma_1 \sigma_2 = 1$  then

$$|m_1|, |m_2| \leq \sqrt{7 \sum_{l \geq 3} \widehat{n}_l^2}.$$

If  $\sigma_1 \sigma_2 = -1$

$$(|m_1| + |m_2|)(|m_1| - |m_2|) = m_1^2 - m_2^2 \leq 7 \sum_{l \geq 3} \widehat{n}_l^2.$$

Now, if  $|m_1| \neq |m_2|$  then

$$|m_1| + |m_2| \leq 7 \sum_{l \geq 3} \widehat{n}_l^2.$$

Conversely, if  $|m_1| = |m_2|$ , by (53),  $m_1 \neq m_2$ , hence  $m_1 = -m_2$ . By substituting this relation into (50), we have

$$2|m_1| \leq \sum_{l \geq 3} |m_l| \leq \sum_{l \geq 3} \widehat{n}_l^2,$$

concluding the proof.  $\square$

**Lemma A.10.** Consider  $\alpha, \beta \in \mathcal{M}$  with  $\alpha \neq \beta$  and  $|\alpha| + |\beta| \geq 3$ . If (44) holds then for all  $j$  such that  $\alpha_j + \beta_j \neq 0$  one has

$$\sum_i |\alpha_i - \beta_i| \langle i \rangle^{\theta/2} \leq C_* \left( \sum_i (\alpha_i + \beta_i) \langle i \rangle^\theta - 2 \langle j \rangle^\theta \right), \quad C_* = \frac{7}{2 - 2^\theta} \quad (55)$$

*Proof.* Let us first consider the case  $D = 0$ , this means that  $\alpha - \beta = (\alpha_0 - \beta_0)\mathbf{e}_0$  and the left hand side of (55) reads  $|\alpha_0 - \beta_0|$ . By (39) and  $N \geq 3$  the right hand side of (55) is at least  $2 - 2^\theta$ , so if  $|\alpha_0 - \beta_0| \leq 7$  the result is trivial. Otherwise we have two cases, if  $j = 0$

$$|\alpha_0 - \beta_0| \leq 2(|\alpha_0 - \beta_0| - 2 \langle j \rangle^\theta) \leq 2 \left( \sum_i (\alpha_i + \beta_i) \langle i \rangle^\theta - 2 \langle j \rangle^\theta \right),$$

Otherwise we remark that if  $j \neq 0$ ,  $\alpha_j + \beta_j \neq 0$  and  $\alpha_j - \beta_j = 0$ , then  $\alpha_j + \beta_j \geq 2$ , then

$$|\alpha_0 - \beta_0| \leq (\alpha_0 + \beta_0) + (\alpha_j + \beta_j - 2) \langle j \rangle^\theta \leq \sum_i (\alpha_i + \beta_i) \langle i \rangle^\theta - 2 \langle j \rangle^\theta.$$

Now we consider indices  $\alpha, \beta$  such that  $N \geq 3$ ,  $D \geq 1$ . Here we apply Lemma A.9 Given  $\alpha, \beta \in \mathbb{N}_f^{\mathbb{Z}}$ , as above we consider  $m = m(\alpha - \beta)$  and  $\hat{n} = \hat{n}(\alpha + \beta)$ .

We have<sup>8</sup>

$$\begin{aligned} \sum_i |\alpha_i - \beta_i| \langle i \rangle^{\theta/2} &\stackrel{(48)}{\leq} 2|m_1|^{\frac{\theta}{2}} + \sum_{l \geq 3} \hat{n}_l^{\frac{\theta}{2}} \\ &\stackrel{(54)}{\leq} 2 \left( 7 \sum_{l \geq 3} \hat{n}_l^2 \right)^{\frac{\theta}{2}} + \sum_{l \geq 3} \hat{n}_l^{\frac{\theta}{2}} \\ &\leq +2(7)^{\frac{\theta}{2}} \sum_{l \geq 3} \hat{n}_l^\theta + \sum_{l \geq 3} \hat{n}_l^{\frac{\theta}{2}} \\ &\leq \frac{2\sqrt{7} + 1}{2 - 2^\theta} \left( (2 - 2^\theta) \left( \sum_{l \geq 3} \hat{n}_l^\theta \right) \right), \end{aligned} \quad (56)$$

Then by Lemma A.5 and (56) we get

$$\begin{aligned} \sum_i |\alpha_i - \beta_i| \langle i \rangle^{\theta/2} &\leq \frac{7}{2 - 2^\theta} \left( \sum_i \langle i \rangle^\theta (\alpha_i + \beta_i) - 2\hat{n}_1^\theta \right) \\ &\leq \frac{7}{2 - 2^\theta} \left[ \sum_i \langle i \rangle^\theta (\alpha_i + \beta_i) - 2 \langle j \rangle^\theta \right], \end{aligned}$$

proving (55). □

<sup>8</sup> Using that for  $x, y \geq 0$  and  $0 \leq c \leq 1$  we get  $(x + y)^c \leq x^c + y^c$ .

*Conclusion of the proof of Lemma 3.8* By applying Lemma A.10, since  $\omega \in \mathcal{D}_\gamma$  we get:

$$\begin{aligned}
\gamma \frac{e^{-\sigma(\sum_i \langle i \rangle^\theta (\alpha_i + \beta_i) - 2\langle j \rangle^\theta)}}{|\omega \cdot (\alpha - \beta)|} &\stackrel{(4)}{\leq} e^{-\sigma(\sum_i \langle i \rangle^\theta (\alpha_i + \beta_i) - 2\langle j \rangle^\theta)} \prod_i \left(1 + (\alpha_i - \beta_i)^2 \langle i \rangle^2\right) \\
&\stackrel{(55)}{\leq} e^{-\frac{\sigma}{C_*} \sum_i |\alpha_i - \beta_i| \langle i \rangle^{\frac{\theta}{2}}} \prod_i \left(1 + (\alpha_i - \beta_i)^2 \langle i \rangle^2\right) \\
&\leq \exp \sum_i \left[ -\frac{\sigma}{C_*} |\alpha_i - \beta_i| \langle i \rangle^{\frac{\theta}{2}} + \ln \left(1 + (\alpha_i - \beta_i)^2 \langle i \rangle^2\right) \right] \\
&= \exp \sum_i f_i(|\alpha_i - \beta_i|) \tag{57}
\end{aligned}$$

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

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

□

Finally, we have

**Lemma A.11.** (Lemma 7.2 of [BMP18]) *Setting*

$$i_\sharp := \left( \frac{24C_*}{\sigma\theta} \ln \frac{12C_*}{\sigma\theta} \right)^{\frac{2}{\theta}},$$

we get

$$\sum_i f_i(|\ell_i|) \leq 18i_\sharp \ln i_\sharp \tag{58}$$

for every  $\ell \in \mathbb{Z}_f^{\mathbb{Z}}$ .

*Proof.* 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<sup>9</sup>

$$f_i(x) \leq -\frac{\sigma}{C_*} \langle i \rangle^{\frac{\theta}{2}} x + 2 \ln(x) + 2 \ln \langle i \rangle + 1, \quad \forall x \geq 1.$$

Now,

$$\max_{x \geq 1} \left( -\frac{\sigma}{C_*} \langle i \rangle^{\frac{\theta}{2}} x + 2 \ln(x) \right) = \begin{cases} -\frac{\sigma}{C_*} \langle i \rangle^{\frac{\theta}{2}} & \text{if } \langle i \rangle \geq i_0, \\ -2 + 2 \ln \frac{2C_*}{\sigma} - \theta \ln \langle i \rangle & \text{if } \langle i \rangle < i_0, \end{cases}$$

where

<sup>9</sup> Using that  $\ln(1+y) \leq 1 + \ln y$  for every  $y \geq 1$ .

$$i_0 := \left( \frac{2C_*}{\sigma} \right)^{\frac{2}{\theta}},$$

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

$$\begin{aligned} \sum_i f_i(|\ell_i|) &= \sum_{i \text{ s.t. } \ell_i \neq 0} f_i(|\ell_i|) \\ &\leq \sum_{\langle i \rangle \geq i_0 \text{ s.t. } \ell_i \neq 0} \left( 2 \ln \langle i \rangle + 1 - \frac{\sigma}{C_*} \langle i \rangle^{\frac{\theta}{2}} \right) + \sum_{\langle i \rangle < i_0} \left( 2 \ln \frac{2C_*}{\sigma} + (2 - \theta) \ln \langle i \rangle \right). \end{aligned}$$

We immediately have that

$$\begin{aligned} \sum_{\langle i \rangle < i_0} \left( 2 \ln \frac{2C_*}{\sigma} + (2 - \theta) \ln \langle i \rangle \right) &\leq 6i_0 \left( \ln \frac{2C_*}{\sigma} + \ln i_0 \right) \\ &= 6 \left( 1 + \frac{2}{\theta} \right) \left( \frac{2C_*}{\sigma} \right)^{\frac{2}{\theta}} \ln \frac{2C_*}{\sigma}. \end{aligned}$$

Moreover, in the case  $\langle i \rangle \geq i_0 \geq e$ ,

$$2 \ln \langle i \rangle + 1 - \frac{\sigma}{C_*} \langle i \rangle^{\frac{\theta}{2}} \leq 3 \ln \langle i \rangle - \frac{\sigma}{C_*} \langle i \rangle^{\frac{\theta}{2}} = \frac{6}{\theta} \left( \ln \langle i \rangle^{\frac{\theta}{2}} - 2\mathfrak{C} \langle i \rangle^{\frac{\theta}{2}} \right)$$

where

$$\mathfrak{C} := \frac{\theta \sigma (2 - 2^\theta)}{84} < 1.$$

We have that<sup>10</sup>

$$\ln \langle i \rangle^{\frac{\theta}{2}} - 2\mathfrak{C} \langle i \rangle^{\frac{\theta}{2}} \leq -\mathfrak{C} \langle i \rangle^{\frac{\theta}{2}}, \quad \text{when} \quad \langle i \rangle \geq i_* := \left( \frac{2}{\mathfrak{C}} \ln \frac{1}{\mathfrak{C}} \right)^{\frac{2}{\theta}}.$$

Note that

$$i_{\sharp} \geq \max\{i_0, i_*\}.$$

Therefore

$$\begin{aligned} \sum_{\langle i \rangle \geq i_0 \text{ s.t. } \ell_i \neq 0} \left( 2 \ln \langle i \rangle + 1 - \frac{\sigma}{C_*} \langle i \rangle^{\frac{\theta}{2}} \right) &\leq \sum_{\langle i \rangle \geq i_0 \text{ s.t. } \ell_i \neq 0} \frac{6}{\theta} \left( \ln \langle i \rangle^{\frac{\theta}{2}} - 2\mathfrak{C} \langle i \rangle^{\frac{\theta}{2}} \right) \\ &\leq \frac{6}{\theta} \left( \sum_{\langle i \rangle < i_{\sharp}} \ln \langle i \rangle^{\frac{\theta}{2}} - \sum_{\langle i \rangle \geq i_{\sharp} \text{ s.t. } \ell_i \neq 0} \left( \mathfrak{C} \langle i \rangle^{\frac{\theta}{2}} \right) \right) \leq 9i_{\sharp} \ln i_{\sharp}. \end{aligned}$$

<sup>10</sup> Using that, for every fixed  $0 < \mathfrak{C} \leq 1$ , we have  $\mathfrak{C}x \geq \ln x$  for every  $x \geq \frac{2}{\mathfrak{C}} \ln \frac{1}{\mathfrak{C}}$ .

In conclusion we get

$$\begin{aligned} \sum_i f_i(|\ell_i|) &\leq 6 \frac{2+\theta}{\theta} \left(\frac{2C_*}{\sigma}\right)^{\frac{2}{\theta}} \ln \frac{2C_*}{\sigma} + 9i_{\#} \ln i_{\#} \\ &\leq 9 \left(\frac{2C_*}{\sigma\theta}\right)^{\frac{2}{\theta}} \ln \left(\frac{2C_*}{\sigma}\right)^{\frac{\theta}{2}} + 9i_{\#} \ln i_{\#} \leq 18i_{\#} \ln i_{\#} \end{aligned}$$

□

The inequality (27) follows from plugging (58) into (57) and evaluating the constant. □

## References

- [BBHM] Baldi, P., Berti, M., Haus, E., Montalto, R.: Time quasi-periodic gravity water waves in finite depth. *Invent. Math.* **214**(2), 739–911 (2018)
- [Bam99a] Bambusi, D.: Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. *Math. Z.* **230**(2), 345–387 (1999)
- [Bam99b] Bambusi, D.: On long time stability in Hamiltonian perturbations of nonresonant linear PDE's. *Nonlinearity* **12**, 823–850 (1999)
- [Bam03] Bambusi, D.: Birkhoff normal form for some nonlinear PDEs. *Commun. Math. Phys.* **234**(2), 253–285 (2003)
- [BDGS07] Bambusi, D., Delort, J.-M., Grébert, B., Szeftel, J.: Almost global existence for Hamiltonian semilinear Klein–Gordon equations with small Cauchy data on Zoll manifolds. *Commun. Pure Appl. Math.* **60**(11), 1665–1690 (2007)
- [BG03] Bambusi, D., Grébert, B.: Forme normale pour NLS en dimension quelconque. *C. R. Math. Acad. Sci. Paris* **337**(6), 409–414 (2003)
- [BG06] Bambusi, D., Grébert, B.: Birkhoff normal form for partial differential equations with tame modulus. *Duke Math. J.* **135**(3), 507–567 (2006)
- [BS20] Bambusi, D., Stolovitch, L.: Convergence to normal forms of integrable PDEs. *Comm. Math. Phys.* **376**(2), 1441–1470 (2020)
- [BCG] Benettin, G., Chierchia, L., Guzzo, M.: The steep Nekhoroshev's theorem. *Commun. Math. Phys.* **342**, 569–601 (2016)
- [BGG85] Benettin, G., Galgani, L., Giorgilli, A.: A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems. *Celest. Mech.* **37**(1), 1–25 (1985)
- [BFG88] Benettin, G., Fröhlich, J., Giorgilli, A.: A Nekhoroshev-type theorem for Hamiltonian systems with infinitely many degrees of freedom. *Commun. Math. Phys.* **119**(1), 95–108 (1988)
- [BKM18] Berti, M., Kappeler, Th., Montalto, R.: Large KAM tori for perturbations of the defocusing NLS equation. *Astérisque* (403), viii+148 (2018)
- [BBP13] Berti, M., Biasco, L., Procesi, M.: KAM theory for the Hamiltonian derivative wave equation. *Annales Scientifiques de l'ENS* **46**(2), 299–371 (2013)
- [BB15] Berti, M., Bolle, P.: Quasi-periodic solutions of nonlinear wave equations on the d-dimensional torus. In: *EMS Series of Lectures in Mathematics* (2020)
- [BBP10] Berti, M., Bolle, P., Procesi, M.: An abstract Nash–Moser theorem with parameters and applications to PDEs. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27**(1), 377–399 (2010)
- [BCP] Berti, M., Corsi, L., Procesi, M.: An abstract Nash–Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds. *Commun. Math. Phys.* (2014). <https://doi.org/10.1007/s00220-014-2128-4>
- [BFP] Berti M., Feola R., Pusateri F., Birkhoff normal form and long time existence for periodic gravity Water Waves, preprint [arXiv:1810.11549](https://arxiv.org/abs/1810.11549)
- [BD18] Berti, M., Delort, J.M.: *Almost Global Existence of Solutions for Capillarity-Gravity Water Waves Equations with Periodic Spatial Boundary Conditions*. Springer, Berlin (2018)
- [BM21] Berti, M., Franzoi, L., Maspero, A.: Traveling quasi-periodic water waves with constant vorticity. *ARMA* **240**(1), 99–202 (2021)
- [BMP18] Biasco, L., Massetti, J.E., Procesi, M.: An abstract Birkhoff normal form theorem and exponential type stability of the 1d NLS. *Commun. Math. Phys.* **375**(3), 2089–2153 (2020)
- [BMP21] Biasco, L., Massetti, J.E., Procesi, M.: Almost-periodic invariant tori for the NLS on the circle. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **38**(3), 711–758 (2021)



- [BFN15] Bounemoura, A., Fayad, B., Niederman, L.: Double exponential stability for generic real-analytic elliptic equilibrium points (2015). Preprint ArXiv [arXiv:1509.00285](https://arxiv.org/abs/1509.00285)
- [Bou98] Bourgain, J.: Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Ann. Math. (2)* **148**(2), 363–439 (1998)
- [Bou96a] Bourgain, J.: Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations. *Geom. Funct. Anal.* **6**(2), 201–230 (1996)
- [Bou05] Bourgain, J.: On invariant tori of full dimension for 1D periodic NLS. *J. Funct. Anal.* **229**(1), 62–94 (2005)
- [Bru72] Bruno, A.D.: Analytical form of differential equations. *Trans. Mosc. Math. Soc.* **25**, 131–288 (1971); **26**, 199–239 (1972), 1971–1972
- [CM18] Corsi, L., Montalto, R.: Quasi-periodic solutions for the forced Kirchhoff equation on  $\mathbb{T}^d$ . *Non-linearity* **31**(11), 5075–5109 (2018)
- [CLSY] Cong, H., Liu, J., Shi, Y., Yuan, X.: The stability of full dimensional KAM tori for nonlinear Schrödinger equation. *J. Differ. Equ.* **264**(7), 4504–4563 (2018)
- [CMW] Cong, H., Mi, L., Wang, P.: A Nekhoroshev type theorem for the derivative nonlinear Schrödinger equation. *J. Differ. Equ.* **268**(9), 5207–5256 (2020)
- [CY21] Cong, H., Yuan, X.: The existence of full dimensional invariant tori for 1-dimensional nonlinear wave equation. *Annales de l'Institut Henri Poincaré (C) Analyse Non Linéaire* **38**(3), 759–786 (2021)
- [CW93] Craig, W., Wayne, C.E.: Newton's method and periodic solutions of nonlinear wave equations. *Commun. Pure Appl. Math.* **46**(11), 1409–1498 (1993)
- [DS04] Delort, J.-M., Szeftel, J.: Long-time existence for small data nonlinear Klein–Gordon equations on tori and spheres. *Int. Math. Res. Not.* **37**, 1897–1966 (2004)
- [DS06] Delort, J.-M., Szeftel, J.: 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)
- [Del12] Delort, J.M.: A quasi-linear Birkhoff normal forms method. application to the quasi-linear Klein–Gordon equation on  $S^1$ . *Astérisque* **341**, vi+113 (2012)
- [D15] Delort, J.-M.: *Quasi-Linear Perturbations of Hamiltonian Klein–Gordon Equations on Spheres*. American Mathematical Society, Providence (2015)
- [EK10] Eliasson, L.H., Kuksin, S.B.: KAM for the nonlinear Schrödinger equation. *Ann. Math. (2)* **172**(1), 371–435 (2010)
- [FG13] Faou, E., Grébert, B.: A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus. *Anal. PDE* **6**(6), 1243–1262 (2013)
- [FG] Feola, R., Giuliani, F.: Quasi-periodic traveling waves on an infinitely deep perfect fluid under gravity. Preprint [arXiv:2005.08280](https://arxiv.org/abs/2005.08280) (2020)
- [FGP] Feola, R., Giuliani, F., Pasquali, S.: On the integrability of Degasperis–Procesi equation: control of the Sobolev norms and Birkhoff resonances. *J. Differ. Equ.* **266**, 3390–3437 (2019)
- [FGPr] Feola, R., Giuliani, F., Procesi, M.: Reducible KAM tori for the Degasperis–Procesi equation. *Commun. Math. Phys.* **377**, 1681–1759 (2020)
- [FI18] Feola, R., Iandoli, F.: Long time existence for fully nonlinear NLS with small Cauchy data on the circle. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **XXII**, 109–182 (2021)
- [FI20] Feola, R., Iandoli, F.: A non-linear Egorov theorem and Poincaré–Birkhoff normal forms for quasi-linear pdes on the circle. Preprint [arXiv:2002.1244](https://arxiv.org/abs/2002.1244) (2020)
- [GYX] Geng, J., Xu, X., You, J.: An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation. *Adv. Math.* **226**(6), 5361–5402 (2011)
- [IoL05] Iooss, G., Lombardi, E.: Polynomial normal forms with exponentially small remainder for analytic vector fields. *J. Differ. Equ.* **212**(1), 1–61 (2005)
- [KP03] Kappeler, Th., Pöschel, J.: *KdV & KAM, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3, vol. 45*. Springer, Berlin (2003)
- [K88] Kuksin, S.B.: Perturbation of conditionally periodic solutions of infinite-dimensional Hamiltonian systems. *Izv. Akad. Nauk SSSR Ser. Mat.* **52**(1), 41–63 (1988)
- [KP96] Kuksin, S., Pöschel, J.: Invariant cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. Math.* **143**(1), 149–179 (1996)
- [K00] Kuksin, S.B.: Analysis of Hamiltonian PDEs. In: *Oxford Lecture Series in Mathematics and its Applications*, vol. 19. Oxford University Press, Oxford (2000)
- [LS10] Lombardi, E., Stolovitch, L.: Normal forms of analytic perturbations of quasihomogeneous vector fields: rigidity, invariant analytic sets and exponentially small approximation. *Ann. Sci. Ec. Norm. Sup.* **43**, 659–718 (2010)
- [N77] Nehorošev, N.N.: An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. *Uspehi Mat. Nauk* **32**(198), 5–66 (1977)

- [Ni04] Niederman, L.: Exponential stability for small perturbations of steep integrable Hamiltonian systems. *Ergod. Theory Dyn. Syst.* **24**(2), 593–608 (2004)
- [Nik86] Nikolenko, N.V.: The method of Poincaré normal forms in problems of integrability of equations of evolution type. *Russ. Math. Surv.* **41**(5), 63–114 (1986)
- [Pös90] Pöschel, J.: Small divisors with spatial structure in infinite dimensional Hamiltonian systems. *Commun. Math. Phys.* **127**(2), 351–393 (1990)
- [Pös86] Pöschel, J.: On invariant manifolds of complex analytic mappings near fixed points. *Expo. Math.* **4**, 97–109 (1986)
- [Pös99] Pöschel, J.: On Nekhoroshev’s estimate at an elliptic equilibrium. *Int. Math. Res. Not.* **1999**(4), 203–215 (1999)
- [PP16] Procesi, C., Procesi, M.: Reducible quasi-periodic solutions for the non linear Schrödinger equation. *BUMI* **9**(2), 189 (2016)
- [Rue77] Rüssmann, H.: On the convergence of power series transformations of analytic mappings near a fixed point into a normal form. Preprint I.H.E.S., M/77/178, 1–44 (1977)
- [Se92] Serre, J.-P.: Lie algebras and lie groups. In: *Lecture Notes in Mathematics*, vol. 1500. Springer (1992)
- [Sie42] Siegel, C.L.: Iteration of analytic functions. *Ann. Math.* **2**(43), 607–612 (1942)
- [Sto00] Stolovitch, L.: Singular complete integrability. *Publ. Math. I.H.E.S.* **91**, 133–210 (2000)
- [Sto15] Stolovitch, L.: Family of intersecting totally real manifolds of  $(\mathbb{C}^n, 0)$  and germs of holomorphic diffeomorphisms. *Bull. Soc. Math. Fr.* **143**(1), 247–263 (2015)
- [Way90] Wayne, C.E.: Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory. *Commun. Math. Phys.* **127**(3), 479–528 (1990)
- [Y21] Yuan, X.: KAM theorem with normal frequencies of finite limit-points for some shallow water equations. *CPAM* **74**(6), 1193–1281 (2021)
- [YZ14] Yuan, Xiaoping, Zhang, Jing: Long time stability of Hamiltonian partial differential equations. *SIAM J. Math. Anal.* **46**(5), 3176–3222 (2014)
- [Zeh77] Zehnder, E.: C. L. Siegel’s linearization theorem in infinite dimensions. *Manuscr. Math.* **23**(4), 363–371 (1977/1978)

Communicated by C.Liverani