ANALYSIS & PDE

Volume 16 No. 5 2023

msp

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2023 is US $405/year for the electronic version, and $630/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 16 (2023), No. 5, pp. 1089–1132
DOI: 10.2140/apde.2023.16.1089

NONUNIFORM STABILITY OF DAMPED CONTRACTION SEMIGROUPS

RALPH CHILL, LASSI PAUNONEN, DAVID SEIFERT, REINHARD STAHN AND YURI TOMILOV

We investigate the stability properties of strongly continuous semigroups generated by operators of the form $A-BB^*$, where A is the generator of a contraction semigroup and B is a possibly unbounded operator. Such systems arise naturally in the study of hyperbolic partial differential equations with damping on the boundary or inside the spatial domain. As our main results we present general sufficient conditions for nonuniform stability of the semigroup generated by $A-BB^*$ in terms of selected observability-type conditions on the pair (B^*, A). The core of our approach consists of deriving resolvent estimates for the generator expressed in terms of these observability properties. We apply the abstract results to obtain rates of energy decay in one-dimensional and two-dimensional wave equations, a damped fractional Klein–Gordon equation and a weakly damped beam equation.

1. Introduction 1089
2. Preliminaries 1093
3. Frequency domain criteria for resolvent bounds and nonuniform stability 1102
4. Time-domain conditions for nonuniform stability 1110
5. Optimality of the decay rates 1115
6. Nonuniform stability of damped partial differential equations 1119
References 1129

1. Introduction

We study the stability properties of abstract differential equations of the form

$$\dot{x}(t) = (A - BB^*)x(t), \quad x(0) = x_0 \in X. \tag{1-1}$$

Here A generates a strongly continuous contraction semigroup, or typically a unitary group, on the Hilbert space X and B is a possibly unbounded operator, defined on a Hilbert space U. This class of dynamical systems includes several types of partial differential equations with damping, especially wave equations [Lebeau 1996; Ammari and Tucsnak 2001; Anantharaman and Léautaud 2014] and other hyperbolic PDE models [Liu and Zhang 2015; Dell'Oro and Pata 2021]. Equations of this form are also often encountered in control theory as a result of feedback interconnections and output feedback stabilisation [Slemrod 1974; Benchimol 1977; Guo and Luo 2002; Lasiecka and Triggiani 2003; Curtain and Weiss 2006; 2019].

The research of Paunonen is funded by the Academy of Finland grants 298182 and 310489. The work of Tomilov was partially supported the NCN grant UMO-2017/27/B/ST1/00078.
MSC2010: primary 47D06, 34D05, 47A10, 35L90; secondary 93D15, 35L05.
Keywords: nonuniform stability, strongly continuous semigroup, resolvent estimate, hyperbolic equation, observability, damped wave equation, Klein–Gordon equation, beam equation.

Our main interest is in studying stability properties of the semigroup $(T_B(t))_{t\geq 0}$ generated by $A - BB^*$ and the asymptotic behaviour of the solution $x(\cdot) = T_B(\cdot)x_0$ of (1-1). One of the key results concerning equations of the form (1-1) is that stability of $(T_B(t))_{t\geq 0}$ can be characterised in terms of *observability* of the pair (B^*, A); see [Slemrod 1974; Benchimol 1977; Curtain and Weiss 2006; 2019]. This relationship is well understood in the context of *exponential stability* and *strong stability*. In this paper we investigate this relationship for semigroups $(T_B(t))_{t\geq 0}$ which are *polynomially stable* or more generally *nonuniformly stable*. Our main results introduce new observability-type conditions which can be used to guarantee and verify the precise nonuniform stability properties of the differential equation (1-1).

The problem in (1-1) and the associated semigroup $(T_B(t))_{t\geq 0}$ are said to be (*uniformly*) *exponentially stable* if $\|x(t)\| \leq Me^{-\omega t}\|x_0\|$ for all $x_0 \in X$ and $t \geq 0$ and for some constants $M, \omega > 0$. The weaker notion of *strong stability* requires only that $\|x(t)\| \to 0$ for $t \to \infty$ for all $x_0 \in X$. The main benefit of exponential stability over strong stability is that the decay of the solutions takes place at a guaranteed *rate* as $t \to \infty$. In this paper we focus on *nonuniform stability* [Batty and Duyckaerts 2008; Borichev and Tomilov 2010; Rozendaal et al. 2019; Chill et al. 2020], where $(T_B(t))_{t\geq 0}$ is strongly stable and all *classical solutions* of (1-1) decay at a specific rate. Nonuniform and polynomial stability have been investigated in detail, especially for damped wave equations on multidimensional domains [Lebeau 1996; Liu and Rao 2005; Burq and Hitrik 2007; Anantharaman and Léautaud 2014; Stahn 2017; Cavalcanti et al. 2019; Datchev and Kleinhenz 2020], coupled partial differential equations [Duyckaerts 2007], and plate equations [Liu and Zhang 2015; Laurent and Léautaud 2021].

Under suitable assumptions on A and B, exponential stability of the semigroup $(T_B(t))_{t\geq 0}$ is equivalent to "exact observability" [Tucsnak and Weiss 2009, Chapter 6] of the pair (B^*, A) [Slemrod 1974; Curtain and Weiss 2006]. In addition, strong stability can be characterised in terms of "approximate observability" of (B^*, A) [Benchimol 1977]. In this paper we show that several modified concepts, each of which may be seen as "quantified approximate observability" of the pair (B^*, A), lead to nonuniform stability of the semigroup $(T_B(t))_{t\geq 0}$. In particular, we say that (B^*, A) satisfies the *nonuniform Hautus test* if there exist functions $M, m : \mathbb{R} \to [r_0, \infty)$ with $r_0 > 0$ such that [Miller 2012, Section 2.3]

$$\|x\|_X^2 \leq M(s)\|(is - A)x\|_X^2 + m(s)\|B^*x\|_U^2, \quad x \in D(A),\ s \in \mathbb{R}.$$

In addition, if A is skew-adjoint we say that the pair (B^*, A) satisfies the *wavepacket condition* if there exist bounded functions $\gamma, \delta : \mathbb{R} \to (0, \infty)$ such that [Miller 2012, Section 2.5]

$$\|B^*x\|_U \geq \gamma(s)\|x\|_X, \quad x \in \mathrm{WP}_{s,\delta(s)}(A),\ s \in \mathbb{R}. \tag{1-2}$$

Here $\mathrm{WP}_{s,\delta(s)}(A)$ denotes the spectral subspace of $-iA$ associated with the interval $(s - \delta(s), s + \delta(s))$ (elements of $\mathrm{WP}_{s,\delta(s)}(A)$ are called *wavepackets* of A).

The following theorem summarises our main results on these two observability concepts. The precise assumptions of Theorem 1.1 are stated in Assumption 2.1 in Section 2A, and they are automatically satisfied whenever A generates a strongly continuous contraction semigroup and $B \in \mathcal{L}(U, X)$. The results employ a function $\mu : \mathbb{R} \to [r_0, \infty)$, $r_0 > 0$, such that

$$\|B^*(1 + is - A)^{-1}B\| \leq \mu(s), \quad s \in \mathbb{R}. \tag{1-3}$$

As shown in Section 2A, we may always choose μ in such a way that $\mu(s) \lesssim 1+s^2,\ s \in \mathbb{R}$. Moreover, in the case where $B \in \mathcal{L}(U, X)$ and in many concrete applications μ may be taken to be constant. Finally, a measurable function $N : [0, \infty) \to (0, \infty)$ is said to have *positive increase* if there exist α, c_α, $s_0 > 0$ such that $N(\lambda s)/N(s) \geq c_\alpha \lambda^\alpha$ for all $\lambda \geq 1$ and $s \geq s_0$.

Theorem 1.1. *Assume that the operators A and B satisfy Assumption 2.1 and that $\mu : \mathbb{R} \to [r_0, \infty)$, $r_0 > 0$, is an even function such that* (1-3) *holds.*

If the pair (B^*, A) satisfies the nonuniform Hautus test for some continuous and even functions M and m, and if the function $N : [0, \infty) \to (0, \infty)$ defined by $N(\,\cdot\,) := M(\,\cdot\,)\mu(\,\cdot\,) + m(\,\cdot\,)\mu(\,\cdot\,)^2$ is strictly increasing and has positive increase, then $(T_B(t))_{t\geq 0}$ is nonuniformly stable and

$$\|T_B(t)x_0\| \leq \frac{C}{N^{-1}(t)} \|(A - BB^*)x_0\|, \quad x_0 \in D(A - BB^*),\ t \geq t_0, \tag{1-4}$$

for some C, $t_0 > 0$, where N^{-1} is the inverse function of N.

If A is skew-adjoint and (B^*, A) satisfies the wavepacket condition (1-2) *for continuous and even functions γ, δ such that $\gamma(\,\cdot\,)^{-1}\delta(\,\cdot\,)^{-1}$ is strictly increasing and has positive increase, then $(T_B(t))_{t\geq 0}$ is nonuniformly stable and* (1-4) *is satisfied for $N(\,\cdot\,) := \gamma(\,\cdot\,)^{-2}\delta(\,\cdot\,)^{-2}\mu(\,\cdot\,)^2$.*

Equations of the form (1-1) in particular include the damped second-order equation

$$\ddot{w}(t) + Lw(t) + DD^*\dot{w}(t) = 0, \quad w(0) \in H_{1/2},\ \dot{w}(0) \in H, \tag{1-5}$$

for a positive operator L on a Hilbert space H and $D \in \mathcal{L}(U, H_{-1/2})$, where $H_{1/2}$ is the domain of the fractional power $L^{1/2}$ and $H_{-1/2}$ is its dual with respect to the pivot space H. Nonuniform stability of such systems has been studied in the literature in the case where $D \in \mathcal{L}(U, H)$, and in particular it was shown in [Anantharaman and Léautaud 2014] and [Joly and Laurent 2020, Appendix B] that for such operators D the problem (1-1) is nonuniformly stable whenever the "Schrödinger group" generated by iL with the observation operator D^* is observable in a certain generalised sense. We subsequently refer to this property as the *Schrödinger group associated with the pair (D^*, iL) being observable*. In this paper we show that the same observability condition for the Schrödinger group generated by iL serves as a sufficient condition for the wavepacket condition and the nonuniform Hautus test for the pair (B^*, A). Moreover, our results generalise the results in [Anantharaman and Léautaud 2014, Theorem 2.3] and [Joly and Laurent 2020, Appendix B] to the case of general damping operators $D \in \mathcal{L}(U, H_{-1/2})$. Finally, the second part of Theorem 1.1 was proved in [Paunonen 2017, Theorem 6.3] in the special case where A is a diagonal operator with uniform spectral gap and $B \in \mathcal{L}(U, X)$.

As our last observability-type concept we introduce *nonuniform observability* of the pair (B^*, A), which requires that there exist $\beta \geq 0$ and τ, $c_\tau > 0$ such that

$$c_\tau \|(I - A)^{-\beta} x\|_X^2 \leq \int_0^\tau \|B^* T(t) x\|_U^2 \, dt, \quad x \in D(A), \tag{1-6}$$

where $(T(t))_{t\geq 0}$ is the contraction semigroup generated by A. Note that if $\beta = 0$, then nonuniform observability reduces to the classical notion of *exact observability* of (B^*, A). The main result of Section 4, Theorem 4.4, shows that if (B^*, A) is nonuniformly observable with parameter $\beta \in (0, 1]$ and if

$B \in \mathcal{L}(U, X)$, then the semigroup $(T_B(t))_{t\geq 0}$ is polynomially stable and (1-4) holds for $N^{-1}(t) = t^{1/(2\beta)}$. Related generalisations of exact observability have previously been used as sufficient conditions for nonuniform stability of damped second-order systems of the form (1-5) in [Ammari and Tucsnak 2001; Ammari and Nicaise 2015; Ammari et al. 2017]. Moreover, in the special case $\beta = \frac{1}{2}$, similar generalised observability conditions were used in [Russell 1975] and [Duyckaerts 2007, Section 5] to prove polynomial stability of (1-1). Finally, nonuniform stability of (1-5) for a special class of dampings satisfying $\|L^{-\beta}x\| \lesssim \|D^*x\| \lesssim \|L^{-\beta}x\|$ for some $\beta > 0$ and all $x \in X$ was studied in [Liu and Zhang 2015], and for $DD^* = f(L)$ with some function f in [Dell'Oro and Pata 2021]. In Section 4 we show that the assumptions in [Liu and Zhang 2015] imply nonuniform observability of the pair (B^*, A), and our results in particular establish a new proof of [loc. cit., Theorem 2.1].

The core of our approach in Sections 3 and 4 consists of deriving upper bounds for the resolvent norms $\|(is - A + BB^*)^{-1}\|$, $s \in \mathbb{R}$, in terms of the different types of observability-type condition. In Section 5 we address optimality of our results. In particular, we present an abstract result which describes how sharpness of the resolvent bound can be used to deduce optimality of the decay rate (1-4) of the semigroup $(T_B(t))_{t\geq 0}$. In addition, in the case where A is skew-adjoint we prove a lower bound for resolvent norms of $A - BB^*$ in terms of the restrictions of B^* to eigenspaces of A. Combining these two results allows us to prove that Theorem 1.1 is optimal in several situations of interest, and in particular if A has compact resolvent and uniformly separated eigenvalues.

In the last part of the paper we apply our main results to derive rates of energy decay for solutions of selected PDE models, namely wave equations on one- and two-dimensional spatial domains with different types of damping, a fractionally damped Klein–Gordon equation, and a weakly damped Euler–Bernoulli beam equation. In most of these examples the wavepackets are simply finite linear combinations of eigenfunctions [Tucsnak and Weiss 2009, Section 6.9]. In our one-dimensional wave and beam equations, the eigenvalues of A have a uniform spectral gap and, as a result, we obtain a particularly simple form of the wavepacket condition (1-2). Moreover, our general optimality results in Section 5 guarantee that the decay estimates we obtain in these cases are sharp. On the other hand, for two-dimensional wave equations with viscous damping our results are typically suboptimal. This is due to the phenomenon that in certain cases the smoothness of the damping profile improves the degree of polynomial stability [Burq and Hitrik 2007; Anantharaman and Léautaud 2014; Datchev and Kleinhenz 2020], whereas observability-type conditions do not in general distinguish between smooth and rough dampings. Indeed, comparing different types of viscous damping reveals natural limitations to optimality of decay rates derived from observability conditions, and we discuss this topic in detail in Section 6A.

The paper is organised as follows. In Section 2 we state the main assumptions on the operators A and B and recall essential results concerning nonuniform stability of strongly continuous semigroups. In Section 3 we present the main results showing that the nonuniform Hautus test and the wavepacket condition imply nonuniform stability of $(T_B(t))_{t\geq 0}$. In particular, in the second part of Section 3 we reformulate these results specifically for damped second-order systems, and present sufficient conditions for nonuniform stability of (1-5) based on observability of the Schrödinger group. Next, in Section 4 we show that nonuniform observability in the sense of (1-6) implies polynomial stability of $(T_B(t))_{t\geq 0}$.

In Section 5 we present a series of abstract results concerning optimality of the stability results in the previous sections. Finally, in Section 6 we study energy decay for several PDE models.

Notation. If X and Y are Banach spaces and $A : D(A) \subseteq X \to Y$ is a linear operator, we denote by $D(A)$, $\mathrm{Ker}(A)$ and $\mathrm{Ran}(A)$ the domain, kernel and range of A, respectively. Moreover, $\sigma(A)$, $\sigma_p(A)$, and $\rho(A)$ denote the spectrum, the point spectrum and the resolvent set of A, respectively. The space of bounded linear operators from X to Y is denoted by $\mathcal{L}(X, Y)$. The notation $X \hookrightarrow Y$ will mean that $X \subseteq Y$ with continuous and dense embedding. We denote the norm on a space X by $\|\cdot\|_X$ and its inner product by $\langle\cdot\,,\,\cdot\rangle_X$, and we omit the subscripts when there is no risk of ambiguity. We assume all our Banach and Hilbert spaces to be complex.

Let $\mathbb{R}_+ := [0, \infty)$, and denote the open right and left half-planes by $\mathbb{C}_+ = \{\lambda \in \mathbb{C} : \mathrm{Re}\,\lambda > 0\}$ and $\mathbb{C}_- = \{\lambda \in \mathbb{C} : \mathrm{Re}\,\lambda < 0\}$, respectively. We denote by χ_E the characteristic function of a set E. For two functions $f : E \subseteq \mathbb{R} \to \mathbb{R}_+$ and $g : \mathbb{R}_+ \to \mathbb{R}_+$ we write $f(t) = O(g(|t|))$ if there exist $C, t_0 > 0$ such that $f(t) \le Cg(|t|)$ whenever $|t| \ge t_0$. If in addition $g(t) > 0$ whenever $|t| \ge t_0$, we write $f(t) = o(g(|t|))$ if $f(t)/g(|t|) \to 0$ as $|t| \to \infty$. For real-valued quantities p and q, we use the notation $p \lesssim q$ if $p \le Cq$ for some constant $C > 0$ which is independent of all the parameters that are free to vary in the given situation. The notation $p \gtrsim q$ is defined analogously.

2. Preliminaries

2A. ***Standing assumptions and well-posedness.*** Let $A : D(A) \subseteq X \to X$ be the generator of a contraction semigroup $(T(t))_{t\ge 0}$ on a Hilbert space X. All semigroups considered in this paper are strongly continuous. For $\lambda_0 \in \rho(A)$ we equip $D(A)$ with the graph norm $\|x\|_1 = \|(\lambda_0 - A)x\|_X$, $x \in D(A)$, and denote the Hilbert space defined in this way by X_1. Defining X_{-1} as the completion of X with respect to the norm $\|x\|_{-1} = \|(\lambda_0 - A)^{-1}x\|_X$, we obtain a Hilbert space X_{-1} such that $X_1 \hookrightarrow X \hookrightarrow X_{-1}$. The operator A has a unique extension A_{-1} to X_{-1}, with domain $D(A_{-1}) = X$, and A_{-1} generates a contraction semigroup $(T_{-1}(t))_{t\ge 0}$ on X_{-1} which is unitarily equivalent to $(T(t))_{t\ge 0}$. In particular, $A_{-1} \in \mathcal{L}(X, X_{-1})$ and the operators A, A_{-1} are unitarily equivalent and thus have the same spectrum. Moreover, any $S \in \mathcal{L}(X)$ commuting with A has a (unique) continuous extension to an operator in $\mathcal{L}(X_{-1})$, unitarily equivalent to S; see [Tucsnak and Weiss 2009, Section 2.10].

To state our main assumptions, we let V be a Hilbert space such that $X_1 \subseteq V \subseteq X$ with continuous embeddings. In particular, V is dense in X and we consider the Gelfand triple $V \hookrightarrow X \hookrightarrow V^*$, where V^* is the dual of V with respect to the pivot space X [Tucsnak and Weiss 2009, Section 2.9]. We denote by $\langle\cdot\,,\,\cdot\rangle_{V^*,V} : V^* \times V \to \mathbb{C}$ the unique continuous extension of the inner product of X, and we define $V_A := \{x \in V : A_{-1}x \in V^*\}$. In the following we state our standing assumptions on the operators $A : D(A) \subseteq X \to X$ and $B \in \mathcal{L}(U, X_{-1})$, where U is another Hilbert space.

Assumption 2.1. The operators $A : D(A) \subseteq X \to X$ and $B \in \mathcal{L}(U, X_{-1})$ have the following properties.

(H1) The generator A of the contraction semigroup $(T(t))_{t\ge 0}$ satisfies $\mathrm{Re}\langle A_{-1}x, x\rangle_{V^*,V} \le 0$ for all $x \in V_A$.

(H2) We have $B \in \mathcal{L}(U, V^*)$ and $\mathrm{Ran}((\lambda_0 - A_{-1})^{-1}B) \subseteq V$ for some (or equivalently all) $\lambda_0 \in \rho(A)$.

Assumption 2.1 in particular requires that $\mathrm{Ran}(B) \subseteq X_{-1} \cap V^*$. Note that when A is not skew-adjoint, the space V^* is not necessarily contained in X_{-1}; it is instead a subspace of X^d_{-1}, the first extrapolation space for the adjoint A^* [Tucsnak and Weiss 2009, Section 2.10]. If $B \in \mathcal{L}(U, X)$, which we will refer to as B being *bounded*, then Assumption 2.1 is automatically satisfied for any generator A of a contraction semigroup $(T(t))_{t\geq 0}$ with the choices $V = V^* = X$.

We write $B^* \in \mathcal{L}(V, U)$ for the adjoint of $B \in \mathcal{L}(U, V^*)$, where V is identified with $(V^*)^*$ via the pivot duality through X. In particular,

$$\langle Bu, x\rangle_{V^*,V} = \langle u, B^*x\rangle_U, \quad x \in V,\ u \in U.$$

Moreover, (H2) in Assumption 2.1 and the closed graph theorem imply that $B^*(\lambda - A_{-1})^{-1}B \in \mathcal{L}(U)$ for all $\lambda \in \rho(A)$. We formally define the operator $A_B = A_{-1} - BB^*$ on X by

$$A_B x = A_{-1}x - BB^*x, \quad x \in D(A_B), \tag{2-1a}$$

$$D(A_B) = \{x \in V : A_{-1}x - BB^*x \in X\}. \tag{2-1b}$$

As shown in the following lemma, Assumption 2.1 guarantees that A_B generates a contraction semigroup $(T_B(t))_{t\geq 0}$ on X. In particular, the orbits of this semigroup are the solutions of the abstract Cauchy problem

$$\dot{x}(t) = A_B x(t), \quad t \geq 0, \tag{2-2a}$$

$$x(0) = x_0 \in X. \tag{2-2b}$$

For $x_0 \in X$ the orbit $x(\cdot) = T_B(\cdot)x_0$ is a *mild solution* of (2-2), and it is a *classical solution* if and only if $x_0 \in D(A_B)$ [Arendt et al. 2011, Chapter 3].

Lemma 2.2. *Let A and B satisfy Assumption 2.1. Then the operator A_B defined in* (2-1) *generates a strongly continuous contraction semigroup $(T_B(t))_{t\geq 0}$ on X. Moreover, we have $\rho(A)\cap\overline{\mathbb{C}}_+ \subseteq \rho(A_B)\cap\overline{\mathbb{C}}_+$,*

$$\mathrm{Re}\langle (is - A_B)x, x\rangle \geq \|B^*x\|^2, \quad s \in \mathbb{R},\ x \in D(A_B), \tag{2-3}$$

and

$$\|(\lambda - A_{-1})^{-1}B\|^2 \leq \frac{1}{\mathrm{Re}\,\lambda}\|B^*(\lambda - A_{-1})^{-1}B\|, \quad \lambda \in \mathbb{C}_+. \tag{2-4}$$

Proof. First note that if $x \in X$ and $u \in U$ are such that $A_{-1}x + Bu =: y \in X$, then condition (H2) implies that for any $\lambda_0 \in \rho(A)$ we have $x = (\lambda_0 - A_{-1})^{-1}(\lambda_0 x - y + Bu) \in V$ and $A_{-1}x = y - Bu \in V^*$. Thus $x \in V_A$ and condition (H1) implies that

$$\mathrm{Re}\langle A_{-1}x + Bu, x\rangle_X = \mathrm{Re}\langle A_{-1}x, x\rangle_{V^*,V} + \mathrm{Re}\langle Bu, x\rangle_{V^*,V} \tag{2-5a}$$

$$\leq \mathrm{Re}\langle B^*x, u\rangle_U. \tag{2-5b}$$

Let $s \in \mathbb{R}$ and $x \in D(A_B)$, and choose $u = -B^*x$. Then (2-5) immediately implies (2-3). In particular, A_B is dissipative.

To prove that $\rho(A)\cap\overline{\mathbb{C}}_+ \subseteq \rho(A_B)\cap\overline{\mathbb{C}}_+$, fix $\lambda \in \rho(A)\cap\overline{\mathbb{C}}_+$, let $u \in U$ and choose $x = (\lambda - A_{-1})^{-1}Bu$. Then $A_{-1}x + Bu = \lambda(\lambda - A_{-1})^{-1}Bu \in X$ and (2-5) implies that

$$(\mathrm{Re}\,\lambda)\|(\lambda - A_{-1})^{-1}Bu\|^2 \leq \mathrm{Re}\langle B^*(\lambda - A_{-1})^{-1}Bu, u\rangle.$$

In particular, this inequality implies (2-4). Moreover, this estimate shows that the operator $G(\lambda) := B^*(\lambda - A_{-1})^{-1}B \in \mathcal{L}(U)$ satisfies $\operatorname{Re} G(\lambda) \geq 0$, and consequently $I + G(\lambda)$ is boundedly invertible in $\mathcal{L}(U)$. A direct verification shows that $\lambda - A_B$ has bounded inverse given by

$$(\lambda - A_B)^{-1} = (\lambda - A_{-1})^{-1}(I - B(I + G(\lambda))^{-1}B^*(\lambda - A)^{-1}), \tag{2-6}$$

and we deduce the required spectral inclusion $\rho(A) \cap \overline{\mathbb{C}}_+ \subseteq \rho(A_B) \cap \overline{\mathbb{C}}_+$. In particular, A_B is closed. Since A_B is dissipative and $\mathbb{C}_+ \subseteq \rho(A_B)$, its domain is dense in X by [Tucsnak and Weiss 2009, Proposition 3.1.6]. Hence A_B is m-dissipative, and by the Lumer–Phillips theorem it generates a strongly continuous contraction semigroup on X. □

Remark 2.3. If Assumption 2.1 holds, then for every $\lambda \in \mathbb{C}_+$ the right-hand side of (2-6) extends uniquely to a mapping from the (not necessarily closed) subspace $X + \operatorname{Ran}(B)$ of X_{-1} to X, simply by replacing $(\lambda - A)^{-1}$ by $(\lambda - A_{-1})^{-1}$. We use this formula to define the extension of $(\lambda - A_B)^{-1}$ to an operator $(\lambda - A_B)^{-1} : X + \operatorname{Ran}(B) \to X$. In particular, we have

$$(\lambda - A_B)^{-1}B = (\lambda - A_{-1})^{-1}B(I + G(\lambda))^{-1} \in \mathcal{L}(U, X)$$

for $\lambda \in \mathbb{C}_+$. The identity $(\lambda - A_B)^{-1} = (I + (1 - \lambda)(\lambda - A_B)^{-1})(1 - A_B)^{-1}$ shows that also for arbitrary $\lambda \in \rho(A_B)$ the operator $(\lambda - A_B)^{-1}$ extends uniquely to a mapping from $X + \operatorname{Ran} B$ into X, and that $(\lambda - A_B)^{-1}B \in \mathcal{L}(U, X)$. For $\lambda \in \rho(A_B)$ and $u \in U$ we have $(\lambda - A_B)^{-1}Bu \in V$ and

$$(\lambda - A_{-1} + BB^*)(\lambda - A_B)^{-1}Bu = Bu,$$

and if $x \in V$ is such that $(\lambda - A_{-1} + BB^*)x \in X + \operatorname{Ran}(B)$ (in particular, if $x \in D(A)$), then

$$(\lambda - A_B)^{-1}(\lambda - A_{-1} + BB^*)x = x.$$

Remark 2.4. Define $X_B := D(A) + \operatorname{Ran}((\lambda_0 - A_{-1})^{-1}B)$, where $\lambda_0 \in \rho(A)$. The space X_B is independent of the choice of λ_0, and $X_B \subseteq V$ by Assumption 2.1. Moreover, the domain of A_B has the useful alternative characterisation

$$D(A_B) = \{x \in X_B : A_{-1}x + BB^*x \in X\}.$$

Here the nontrivial inclusion can be verified as in the beginning of the proof of Lemma 2.2.

Our results in Section 3 employ a parameter which describes the growth of the operator-valued function $\lambda \mapsto B^*(\lambda - A_{-1})^{-1}B$ on a vertical line in $\mathbb{C}_+$. In particular, we take $\mu : \mathbb{R} \to [r_0, \infty)$, $r_0 > 0$, to be a function such that

$$\|B^*(1 + is - A_{-1})^{-1}B\| \leq \mu(s), \quad s \in \mathbb{R}, \tag{2-7}$$

and the rate of growth of μ affects the resolvent estimates in our results. The following lemma shows that μ can be taken to be uniformly bounded whenever $B \in \mathcal{L}(U, X)$, and that estimate (2-7) always holds for a quadratic function μ.

Lemma 2.5. *If A and B satisfy Assumption 2.1, then the following hold*:

(a) *The estimate* (2-7) *holds for $\mu(s) = c(1 + s^2)$, $s \in \mathbb{R}$, for some $c > 0$.*

(b) *If* $B \in \mathcal{L}(U, X)$, *then* (2-7) *holds for* $\mu(s) \equiv c$ *with some* $c > 0$.

(c) *If* (2-7) *holds, then* $\|(1+is-A_{-1})^{-1}B\| \le \mu(s)^{1/2}$ *for* $s \in \mathbb{R}$.

Proof. Part (b) follows directly from the assumption that A generates a contraction semigroup, which implies that $\|(1+is-A)^{-1}\| \le 1$ for all $s \in \mathbb{R}$. Moreover, part (c) follows from (2-4) in Lemma 2.2. To prove part (a), fix $s \in \mathbb{R}$ and let $R = (1+is-A_{-1})^{-1}$. Using the identity $R = (I-A_{-1})^{-1} - is(I-A)^{-1}R$ we see that

$$\|B^*RB\| \le \|B^*(I-A_{-1})^{-1}B\| + |s|\,\|B^*(I-A)^{-1}\|\,\|RB\| \lesssim 1+|s|\|RB\|$$

and similarly

$$\|RB\| \le \|(I-A_{-1})^{-1}B\| + |s|\,\|(1+is-A)^{-1}\|\,\|(I-A_{-1})^{-1}B\| \lesssim 1+|s|.$$

Together these estimates give $\|B^*(1+is-A_{-1})^{-1}B\| \lesssim 1+s^2,\ s \in \mathbb{R}$. □

Estimates of the form (2-7) have been studied extensively in the control theory literature. In particular, for a *bounded* function μ the estimate in (2-7) is known as the property of *well-posedness* of the operator-valued "transfer function" $\lambda \mapsto B^*(\lambda - A_{-1})^{-1}B$; see [Salamon 1987; Guo and Luo 2002; Staffans 2002; Tucsnak and Weiss 2014]. This property has been verified in the literature for several different types of PDE systems; see for instance [Ammari and Tucsnak 2001; Guo and Luo 2002; Lasiecka and Triggiani 2003; Tucsnak and Weiss 2014; Ammari and Nicaise 2015]. As shown in the next lemma, validity of (2-7) for a bounded function μ moreover implies that B^* is an *admissible observation operator* for the semigroup $(T(t))_{t\ge 0}$, which is to say that $B^*T(\cdot)x \in L^2(0,\tau;U)$ for all $x \in D(A)$ and $\tau > 0$. This property will be useful in discussing the relationship between our results and existing results in the literature. In addition, the following lemma shows that under the same assumption B is an *admissible control operator* in the sense that $\int_0^\tau T_{-1}(\tau-t)Bu(t)\,dt \in X$ for all $u \in L^2(0,\tau;U)$ and $\tau > 0$.

Lemma 2.6. *Let A and B satisfy Assumption 2.1. If* (2-7) *is satisfied for a bounded function μ, then B and B^* are, respectively, admissible control and observation operators for the semigroup $(T(t))_{t\ge 0}$ generated by A.*

Proof. Since A and B satisfy Assumption 2.1, it is straightforward to verify that the operator $S : D(S) \subseteq X \times U \to X \times U$ defined by

$$S = \begin{pmatrix} A_{-1} & B \\ B^* & 0 \end{pmatrix}, \qquad D(S) = \left\{ \begin{pmatrix} x \\ u \end{pmatrix} \in X \times U : A_{-1}x + Bu \in X \right\}$$

is a system node on (U, X, U) in the sense of [Staffans 2002, Definition 2.1]. Moreover, estimate (2-5) for $(x, u) \in D(S)$ and [loc. cit., Theorem 4.2] imply that the system node S is impedance passive in the sense of [loc. cit., Definition 4.1]. The transfer function of the system node S is given by $G(\lambda) = B^*(\lambda - A_{-1})^{-1}B$ for $\lambda \in \rho(A)$. Hence the assumption that (2-7) is satisfied for a bounded function μ together with [loc. cit., Theorem 5.1] imply that the system node S is well-posed in the sense of [loc. cit., Definition 2.6]. In particular, $B \in \mathcal{L}(U, X_{-1})$ and $B^* \in \mathcal{L}(X_1, U)$ are, respectively, admissible control and observation operators for the semigroup generated by A. □

2B. ***Damped second-order problems.*** In this section we wish to use the framework introduced in Section 2A to study a class of abstract second-order equations with damping. To this end, we consider a positive self-adjoint and boundedly invertible operator $L : D(L) \subseteq H \to H$ on a Hilbert space H. We write H_1 for the domain of L equipped with the norm $\|x\|_{H_1} = \|Lx\|_H,\ x \in H_1$, and define $H_{1/2}$ to be the domain of the fractional power $L^{1/2}$ equipped with the norm $\|x\|_{H_{1/2}} = \|L^{1/2}x\|_H,\ x \in H_{1/2}$. We denote by $H_{-1/2}$ the dual of $H_{1/2}$ with respect to the pivot space H. For an operator $D \in \mathcal{L}(U, H_{-1/2})$, where U is another Hilbert space, we consider the differential equation

$$\ddot{w}(t) + Lw(t) + DD^*\dot{w}(t) = 0, \quad t \geq 0, \tag{2-8a}$$

$$w(0) = w_0 \in H_{1/2}, \quad \dot{w}(0) = w_1 \in H. \tag{2-8b}$$

Such systems have been studied extensively; see for instance [Lasiecka and Triggiani 2000; Guo and Luo 2002; Anantharaman and Léautaud 2014; Ammari and Nicaise 2015; Ammari and Tucsnak 2001] and the references therein. This class of systems in particular contains the wave equation with viscous damping on a two-dimensional bounded and convex domain $\Omega \subseteq \mathbb{R}^2$ with (necessarily Lipschitz) boundary $\partial\Omega$,

$$w_{tt}(\xi, t) - \Delta w(\xi, t) + b(\xi)^2 w_t(\xi, t) = 0, \quad t > 0,$$

where $b \in L^\infty(\Omega)$ is a nonnegative function and we impose Dirichlet boundary conditions. In this situation we may choose $H = U = L^2(\Omega)$, let $L = -\Delta$ be the (negative) Laplacian on H with Dirichlet boundary conditions, and define $D \in \mathcal{L}(U, H)$ by $Du = bu$ for all $u \in U$. This partial differential equation will be studied in detail in Section 6A.

In order to formulate the abstract system (2-8) as a first-order abstract Cauchy problem of the form (2-2), we proceed as in [Tucsnak and Weiss 2014, Section 6]. In particular, we let $x(\,\cdot\,) = (w(\,\cdot\,), \dot{w}(\,\cdot\,))$ and take X to be the Hilbert space $X = H_{1/2} \times H$ equipped with the inner product $\langle x, y\rangle_X = \langle x_1, y_1\rangle_{H_{1/2}} + \langle x_2, y_2\rangle_H$ for $x = (x_1, x_2),\ y = (y_1, y_2) \in X$. The operators $A : D(A) \subseteq X \to X$ and $B : U \to X_{-1}$ in Section 2A are defined as

$$A = \begin{pmatrix} 0 & I \\ -L & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 \\ D \end{pmatrix},$$

with $D(A) = H_1 \times H_{1/2}$ and $X_{-1} = H \times H_{-1/2}$. Then A is a skew-adjoint operator and thus it generates a unitary group $(T(t))_{t\in\mathbb{R}}$ on X. We may choose $V = H_{1/2} \times H_{1/2}$, which has the corresponding dual space $V^* = H_{1/2} \times H_{-1/2}$. The dual pairing of V and V^* is given by

$$\langle x, y\rangle_{V^*,V} = \langle x_1, y_1\rangle_{H_{1/2}} + \langle x_2, y_2\rangle_{H_{-1/2},H_{1/2}}$$

for $x = (x_1, x_2) \in V^*,\ y = (y_1, y_2) \in V$.

Condition (H1) is satisfied since $\operatorname{Re}\langle A_{-1}x, x\rangle_{V^*,V} = 0$ for $x \in V = V_A$, as is easily verified. In addition, we have both $B \in \mathcal{L}(U, X_{-1})$ and $B \in \mathcal{L}(U, V^*)$. For $\lambda \in \rho(A)$ the resolvent of A has the form

$$(\lambda - A)^{-1} = \begin{pmatrix} \lambda(\lambda^2+L)^{-1} & (\lambda^2+L)^{-1} \\ -L(\lambda^2+L)^{-1} & \lambda(\lambda^2+L)^{-1} \end{pmatrix},$$

and an analogous formula holds for $(\lambda - A_{-1})^{-1}$. Therefore we in particular have $\operatorname{Ran}(A_{-1}^{-1}B) \subseteq V$, and thus condition (H2) in Assumption 2.1 is satisfied. By Lemma 2.2 the operator A_B defined in (2-1) generates a contraction semigroup on X, as also shown in [Lasiecka and Triggiani 2000, Proposition 7.6.1] and [Guo and Luo 2002, Theorem 1].

It is straightforward to see that $B^* = (0, D^*) \in \mathcal{L}(V, U)$, where $D^* \in \mathcal{L}(H_{1/2}, U)$ is the adjoint of $D \in \mathcal{L}(U, H_{-1/2})$. Therefore the formula for $(\lambda - A_{-1})^{-1}$ implies that

$$B^*(\lambda - A_{-1})^{-1}B = \lambda D^*(\lambda^2 + L_{-1})^{-1}D, \quad \lambda \in \mathbb{C}_+.$$

Moreover, $\|D^*((1+is)^2 + L_{-1})^{-1}D\| = \|D^*((1-is)^2 + L_{-1})^{-1}D\|$, $s \in \mathbb{R}$. Hence if

$$s\|D^*((1+is)^2 + L_{-1})^{-1}D\| \le \mu_0(s), \quad s \in \mathbb{R}_+, \tag{2-9}$$

for some $\mu_0 : \mathbb{R}_+ \to [r_0', \infty)$, $r_0' > 0$, then condition (2-7) holds for some even function $\mu : \mathbb{R} \to [r_0, \infty)$, $r_0 > 0$, satisfying $\mu(s) \lesssim \mu_0(|s|)$, $s \in \mathbb{R}$. Conversely, property (2-7) implies the above estimate for $\mu_0 : \mathbb{R}_+ \to [r_0, \infty)$ defined by $\mu_0(s) = \mu(s)$, $s \in \mathbb{R}_+$. The estimate (2-9) has been shown to hold for a bounded function μ_0 for several PDE models having our second-order form (2-8); see for instance [Ammari and Tucsnak 2001; Guo and Luo 2002; Lasiecka and Triggiani 2003]. On the other hand, as shown in [Lasiecka and Triggiani 1981] and [Weiss 2003, Section 4], unbounded functions μ_0 are needed in some cases including wave equations with boundary damping. In the case where $D \in \mathcal{L}(U, H)$, we have $B \in \mathcal{L}(U, X)$ and, in particular, (2-7) holds for a bounded function μ by Lemma 2.5.

2C. ***Resolvent estimates and nonuniform stability.*** Throughout the paper we are interested in finding sufficient conditions for the spectrum of the operator A_B defined in (2-1) to be contained in $\mathbb{C}_-$ and in obtaining a resolvent estimate of the form

$$\|(is - A_B)^{-1}\| \le N(s), \quad s \in \mathbb{R}, \tag{2-10}$$

for an explicit function $N : \mathbb{R} \to (0, \infty)$.

In order to pass from the resolvent estimate (2-10) to sharp rates of decay for the semigroup $(T_B(t))_{t\ge0}$ we make use of the following abstract result from [Rozendaal et al. 2019, Theorem 3.2]; see [Borichev and Tomilov 2010, Theorem 2.4] for the case where N is a polynomial. Recall that a measurable function $N : \mathbb{R}_+ \to (0, \infty)$ is said to have *positive increase* if there exist constants α, $s_0 > 0$ and $c_\alpha \in (0, 1]$ such that

$$\frac{N(\lambda s)}{N(s)} \ge c_\alpha \lambda^\alpha, \quad \lambda \ge 1, \ s \ge s_0. \tag{2-11}$$

When $N : \mathbb{R}_+ \to (0, \infty)$ is nondecreasing but not necessarily strictly increasing we take N^{-1} to denote the right-continuous right-inverse of N defined by $N^{-1}(t) = \sup\{s \ge 0 : N(s) \le t\}$ for $t \ge N(0)$.

Theorem 2.7 [Rozendaal et al. 2019, Theorem 3.2]. *Let $(T(t))_{t\ge0}$ be a strongly continuous contraction semigroup on a Hilbert space X, with generator A. If $i\mathbb{R} \subseteq \rho(A)$ and if $\|(is - A)^{-1}\| \le N(|s|)$ for all $s \in \mathbb{R}$, where $N : \mathbb{R}_+ \to (0, \infty)$ is a continuous nondecreasing function of positive increase, then*

$$\|T(t)A^{-1}\| = O\left(\frac{1}{N^{-1}(t)}\right), \quad t \to \infty. \tag{2-12}$$

The class of functions satisfying (2-11) contains all regularly varying functions $N : \mathbb{R}_+ \to (0, \infty)$ which have positive index [Rozendaal et al. 2019, Section 2], and in particular it contains any measurable function $N : \mathbb{R}_+ \to (0, \infty)$ defined for all sufficiently large values of $s \geq 0$ by $N(s) = s^\alpha \log(s)^\beta$, where $\alpha > 0$ and $\beta \in \mathbb{R}$. As discussed in [Borichev and Tomilov 2010; Rozendaal et al. 2019; Debruyne and Seifert 2019], Theorem 2.7 is optimal in several senses, and for a large class of semigroups the condition of positive increase is even a *necessary* condition for (2-12) to hold.

Remark 2.8. If $N(s) = C(1+|s|)^\alpha$ in Theorem 2.7 for some constants $C, \alpha > 0$, then (2-12) becomes $\|T(t)A^{-1}\| = O(t^{1/\alpha})$ as $t \to \infty$. It is shown in [Borichev and Tomilov 2010, Theorem 2.4] that for individual orbits of $(T(t))_{t\geq 0}$ one obtains the even better decay rate $\|T(t)x\| = o(t^{-1/\alpha})$ as $t \to \infty$ for all $x \in D(A)$.

In subsequent sections we shall repeatedly make use of the following lemma when proving resolvent estimates; see, e.g., [Arendt et al. 2011, Proposition 4.3.6] for a proof of a more general result.

Lemma 2.9. *Let A be the generator of a contraction semigroup on a Hilbert space X and let $s \in \mathbb{R}$. If there exists $c_s > 0$ such that*

$$\|x\| \leq c_s \|(is - A)x\|, \quad x \in D(A), \tag{2-13}$$

then $is \in \rho(A)$ and $\|(is - A)^{-1}\| \leq c_s$.

We shall also make use of the following lemma on adjoints in the case where A is a skew-adjoint operator. Here the composition $(\lambda - A_B)^{-1}B$ in part (b) is defined as in Remark 2.3.

Lemma 2.10. *Let A and B satisfy Assumption 2.1 and assume that A is skew-adjoint.*

(a) *We have*

$$((\lambda - A_{-1})^{-1}B)^* = B^*(\bar{\lambda} + A)^{-1}, \quad \lambda \in \rho(A).$$

(b) *If* $\mathrm{Re}\langle A_{-1}x, x\rangle_{V^*,V} = 0$ *for all $x \in V_A$, then the adjoint A_B^* of A_B defined in* (2-1) *is given by*

$$A_B^* x = -A_{-1}x - BB^*x, \quad x \in D(A_B^*), \tag{2-14a}$$

$$D(A_B^*) = \{x \in V : A_{-1}x + BB^*x \in X\}. \tag{2-14b}$$

Moreover, $((\lambda - A_B)^{-1}B)^* = B^*(\bar{\lambda} - A_B^*)^{-1}$ *for $\lambda \in \rho(A_B) \cap \overline{\mathbb{C}}_+$.*

Proof. To prove part (a), let $\lambda \in \rho(A)$, $x \in X$ and $u \in U$. By density of X in X_{-1}, we may find a sequence $(y_k)_{k\in\mathbb{N}} \subseteq X$ such that $\|y_k - Bu\|_{X_{-1}} \to 0$ as $k \to \infty$. Since $(\bar{\lambda} + A_{-1})^{-1} \in \mathcal{L}(X_{-1}, X)$, we also have

$$\|(\bar{\lambda} + A_{-1})^{-1}Bu - (\bar{\lambda} + A)^{-1}y_k\|_X \to 0, \quad k \to \infty.$$

Hence the definition of B^* and skew-adjointness of A imply that

$$\begin{aligned}\langle u, B^*(\lambda - A)^{-1}x\rangle_U &= \langle Bu, (\lambda - A)^{-1}x\rangle_{V^*,V} = \langle Bu, (\lambda - A)^{-1}x\rangle_{X_{-1},X_1} \\ &= \lim_{k\to\infty} \langle y_k, (\lambda - A)^{-1}x\rangle_{X_{-1},X_1} = \lim_{k\to\infty} \langle y_k, (\lambda - A)^{-1}x\rangle_X \\ &= \lim_{k\to\infty} \langle (\bar{\lambda} + A)^{-1}y_k, x\rangle_X = \langle (\bar{\lambda} + A_{-1})^{-1}Bu, x\rangle_X.\end{aligned}$$

Since x and u were arbitrary, we have $(B^*(\lambda - A)^{-1})^* = (\bar{\lambda} + A_{-1})^{-1}B$.

To prove (b), we define

$$\tilde{A}_B x = -A_{-1}x - BB^*x, \quad x \in D(\tilde{A}_B),$$
$$D(\tilde{A}_B) = \{x \in V : A_{-1}x + BB^*x \in X\}.$$

Since $-A$ and B satisfy Assumption 2.1 (with the same choice of V), $\tilde{A}_B$ generates a contraction semigroup on X by Lemma 2.2. The assumption that $\operatorname{Re}\langle A_{-1}x, x\rangle_{V^*,V} = 0$ for $x \in V_A$ and a simple polarisation argument imply that $\langle A_{-1}x, y\rangle_{V^*,V} = -\langle x, A_{-1}y\rangle_{V,V^*}$ for $x, y \in V_A$, where we define $\langle z_1, z_2\rangle_{V,V^*} := \overline{\langle z_2, z_1\rangle_{V^*,V}}$ for $z_1 \in V$, $z_2 \in V^*$. Hence if $x \in D(A_B) \subseteq V_A$ and $y \in D(\tilde{A}_B) \subseteq V_A$, then

$$\langle A_B x, y\rangle_X = \langle A_{-1}x - BB^*x, y\rangle_{V^*,V} = \langle x, (-A_{-1} - BB^*)y\rangle_{V,V^*} = \langle x, \tilde{A}_B y\rangle_X.$$

Thus A_B^* is an extension of $\tilde{A}_B$, and since $\rho(A_B^*) \cap \rho(\tilde{A}_B) \neq \varnothing$ we further see that $A_B^* = \tilde{A}_B$.

Now let $\lambda \in \rho(A_B) \cap \overline{\mathbb{C}}_+$, $x \in X$ and $u \in U$. We have $(\bar{\lambda} - A_B^*)^{-1}x \in D(A_B^*) \subseteq V_A$. Moreover, by Remark 2.3 we have $(\lambda - A_B)^{-1}Bu \in V_A$ and

$$\begin{aligned}\langle u, B^*(\bar{\lambda} - A_B^*)^{-1}x\rangle_U &= \langle Bu, (\bar{\lambda} - A_B^*)^{-1}x\rangle_{V^*,V}\\ &= \langle(\lambda - A_{-1} + BB^*)(\lambda - A_B)^{-1}Bu, (\bar{\lambda} - A_B^*)^{-1}x\rangle_{V^*,V}\\ &= \langle(\lambda - A_B)^{-1}Bu, (\bar{\lambda} + A_{-1} + BB^*)(\bar{\lambda} - A_B^*)^{-1}x\rangle_{V,V^*}\\ &= \langle(\lambda - A_B)^{-1}Bu, x\rangle_X.\end{aligned}$$

Since $\lambda \in \rho(A_B) \cap \overline{\mathbb{C}}_+$, $x \in X$ and $u \in U$ were arbitrary, the proof is complete. □

The following proposition presents some general consequences of resolvent estimates of the form (2-10). In particular, part (c) concerns the effect of scaling the operator B on the resulting resolvent estimate. Once again, the composition $(is - A_B)^{-1}B$ for $s \in \mathbb{R}$ is defined as in Remark 2.3. As noted in Section 2B, the additional assumptions in (b) are in particular satisfied for the class of second-order systems considered there.

Lemma 2.11. *Let A and B satisfy Assumption 2.1 and let A_B be as defined in* (2-1). *If $i\mathbb{R} \subseteq \rho(A_B)$ and if $N : \mathbb{R} \to (0, \infty)$ is such that* (2-10) *holds, then the following are true*:

(a) *For $s \in \mathbb{R}$, we have*
$$\begin{aligned}\|B^*(is - A_B)^{-1}\| &\le N(s)^{1/2},\\ \|(is - A_B)^{-1}B\| &\lesssim 1 + N(s),\\ \|B^*(is - A_B)^{-1}B\| &\le 1.\end{aligned}$$

(b) *If either $B \in \mathcal{L}(U, X)$, or*
$$A^* = -A \qquad \text{and} \qquad \operatorname{Re}\langle A_{-1}x, x\rangle_{V^*,V} = 0, \quad x \in V_A,$$
then $\|(is - A_B)^{-1}B\| \le N(s)^{1/2}$ for all $s \in \mathbb{R}$.

(c) *Let $\kappa > 0$ and consider the operator $A_{B,\kappa} : D(A_{B,\kappa}) \subseteq X \to X$ defined by*
$$A_{B,\kappa}x = A_{-1}x - \kappa^2 BB^*x, \quad x \in D(A_{B,\kappa}),$$
$$D(A_{B,\kappa}) = \{x \in V : A_{-1}x - \kappa^2 BB^*x \in X\}.$$
Then $i\mathbb{R} \subseteq \rho(A_{B,\kappa})$ and $\|(is - A_{B,\kappa})^{-1}\| \lesssim 1 + N(s)^2$ for $s \in \mathbb{R}$. If the assumptions in part (b) *hold, then $\|(is - A_{B,\kappa})^{-1}\| \lesssim N(s)$ for $s \in \mathbb{R}$.*

Proof. To prove the first estimate in (a), fix $s \in \mathbb{R}$ and $y \in X$, and let $x = (is - A_B)^{-1} y \in D(A_B)$. Then $\|x\| \le N(s)\|y\|$ and $(is - A_B)x = y$, and hence, by (2-3) in Lemma 2.2,

$$\|B^*x\|^2 \le \operatorname{Re}\langle y, x\rangle \le \|y\|\,\|x\| \le N(s)\|y\|^2.$$

Since $s \in \mathbb{R}$ and $y \in X$ were arbitrary, the first estimate in part (a) follows.

To prove the second and third estimates in (a), we begin by deriving a preliminary estimate. Let $\lambda \in \overline{\mathbb{C}}_+$ and $u \in U$. If we define the composition $(\lambda - A_B)^{-1}B$ as in Remark 2.3 and let $x = (\lambda - A_B)^{-1}Bu \in X$, then Remark 2.3 implies that $x \in V$ and $A_{-1}x + B(u - B^*x) = \lambda x \in X$. Estimate (2-5) in the proof of Lemma 2.2 shows that

$$\begin{aligned}(\operatorname{Re}\lambda)\|x\|^2 &= \operatorname{Re}\langle A_{-1}x + B(u - B^*x), x\rangle_X \le \operatorname{Re}\langle B^*x, u - B^*x\rangle_U \\ &= \operatorname{Re}\langle B^*x, u\rangle_U - \|B^*x\|_U^2.\end{aligned}$$

In particular, $\|B^*(\lambda - A_B)^{-1}Bu\| = \|B^*x\| \le \|u\|$ for all $\lambda \in \overline{\mathbb{C}}_+$, which implies the third estimate in (a). On the other hand, for $\lambda = 1 + is$ with $s \in \mathbb{R}$, the same estimate shows that

$$\begin{aligned}\|(1 + is - A_B)^{-1}Bu\|^2 &\le \operatorname{Re}\langle B^*x, u\rangle_U - \|B^*x\|_U^2 \\ &\le \operatorname{Re}\langle B^*(1 + is - A_B)^{-1}Bu, u\rangle_U \le 1.\end{aligned}$$

This inequality together with the property that (see Remark 2.3)

$$(is - A_B)^{-1}Bu = (I + (is - A_B)^{-1})(1 + is - A_B)^{-1}Bu, \quad s \in \mathbb{R},$$

finally implies the second estimate in (a).

In order to prove (b), we first note that under the additional assumptions it follows either from boundedness of B or from Lemma 2.10(b) that the adjoint A_B^* is given by (2-14) and that $((is - A_B)^{-1}B)^* = B^*(-is - A_B^*)^{-1}$, $s \in \mathbb{R}$. Proceeding as in the case of the first estimate in part (a), we may use the structure of A_B^* to show that $\|B^*(-is - A_B^*)^{-1}\|^2 \le \|(-is - A_B^*)^{-1}\|$ for $s \in \mathbb{R}$. Hence for all $s \in \mathbb{R}$ we have

$$\|(is - A_B)^{-1}B\| = \|B^*(-is - A_B^*)^{-1}\| \le \|(is - A_B)^{-1}\|^{1/2} \le N(s)^{1/2}.$$

To show (c), let $\kappa > 0$ and $s \in \mathbb{R}$ be fixed. Moreover, let $x \in D(A_{B,\kappa})$ and $y = (is - A_{B,\kappa})x \in X$. Estimate (2-3) in Lemma 2.2 (applied to the operators A and κB) implies that $\|B^*x\|^2 \le \kappa^{-2}\|x\|\,\|y\|$. We have

$$y = (is - A_{-1} + \kappa^2 BB^*)x = (is - A_{-1} + BB^*)x + (\kappa^2 - 1)BB^*x,$$

and since $x \in V$ and $(is - A_{-1} + BB^*)x \in X + \operatorname{Ran}(B)$, Remark 2.3 gives

$$x = (is - A_B)^{-1}y + (1 - \kappa^2)(is - A_B)^{-1}BB^*x.$$

Using Young's inequality we obtain

$$\begin{aligned}\|x\|^2 &\le 2N(s)^2\|y\|^2 + 2(1 - \kappa^2)^2\|(is - A_B)^{-1}B\|^2\,\|B^*x\|^2 \\ &\le 2N(s)^2\|y\|^2 + 2\frac{(1 - \kappa^2)^2}{\kappa^2}\|(is - A_B)^{-1}B\|^2\,\|x\|\,\|y\| \\ &\le 2N(s)^2\|y\|^2 + \frac{1}{2}\|x\|^2 + \frac{2(1 - \kappa^2)^4}{\kappa^4}\|(is - A_B)^{-1}B\|^4\,\|y\|^2.\end{aligned}$$

Since $A_{B,\kappa}$ generates a contraction semigroup by Lemma 2.2, the claims follow from parts (a) and (b) together with Lemma 2.9. □

The estimate $\|B^*(is-A_B)^{-1}B\| \le 1$, $s\in\mathbb{R}$, in part (a) was proved in [Oostveen 2000, Lemma 2.2.6, P6] in the case where $B \in \mathcal{L}(U, X)$, and a similar result for general B in the case of second-order systems was presented in [Weiss and Tucsnak 2003, Theorem 1.3].

3. Frequency domain criteria for resolvent bounds and nonuniform stability

3A. ***Criteria for first-order problems.*** In this section we consider the semigroup $(T_B(t))_{t\ge 0}$ generated by the operator A_B defined in (2-1), and present sufficient conditions for nonuniform stability of this semigroup in terms of observability properties of the pair (B^*, A). Theorem 2.7 allows us to focus on estimating the resolvent of A_B on the imaginary axis, and shows that whenever $\|(is - A_B)^{-1}\| \le N(|s|)$, $s \in \mathbb{R}$, for some continuous nondecreasing $N : \mathbb{R}_+ \to (0, \infty)$ with positive increase, the classical solutions $x(\cdot) = T_B(\cdot)x_0$, $x_0 \in D(A_B)$, of (2-2) satisfy

$$\|T_B(t)x_0\| \le \frac{C}{N^{-1}(t)}\|A_B x_0\|, \quad t \ge t_0, \tag{3-1}$$

for some constants $C, t_0 > 0$.

Our first main result is based on the following Hautus-type condition with variable parameters. The same condition with bounded functions M and m was used in [Miller 2012] to study observability properties of the pair (B^*, A).

Definition 3.1. The pair (B^*, A) is said to satisfy the *nonuniform Hautus test* if there exist $M, m : \mathbb{R} \to [r_0, \infty)$, $r_0 > 0$, such that

$$\|x\|_X^2 \le M(s)\|(is - A)x\|_X^2 + m(s)\|B^*x\|_U^2, \quad x \in D(A),\ s \in \mathbb{R}. \tag{3-2}$$

The following theorem presents a norm bound for the resolvent of A_B on $i\mathbb{R}$ when the pair (B^*, A) satisfies the nonuniform Hautus test. General properties of the function μ in condition (3-3) were discussed in Section 2A and in Lemma 2.5.

Theorem 3.2. *Let A and B satisfy Assumption 2.1. Assume further that $M, m, \mu : \mathbb{R} \to [r_0, \infty)$, $r_0 > 0$, are such that the pair (B^*, A) satisfies the nonuniform Hautus test for the functions M and m, and*

$$\|B^*(1 + is - A_{-1})^{-1}B\| \le \mu(s), \quad s \in \mathbb{R}. \tag{3-3}$$

Then the operator A_B defined in (2-1) *satisfies $i\mathbb{R} \subseteq \rho(A_B)$ and*

$$\|(is - A_B)^{-1}\| \lesssim M(s)\mu(s) + m(s)\mu(s)^2, \quad s \in \mathbb{R}.$$

Conversely, if $N : \mathbb{R} \to (0, \infty)$ is such that $\|(is - A_B)^{-1}\| \le N(s)$ for all $s \in \mathbb{R}$, then (3-2) *holds for $M(\cdot) = 2N(\cdot)^2$ and a function m such that $m(s) \lesssim 1 + N(s)^2$ for $s \in \mathbb{R}$. If, in addition, either $B \in \mathcal{L}(U, X)$, or $A^* = -A$ and $\mathrm{Re}\langle A_{-1}x, x\rangle_{V^*,V} = 0$ for all $x \in V_A$, then one may choose $m = 2N$.*

Proof. Since A_B generates a contraction semigroup on X by Lemma 2.2, Lemma 2.9 shows that the inclusion $i\mathbb{R} \subseteq \rho(A_B)$ and the resolvent estimate will follow from a suitable lower bound for $is - A_B$, $s \in \mathbb{R}$. To this end, let $s \in \mathbb{R}$ and $x \in D(A_B)$ be fixed and let $y = (is - A_B)x$. If we let $R = (1 + is - A_{-1})^{-1}$ and define $x_1 = x + RBB^*x$, then $(is - A_{-1})x_1 = y - RBB^*x \in X$ and hence $x_1 \in D(A)$. Applying (3-2) and using the identity $B^*x_1 = (I + B^*RB)B^*x$ shows that

$$\begin{aligned}
\|x_1\|^2 &\le M(s)\|(is - A)x_1\|^2 + m(s)\|B^*x_1\|^2 \\
&\le M(s)(\|y\| + \|RB\|\|B^*x\|)^2 + m(s)(1 + \|B^*RB\|)^2\|B^*x\|^2 \\
&\lesssim M(s)\|y\|^2 + \big(M(s)\|RB\|^2 + m(s)(1 + \|B^*RB\|^2)\big)\|B^*x\|^2.
\end{aligned}$$

Since $\|B^*x\|^2 \le \operatorname{Re}\langle y, x\rangle \le \|y\|\,\|x\|$ by Lemma 2.2, we may further estimate the norm of $x = x_1 - RBB^*x$ by

$$\begin{aligned}
\|x\|^2 &\lesssim \|x_1\|^2 + \|RB\|^2\|B^*x\|^2 \\
&\lesssim M(s)\|y\|^2 + \big(M(s)\|RB\|^2 + m(s)(1 + \|B^*RB\|^2)\big)\|x\|\,\|y\| \\
&\le M(s)\|y\|^2 + \varepsilon\|x\|^2 + \frac{1}{4\varepsilon}\big(M(s)\|RB\|^2 + m(s)(1 + \|B^*RB\|^2)\big)^2\|y\|^2,
\end{aligned}$$

where $\varepsilon > 0$. We have $\|B^*RB\| \le \mu(s)$ by assumption, and Lemma 2.2 further implies that $\|RB\|^2 \le \|B^*RB\| \le \mu(s)$. Letting ε be sufficiently small we obtain

$$\begin{aligned}
\|x\|^2 &\lesssim (M(s) + M(s)^2\|RB\|^4 + m(s)^2(1 + \|B^*RB\|^2)^2)\|y\|^2 \\
&\lesssim (M(s)^2\mu(s)^2 + m(s)^2\mu(s)^4)\|y\|^2 \\
&\lesssim (M(s)\mu(s) + m(s)\mu(s)^2)^2\|(is - A_B)x\|^2.
\end{aligned}$$

Since $x \in D(A_B)$ was arbitrary, Lemma 2.9 implies that $is \in \rho(A_B)$ and $\|(is - A_B)^{-1}\| \lesssim M(s)\mu(s) + m(s)\mu(s)^2$.

To prove the other claims, assume that $\|(is - A_B)^{-1}\| \le N(s)$ and let $s \in \mathbb{R}$ and $x \in D(A)$ be arbitrary. Using the properties in Remark 2.3, the claims follow from the estimate

$$\begin{aligned}
\|x\|^2 &= \|(is - A_B)^{-1}(is - A)x + (is - A_B)^{-1}BB^*x\|^2 \\
&\le 2\|(is - A_B)^{-1}\|^2\|(is - A)x\|^2 + 2\|(is - A_B)^{-1}B\|^2\|B^*x\|^2
\end{aligned}$$

and Lemma 2.11. □

Remark 3.3. In the case where μ is a bounded function the resolvent estimate in Theorem 3.2 takes the form $\|(is - A_B)^{-1}\| \lesssim M(s) + m(s)$, $s \in \mathbb{R}$. As shown in Lemma 2.5, if A and B satisfy Assumption 2.1, then condition (3-3) is always satisfied for $\mu(s) = c(1 + s^2)$, $s \in \mathbb{R}$, with some $c > 0$. However, in the absence of a more precise bound for $\|B^*(1 + is - A_{-1})^{-1}B\|$ the proof of Theorem 3.2 can be modified to derive an alternative resolvent growth bound. Indeed, if the operator R in the proof is redefined as $R = (I - A_{-1})^{-1}$ and if x_1 is defined as before, then we have $(is - A_{-1})x_1 = y + (is - 1)RBB^*x$, and estimates analogous to those in the original proof show that $i\mathbb{R} \subseteq \rho(A_B)$ and

$$\|(is - A_B)^{-1}\| \lesssim M(s)(1 + s^2) + m(s), \quad s \in \mathbb{R}.$$

This estimate is in general sharper than what is obtained from Theorem 3.2 with a quadratic upper bound for μ. Finally, for general μ the estimates in the proof of Theorem 3.2 also establish the more precise bound

$$\|(is - A_B)^{-1}\| \lesssim M(s)^{1/2} + M(s)\|(1+is-A_{-1})^{-1}B\|^2 + m(s)\mu(s)^2$$

for $s \in \mathbb{R}$. This improves on the original estimate if $\|(1+is-A_{-1})^{-1}B\| \to 0$ as $|s| \to \infty$. The latter holds, for instance, if $B \in \mathcal{L}(U, X)$ is compact.

Recall that the pair (B^*, A) is said to be *exactly observable* if

$$\int_0^\tau \|B^*T(t)x\|^2\,dt \ge c_\tau\|x\|^2, \quad x \in D(A),$$

for some $\tau > 0$ and $c_\tau > 0$ [Tucsnak and Weiss 2009, Definition 6.1.1]. If (3-3) is satisfied for a bounded function μ, then Lemma 2.6 and [Miller 2012, Theorem 2.4] imply that the nonuniform Hautus test is satisfied for some bounded functions M and m if and only if the pair (B^*, A) is exactly observable. In this situation Theorem 3.2 and the Gearhart–Prüss theorem imply that $(T_B(t))_{t\ge0}$ is exponentially stable, as in [Slemrod 1974; Curtain and Weiss 2006].

Our next resolvent estimate for a skew-adjoint operator A is based on lower bounds for B^* restricted to so-called *wavepackets* of A. Similar conditions have previously been used to study exact observability of the pair (B^*, A), for example in [Chen et al. 1991; Ramdani et al. 2005; Miller 2012].

Definition 3.4. Let A be a self-adjoint operator on X. For $s \in \mathbb{R}$ and $\delta(s) > 0$ we define $\mathrm{WP}_{s,\delta(s)}(A)$ to be the spectral subspace of A associated with the interval $(s-\delta(s), s+\delta(s)) \subseteq \mathbb{R}$. The elements $x \in \mathrm{WP}_{s,\delta(s)}(A)$ are called $(s, \delta(s))$-*wavepackets of* A. If A is skew-adjoint, then we define $\mathrm{WP}_{s,\delta(s)}(A)$ to be $\mathrm{WP}_{s,\delta(s)}(-iA)$.

The following proposition presents a sufficient condition for nonuniform stability of $(T_B(t))_{t\ge0}$ given in terms of the action of B^* on wavepackets of A. In the case where μ is a bounded function and the pair (B^*, A) is exactly observable, it is possible by Lemma 2.6 and [Miller 2012, Corollary 2.17] to choose $\delta(s) \equiv \delta_0 > 0$ and $\gamma(s) \equiv \gamma_0 > 0$, and our result then implies exponential stability of $(T_B(t))_{t\ge0}$.

Theorem 3.5. *Let A and B satisfy Assumption 2.1 and suppose that A is skew-adjoint. Suppose further that $\mu : \mathbb{R} \to [r_0, \infty)$, $r_0 > 0$, is such that*

$$\|B^*(1+is-A_{-1})^{-1}B\| \le \mu(s), \quad s \in \mathbb{R}.$$

If there exist bounded functions $\gamma, \delta : \mathbb{R} \to (0, \infty)$ such that

$$\|B^*x\|_U \ge \gamma(s)\|x\|_X, \quad x \in \mathrm{WP}_{s,\delta(s)}(A),\ s \in \mathbb{R}, \tag{3-4}$$

then $i\mathbb{R} \subseteq \rho(A_B)$ and

$$\|(is - A_B)^{-1}\| \lesssim \frac{\mu(s)^2}{\gamma(s)^2\delta(s)^2}, \quad s \in \mathbb{R}.$$

Proof. By Lemma 2.2, A_B generates a contraction semigroup on X. Thus by Lemma 2.9 the claims will follow from suitable lower bounds for the operators $is - A_B$, $s \in \mathbb{R}$. Let $s \in \mathbb{R}$ and $x \in D(A_B)$ be fixed and let $y = (is - A_B)x$. Further let $P_0 \in \mathcal{L}(X)$ be the orthogonal projection onto $\mathrm{WP}_{s,\delta(s)}(A)$, and let $P_\infty = I - P_0$. Define

$$x_0 = P_0 x, \quad x_\infty = P_\infty x, \quad y_0 = P_0 y, \quad \text{and} \quad y_\infty = P_\infty y.$$

Since $x_0 \in \mathrm{WP}_{s,\delta(s)}(A)$ and $B^* x_0 = B^* x - B^* x_\infty$, (3-4) implies that

$$\|x\|^2 = \|x_0\|^2 + \|x_\infty\|^2 \lesssim \gamma(s)^{-2}(\|B^* x\|^2 + \|B^* x_\infty\|^2) + \|x_\infty\|^2. \tag{3-5}$$

We now estimate $\|x_\infty\|$ and $\|B^* x_\infty\|$ in turn. We begin by introducing the operator $R = (1 + is - A_{-1})^{-1}$, noting that $\|R\| \le 1$ since A generates a contraction semigroup. Applying $P_\infty R$ to both sides of the identity $y = (is - A_B)x$ we obtain

$$(is - A)R x_\infty = R y_\infty - P_\infty R B B^* x, \tag{3-6}$$

and hence

$$x_\infty = R x_\infty + R y_\infty - P_\infty R B B^* x. \tag{3-7}$$

Now since R and P_∞ commute, we have $R x_\infty \in \mathrm{Ran}(P_\infty)$, and the spectral theorem for self-adjoint operators implies that $\|R x_\infty\| \le \delta(s)^{-1} \|(is - A) R x_\infty\|$. Thus

$$\|x_\infty\| \lesssim \delta(s)^{-1} \|(is - A) R x_\infty\| + \|y\| + \|RB\| \, \|B^* x\|.$$

By (3-6) we have

$$\|(is - A) R x_\infty\| \le \|R y_\infty\| + \|P_\infty R B B^* x\| \le \|y\| + \|RB\| \, \|B^* x\|,$$

and therefore

$$\|x_\infty\| \lesssim \delta(s)^{-1} (\|y\| + \|RB\| \, \|B^* x\|). \tag{3-8}$$

In order to estimate $\|B^* x_\infty\|$ we begin by observing that, by (3-7),

$$\|B^* x_\infty\| \le \|B^* R\| \, \|x_\infty\| + \|B^* R\| \, \|y\| + \|B^*(I - P_0) R B\| \, \|B^* x\|. \tag{3-9}$$

Since A is skew-adjoint, we have $B^*(1 + is - A)^{-1} = ((1 - is + A_{-1})^{-1} B)^*$ by Lemma 2.10. Hence the resolvent identity gives

$$\|B^* R\| = \|(1 - is + A_{-1})^{-1} B\| = \|RB - 2(1 - is + A)^{-1} RB\| \le 3\|RB\|,$$

and since $\|(1 + is - A) P_0\| \lesssim 1 + \delta(s) \lesssim 1$ we see using (2-4) in Lemma 2.2 that

$$\begin{aligned} \|B^*(I - P_0) R B\| &\le \|B^* R B\| + \|B^* R (1 + is - A) P_0 R B\| \\ &\lesssim \|B^* R B\| + \|RB\|^2 \lesssim \|B^* R B\|. \end{aligned}$$

Using these estimates and (3-8), we obtain from (3-9) that

$$\begin{aligned} \|B^* x_\infty\| &\lesssim \|RB\| \, \|x_\infty\| + \|RB\| \, \|y\| + \|B^* R B\| \, \|B^* x\| \\ &\lesssim \delta(s)^{-1} \|RB\| \, \|y\| + (\delta(s)^{-1} \|RB\|^2 + \|B^* R B\|) \|B^* x\|. \end{aligned}$$

Inserting our bounds for $\|x_\infty\|$ and $\|B^*x_\infty\|$ into (3-5), and using the estimate $\|B^*x\|^2 \le \|x\|\,\|y\|$ implied by (2-3) in Lemma 2.2, we deduce after a straightforward calculation that

$$\begin{aligned}\|x\|^2 &\lesssim \gamma(s)^{-2}(\|B^*x\|^2+\|B^*x_\infty\|^2)+\|x_\infty\|^2\\ &\lesssim \delta(s)^{-2}(1+\gamma(s)^{-2}\|RB\|^2)\|y\|^2+\big(\gamma(s)^{-2}(1+\delta(s)^{-2}\|RB\|^4+\|B^*RB\|^2)+\delta(s)^{-2}\|RB\|^2\big)\|x\|\,\|y\|.\end{aligned}$$

Since $\|RB\|^2 \le \|B^*RB\| \le \mu(s)$ by Lemma 2.2 and our assumption we obtain, after dropping dominated terms, the estimate

$$\|x\|^2 \lesssim \gamma(s)^{-2}\delta(s)^{-2}\mu(s)\|y\|^2+\gamma(s)^{-2}\delta(s)^{-2}\mu(s)^2\|x\|\,\|y\|.$$

An application of Young's inequality now yields

$$\|x\|^2 \lesssim \gamma(s)^{-4}\delta(s)^{-4}\mu(s)^4\|y\|^2,$$

and the claim follows from Lemma 2.9. □

Remark 3.6. In the situation where μ is a bounded function, Theorem 3.5 can alternatively be proved by combining Theorem 3.2, Lemma 2.6 and results in [Miller 2012]. Indeed, in this case Lemma 2.6 implies that B^* is admissible and by [loc. cit., Proposition 2.16] the pair (B^*, A) satisfies the nonuniform Hautus test (3-2) for some functions M and m such that $M(s) \lesssim \gamma(s)^{-2}\delta(s)^{-2}$ and $m(s) \lesssim \gamma(s)^{-2}$ for $s \in \mathbb{R}$. The claim of Theorem 3.5 then follows from Theorem 3.2. Similarly to Remark 3.3, the end of the proof of Theorem 3.5 can be modified to establish the potentially sharper resolvent estimate

$$\|(is-A_B)^{-1}\| \lesssim \nu(s)+\nu(s)^2\|(1+is-A_{-1})^{-1}B\|^2+\frac{\mu(s)^2}{\gamma(s)^2}, \quad s\in\mathbb{R},$$

where $\nu(s)=\delta(s)^{-1}(1+\gamma(s)^{-1}\|(1+is-A_{-1})^{-1}B\|)$.

Remark 3.7. It is easy to see from the proofs of Theorems 3.2 and 3.5 that if the assumptions are satisfied only for $|s| \ge s_0$ for some $s_0 > 0$, then $i\mathbb{R}\setminus(-is_0, is_0) \subseteq \rho(A_B)$ and the resolvent estimate will hold for $|s| \ge s_0$. The same comment applies to the results in the remainder of this paper. Since the nonuniform decay rate is determined only by the resolvent norms for large values of $|s|$, this property is useful in situations where $i\mathbb{R} \subseteq \rho(A_B)$ is already known or can be shown using other methods.

3B. *Criteria for second-order problems.* In this section we focus on studying the resolvent growth for the operator A_B defined in (2-1) in the case where the operators

$$A=\begin{pmatrix}0 & I\\ -L & 0\end{pmatrix} \quad\text{and}\quad B=\begin{pmatrix}0\\ D\end{pmatrix}$$

on X and U, respectively, satisfy the assumptions in Section 2B. In particular, $L : H_1 \subseteq H \to H$ is a positive self-adjoint and boundedly invertible operator and $D \in \mathcal{L}(U, H_{-1/2})$. We shall reformulate the conditions of Theorems 3.2 and 3.5 in terms of the operators L and D. In addition, we shall present further sufficient conditions for nonuniform stability in terms of generalised observability properties of the "Schrödinger group" generated by iL.

In the proofs of our results we shall employ a change of variables which transforms A into a block-diagonal operator A_{diag}; see for instance the proof of [Miller 2012, Theorem 3.8]. Recalling that $V = H_{1/2} \times H_{1/2}$, we define a unitary operator $Q \in \mathcal{L}(V, X)$ by

$$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ iL^{1/2} & -iL^{1/2} \end{pmatrix}, \quad \text{with } Q^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} I & -iL^{-1/2} \\ I & iL^{-1/2} \end{pmatrix}. \tag{3-10}$$

We then have $A = QA_{\mathrm{diag}}Q^{-1}$, where

$$A_{\mathrm{diag}} = \begin{pmatrix} iL^{1/2} & 0 \\ 0 & -iL^{1/2} \end{pmatrix} : D(A_{\mathrm{diag}}) \subseteq V \to V,$$

with domain $D(A_{\mathrm{diag}}) = H_1 \times H_1$. The following lemma describes the wavepackets of A in terms of the wavepackets of $L^{1/2}$.

Lemma 3.8. *Let L and A be as in Section 2B and let $\delta : \mathbb{R} \to (0, \infty)$ be such that $\sup_{s\in\mathbb{R}} \delta(s) \le \|L^{-1/2}\|$. Then for every $s \in \mathbb{R}$ we have*

$$\mathrm{WP}_{s,\delta(s)}(A) = \left\{ \begin{pmatrix} w \\ i\,\mathrm{sign}(s) L^{1/2} w \end{pmatrix} : w \in \mathrm{WP}_{|s|,\delta(s)}(L^{1/2}) \right\}. \tag{3-11}$$

Proof. Let $s > 0$ be fixed. We have $\mathrm{WP}_{s,\delta(s)}(A) = \mathrm{Ran}(\chi_{I_{s,\delta(s)}}(-iA))$, where $I_{s,\delta(s)} = (s-\delta(s), s+\delta(s))$. Using the decomposition $A = QA_{\mathrm{diag}}Q^{-1}$ and the upper bound for δ we see that

$$\chi_{I_{s,\delta_s}}(-iA) = Q \begin{pmatrix} \chi_{I_{s,\delta(s)}}(L^{1/2}) & 0 \\ 0 & 0 \end{pmatrix} Q^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} \chi_{I_{s,\delta(s)}}(L^{1/2}) & 0 \\ iL^{1/2}\chi_{I_{s,\delta(s)}}(L^{1/2}) & 0 \end{pmatrix} Q^{-1}.$$

The functional calculus for the positive and boundedly invertible operator L implies that

$$\chi_{I_{s,\delta(s)}}(L^{1/2}) H_{1/2} = \mathrm{Ran}(\chi_{I_{s,\delta(s)}}(L^{1/2})),$$

and hence (3-11) follows from surjectivity of Q^{-1}. The proof in the case $s < 0$ is analogous. □

The next result is a counterpart of Theorem 3.5 for damped second-order systems. We refer to [Russell 1975, Section 3] for a related result on polynomial stability of second-order systems in the case where L has discrete spectrum and $D \in \mathcal{L}(U, H)$.

Theorem 3.9. *Let L, D, A and B be as in Section 2B and assume $\mu_0 : \mathbb{R}_+ \to [r_0, \infty)$, $r_0 > 0$, is such that*

$$s\|D^*((1+is)^2 + L_{-1})^{-1}D\| \le \mu_0(s), \quad s \in \mathbb{R}_+.$$

If there exist bounded functions $\gamma_0, \delta_0 : \mathbb{R}_+ \to (0, \infty)$ such that

$$\|D^*w\|_U \ge \gamma_0(s)\|w\|_H, \quad w \in \mathrm{WP}_{s,\delta_0(s)}(L^{1/2}),\ s \ge 0,$$

then $i\mathbb{R} \subseteq \rho(A_B)$ and

$$\|(is - A_B)^{-1}\| \lesssim \frac{\mu_0(|s|)^2}{\gamma_0(|s|)^2\delta_0(|s|)^2}, \quad s \in \mathbb{R}.$$

Proof. If we let $s_0 = \min\{\|L^{-1/2}\|, 1\}$ then $\sigma(L^{1/2}) \subseteq [s_0, \infty)$. Define $\delta : \mathbb{R} \to (0, \infty)$ by

$$\delta(s) = \frac{s_0 \delta_0(|s|)}{2 \sup_{s \geq 0} \delta_0(s)}, \quad s \in \mathbb{R}. \tag{3-12}$$

Fix $s \in \mathbb{R}$ and let $x \in \mathrm{WP}_{s,\delta(s)}(A)$ be arbitrary. Lemma 3.8 implies that $x = (w, i\,\mathrm{sign}(s) L^{1/2} w)$ for some $w \in \mathrm{WP}_{|s|,\delta(s)}(L^{1/2})$. Noting that $L^{1/2} w \in \mathrm{WP}_{|s|,\delta(s)}(L^{1/2})$, our assumptions imply that

$$\|B^* x\|_U = \|D^* L^{1/2} w\|_U \geq \gamma_0(|s|)\|L^{1/2} w\|_H = \frac{\gamma_0(|s|)}{\sqrt{2}}\|x\|_X.$$

Thus the conditions of Theorem 3.5 hold for $\delta : \mathbb{R}_+ \to (0, \infty)$ defined in (3-12) and for $\gamma : \mathbb{R}_+ \to (0, \infty)$ defined by $\gamma(s) = \gamma_0(|s|)/\sqrt{2}$ for $s \in \mathbb{R}$. Since (2-9) holds by assumption, the arguments in Section 2B show that $\|B^*(1 + is - A_{-1})^{-1} B\| \lesssim \mu_0(|s|)$, $s \in \mathbb{R}$. Thus the claims follow from Theorem 3.5. □

The recent literature contains several studies of nonuniform stability for second-order systems based on observability properties of the Schrödinger group associated with (D^*, iL) when $D \in \mathcal{L}(U, H)$ is a bounded operator. In particular, the Hautus-type condition (3-13) in the following proposition was used as a starting point for deriving resolvent estimates for A_B in [Anantharaman and Léautaud 2014, Theorem 2.3] in the case of constant parameters M_0 and m_0, and with variable parameters in [Joly and Laurent 2020, Appendix B]; see also [Laurent and Léautaud 2021]. In both cases the results were used to prove nonuniform stability of wave equations with viscous damping. The following result generalises the results on resolvent growth in [Joly and Laurent 2020, Appendix B] to operators L with possibly noncompact resolvent and operators $D \in \mathcal{L}(U, H_{-1/2})$.

Proposition 3.10. *Let L, D, A and B be as in Section 2B. Moreover, let $M_0 : \mathbb{R}_+ \to (0, \infty)$ and $m_0 : \mathbb{R}_+ \to [r_0, \infty)$, $r_0 > 0$, be such that*

$$\|w\|_H^2 \leq M_0(s)\|(s^2 - L)w\|_H^2 + m_0(s)\|D^* w\|_U^2, \quad w \in H_1,\ s \geq 0, \tag{3-13}$$

and define $\eta := \inf_{s \geq 0} M_0(s)(1+s)^2 > 0$. Then the conditions of Theorem 3.9 are satisfied for the functions $\gamma_0, \delta_0 : \mathbb{R}_+ \to (0, \infty)$ defined by

$$\delta_0(s) = \frac{\min\{\sqrt{\eta}, \frac{1}{2}\}}{\sqrt{2M_0(s)}(1+s)} \quad \text{and} \quad \gamma_0(s) = \frac{1}{\sqrt{2m_0(s)}} \tag{3-14}$$

for $s \geq 0$. If, in addition, $\mu_0 : \mathbb{R}_+ \to [r_0, \infty)$, $r_0 > 0$, is such that

$$s\|D^*((1 + is)^2 + L_{-1})^{-1} D\| \leq \mu_0(s), \quad s \in \mathbb{R}_+, \tag{3-15}$$

then $i\mathbb{R} \subseteq \rho(A_B)$ and

$$\|(is - A_B)^{-1}\| \lesssim (1 + s^2) M_0(|s|) m_0(|s|) \mu_0(|s|)^2, \quad s \in \mathbb{R}.$$

Proof. Let $s \geq 0$. The function δ_0 in (3-14) is bounded and for every $r \in (s - \delta_0(s), s + \delta_0(s))$ we have

$$|s^2 - r^2| = |s - r||s + r| \leq \frac{\min\{\sqrt{\eta}, \frac{1}{2}\}(2s + \delta_0(s))}{\sqrt{M_0(s)}(1+s)} \leq \frac{1}{\sqrt{2M_0(s)}}.$$

If $w \in \mathrm{WP}_{s,\delta_0(s)}(L^{1/2})$, this estimate and the functional calculus for L imply that $\|(s^2 - L)w\|^2 \le (2M_0(s))^{-1}\|w\|^2$. Hence (3-13) yields

$$\|D^* w\|^2 \ge \frac{1}{2m_0(s)}\|w\|^2.$$

Since $s \ge 0$ and the wavepacket w were arbitrary, the conditions of Theorem 3.9 are satisfied for the functions δ_0 and γ_0 defined by (3-14), and the remaining claims follow from Theorem 3.9. □

Our result shows in particular that if (3-13) holds for constant functions M_0 and m_0 and if (3-15) holds for a bounded function μ_0, then $\|(is - A_B)^{-1}\| \lesssim 1 + s^2$ for $s \in \mathbb{R}$. The same result was previously proved for $D \in \mathcal{L}(U, H)$ in [Anantharaman and Léautaud 2014, Theorem 2.3], and we shall discuss this result further in the context of damped waves in Section 6A below. A result closely related to Proposition 3.10 and, in particular, allowing nonconstant functions M_0 and m_0 was proved in [Joly and Laurent 2020, Proposition B.3], once again in the simpler setting where $D \in \mathcal{L}(U, H)$; see also [Laurent and Léautaud 2021]. Proposition 3.10 not only generalises and extends these earlier results, it moreover allows us to see that observability conditions of the type considered in (3-13) and in [Joly and Laurent 2020, Appendix B] serve as sufficient conditions for the wavepacket condition in Theorem 3.5. Finally, in the case where μ_0 is a bounded function, Lemma 2.6 and [Miller 2012, Proposition 2.16] show that the same conditions further imply the nonuniform Hautus test in Definition 3.1 for the associated first-order equation.

We conclude this section by presenting an equivalent characterisation for the nonuniform Hautus test of pairs (B^*, A) stemming from second-order systems.

Proposition 3.11. *Let L, D, A and B be as in Section 2B. If $M_0,\ m_0 : \mathbb{R}_+ \to [r_0, \infty),\ r_0 > 0$, are such that*

$$\|w\|_H^2 \le M_0(s)\|(s - L^{1/2})w\|_H^2 + m_0(s)\|D^* w\|_U^2 \tag{3-16}$$

for all $w \in H_{1/2}$ and $s \ge 0$, then (B^*, A) satisfies the nonuniform Hautus test for some function M such that $M(s) \lesssim M_0(|s|) + m_0(|s|)$ and for m given by $m(s) = 4m_0(|s|),\ s \in \mathbb{R}$. If, in addition, $\mu_0 : \mathbb{R}_+ \to [r_0, \infty)$, $r_0 > 0$, is such that

$$s\|D^*((1 + is)^2 + L_{-1})^{-1} D\| \le \mu_0(s), \quad s \in \mathbb{R}_+,$$

then $i\mathbb{R} \subseteq \rho(A_B)$ and

$$\|(is - A_B)^{-1}\| \lesssim M_0(|s|)\mu_0(|s|) + m_0(|s|)\mu_0(|s|)^2, \quad s \in \mathbb{R}.$$

Conversely, if (B^*, A) satisfies the nonuniform Hautus test for some $M, m : \mathbb{R} \to [r_0, \infty),\ r_0 > 0$, then (3-16) holds for M_0 and m_0 defined by $M_0(s) = M(s)$ and $m_0(s) = m(s)/2$ for $s \ge 0$.

Proof. Since $L^{1/2}$ is boundedly invertible by definition, similarly as in [Miller 2012, Theorem 3.8] the decomposition $A = QA_{\mathrm{diag}}Q^{-1}$ with Q as in (3-10) implies that (3-2) holds if and only if

$$\|y_1\|_H^2 + \|y_2\|_H^2 \le M(s)\big(\|(s - L^{1/2})y_1\|_H^2 + \|(s + L^{1/2})y_2\|_H^2\big) + \frac{m(s)}{2}\|D^*(y_1 - y_2)\|_U^2$$

for all $y_1, y_2 \in H_{1/2}$ and $s \in \mathbb{R}$. Thus if (3-2) holds, then choosing $y_2 = 0$ and $s \ge 0$ in the above inequality implies the last claim of the proposition.

To prove the first claim, let $s \geq 0$ and $y_1, y_2 \in H_{1/2}$ be arbitrary. Our assumptions imply that $L^{1/2}$ is boundedly invertible and $D^*L^{-1/2} \in \mathcal{L}(H, U)$. Thus the estimates $\|L^{1/2}(s + L^{1/2})^{-1}\| \leq 1$, $\|(s + L^{1/2})^{-1}\| \leq \|L^{-1/2}\|^{-1}$ and (3-16) imply that

$$\begin{aligned}
&\|y_1\|_H^2+\|y_2\|_H^2\\
&\quad\leq M_0(s)\|(s-L^{1/2})y_1\|_H^2+m_0(s)\|D^*y_1\|_U^2+\|y_2\|_H^2\\
&\quad\leq M_0(s)\|(s-L^{1/2})y_1\|_H^2+2m_0(s)\|D^*(y_1-y_2)\|_U^2+2m_0(s)\|D^*L^{-1/2}\|^2\|L^{1/2}y_2\|_H^2+\|y_2\|_H^2\\
&\quad\leq M_0(s)\|(s-L^{1/2})y_1\|_H^2+2m_0(s)\|D^*(y_1-y_2)\|_U^2+\big(2m_0(s)\|D^*L^{-1/2}\|^2+\|L^{-1/2}\|^{-2}\big)\|(s+L^{1/2})y_2\|_H^2.
\end{aligned}$$

Thus (3-2) holds for $s \geq 0$ with M and m as described in the claim. For $s < 0$ we get an analogous estimate by applying (3-16) to $\|y_2\|^2$ with s replaced by $|s|$, and combining the estimates shows that (3-2) holds for $s \in \mathbb{R}$ with functions $M, m : \mathbb{R} \to [r_0, \infty)$ satisfying $m(s) = 4m_0(|s|)$ and $M(s) \lesssim M_0(|s|) + m_0(|s|)$ for $s \in \mathbb{R}$. Finally, as shown in Section 2B, the fact that (2-9) holds by assumption implies $\|B^*(1+is-A_{-1})^{-1}B\| \lesssim \mu_0(|s|)$, $s \in \mathbb{R}$, and thus the remaining claims follow from Theorem 3.2. $\square$

4. Time-domain conditions for nonuniform stability

4A. ***Conditions for first-order problems.*** In this section we present sufficient conditions for polynomial stability of the semigroup $(T_B(t))_{t\geq 0}$ generated by A_B in terms of the following generalised observability concept. Related generalisations of exact observability have previously been used in [Ammari and Tucsnak 2001; Ammari and Nicaise 2015; Ammari et al. 2017] to study nonuniform stability of damped second-order systems.

Definition 4.1. Let $(T(t))_{t\geq 0}$ be a contraction semigroup on X, with generator A, and let $C \in \mathcal{L}(X_1, U)$, where X and U are Hilbert spaces. The pair (C, A) is said to be *nonuniformly observable* (*with parameters* $\beta \geq 0$ *and* $\tau > 0$) if there exists $c_\tau > 0$ such that

$$c_\tau\|(I-A)^{-\beta}x\|_X^2 \leq \int_0^\tau \|CT(t)x\|_U^2\,dt, \quad x \in D(A). \tag{4-1}$$

Note that by [Kato 1961, Corollary] the norm $\|(I-A)^{-\beta}x\|$ in (4-1) can be replaced by $\|(\lambda_0-A)^{-\beta}x\|$ for any fixed $\lambda_0 \in \rho(A)\cap\overline{\mathbb{C}}_+$ (and a possibly different $c_\tau > 0$), and in particular the choice $\lambda_0 = 0$ is possible if $0 \in \rho(A)$. By injectivity of $(I-A)^{-\beta}$, nonuniform observability also implies *approximate observability* of the pair (C, A) in the sense that if $CT(t)x = 0$ for all $t \in [0, \tau]$, then necessarily $x = 0$. The case $\beta = 0$ corresponds to exact observability of the pair (C, A).

Throughout this section we consider the setting of Section 2A in the case where B is a bounded operator. In particular, $A : D(A) \subseteq X \to X$ generates a contraction semigroup $(T(t))_{t\geq 0}$ on a Hilbert space X and $B \in \mathcal{L}(U, X)$, where U is another Hilbert space. In this situation the generator of the semigroup $(T_B(t))_{t\geq 0}$ is $A_B = A - BB^*$ with $D(A_B) = D(A)$. The following consequence of the Heinz inequality for dissipative operators due to Kato will be important for the arguments in this section. The result in particular allows us to compare fractional powers of $I - A$ and $I - A_B$.

Theorem 4.2 [Kato 1961, Corollary]. *Let A_1 and A_2 be generators of contraction semigroups on X, and suppose that $D(A_1) \subseteq D(A_2)$ and $\|A_2 x\| \lesssim \|A_1 x\|$ for all $x \in D(A_1)$. Then for every $\alpha \in [0, 1]$ we have $D((-A_1)^\alpha) \subseteq D((-A_2)^\alpha)$ and $\|(-A_2)^\alpha x\| \lesssim \|(-A_1)^\alpha x\|$ for all $x \in D((-A_1)^\alpha)$.*

We shall also require the following lemma. A similar result for second-order systems of the form in Section 2B (and a possibly unbounded operator B) was presented in [Ammari and Tucsnak 2001, Lemma 4.1].

Lemma 4.3. *Let $A : D(A) \subseteq X \to X$ be a skew-adjoint operator generating a unitary group $(T(t))_{t\geq 0}$ and let $B \in \mathcal{L}(U, X)$.*

(a) *For every $\tau > 0$ there exists $C_\tau > 0$ such that*

$$\int_0^\tau \|B^* T_B(t)x\|^2 \, dt \leq \int_0^\tau \|B^* T(t)x\|^2 \, dt \leq C_\tau \int_0^\tau \|B^* T_B(t)x\|^2 \, dt \tag{4-2}$$

for all $x \in X$. Moreover, the second inequality in (4-2) *remains valid when A is merely a generator of a contraction semigroup.*

(b) *The pair (B^*, A) is nonuniformly observable with parameters $\beta \in [0, 1]$ and $\tau > 0$ if and only if (B^*, A_B) is nonuniformly observable with the same parameters β and τ.*

Proof. We begin by the second statement in (a). Suppose therefore that $(T(t))_{t\geq 0}$ is a contraction semigroup and let $\tau > 0$ be fixed. Define $\Psi, \Psi_B \in \mathcal{L}(X, L^2(0, \tau; U))$ by $\Psi x := B^* T(\cdot)x$ and $\Psi_B x := B^* T_B(\cdot)x$ for all $x \in X$. If we define $\mathbb{F}_\tau \in \mathcal{L}(L^2(0, \tau; U))$ by

$$(\mathbb{F}_\tau u)(t) = \int_0^t B^* T(t-s) B u(s) \, ds, \quad u \in L^2(0, \tau; U),$$

then the variation of parameters formula for $(T_B(t))_{t\geq 0}$ implies that

$$(I + \mathbb{F}_\tau)\Psi_B = \Psi.$$

Hence the second inequality in (4-2) holds with $C_\tau = (1 + \|\mathbb{F}_\tau\|)^2$. To complete the proof of (a), assume that A is skew-adjoint in which case $(T(t))_{t\geq 0}$ is a unitary group. Direct computations may be used to show that $\operatorname{Re}\langle \mathbb{F}_\tau u, u\rangle \geq 0$ for all $u \in L^2(0, \tau; U)$, and therefore the operator $I + \mathbb{F}_\tau$ is boundedly invertible with $\|(I + \mathbb{F}_\tau)^{-1}\| \leq 1$. This implies the first inequality in (4-2) and thus completes the proof of (a).

To prove (b), fix $\beta \in [0, 1]$ and $\tau > 0$. Both $(A - I)^{-1}$ and $(A_B - I)^{-1}$ are bounded operators generating contraction semigroups on X. Since $\|(A - I)^{-1}x\| \lesssim \|(A_B - I)^{-1}x\| \lesssim \|(A - I)^{-1}x\|$ for all $x \in X$, Theorem 4.2 implies that $\|(I - A)^{-\beta}x\| \lesssim \|(I - A_B)^{-\beta}x\| \lesssim \|(I - A)^{-\beta}x\|$ for all $x \in X$. Now the claim follows directly from (a). ☐

As our first main result of this section we show that if $D(A^*) = D(A)$ and $B \in \mathcal{L}(U, X)$, then nonuniform observability of (B^*, A) implies polynomial stability of the semigroup $(T_B(t))_{t\geq 0}$ generated by A_B. The theorem is similar in nature to the results presented in [Ammari and Tucsnak 2001; Ammari et al. 2017] and [Ammari and Nicaise 2015, Chapter 2]. In particular, these references introduce generalised versions of exact observability of (B^*, A) for second-order equations of the form in Section 2B, and deduce nonuniform stability of the semigroup $(T_B(t))_{t\geq 0}$. If $\beta = 0$ in our result, then the pair (B^*, A) is exactly observable and we obtain exponential stability, similarly to [Slemrod 1974].

Theorem 4.4. *Let A be the generator of a contraction semigroup on X such that $D(A^*) = D(A)$, and let $B \in \mathcal{L}(U, X)$. If the pair (B^*, A) is nonuniformly observable with parameters $\beta \in [0, 1]$ and $\tau > 0$, then $i\mathbb{R} \subseteq \rho(A_B)$ and*

$$\|(is - A_B)^{-1}\| \lesssim 1 + |s|^{2\beta}, \quad s \in \mathbb{R}.$$

In particular, if $0 < \beta \leq 1$ then the semigroup $(T_B(t))_{t\geq 0}$ is polynomially stable and there exists a constant $C > 0$ such that

$$\|T_B(t)x\| \leq \frac{C}{t^{1/(2\beta)}}\|A_B x\|, \quad x \in D(A_B),\ t > 0. \tag{4-3}$$

If $\beta = 0$ then the semigroup $(T_B(t))_{t\geq 0}$ is exponentially stable.

Proof. Let $\beta \in [0, 1]$ and $\tau > 0$ be such that (4-1) holds for some $c_\tau > 0$. By Lemma 2.2 the semigroup $(T_B(t))_{t\geq 0}$ is contractive and $1 \in \rho(A_B)$. Moreover, both $(A - I)^{-1}$ and $(A_B - I)^{-1}$ are bounded operators generating contraction semigroups on X. Since $\|(A_B - I)^{-1}x\| \lesssim \|(A - I)^{-1}x\|$ for all $x \in X$, we have $\|(I - A_B)^{-\beta}x\| \lesssim \|(I - A)^{-\beta}x\|$ for all $x \in X$, by Theorem 4.2. Let $\lambda \in \mathbb{C}_+$ and $x \in D(A)$. The previous estimate, together with nonuniform observability of (B^*, A), Lemma 4.3(a) and the estimate $\operatorname{Re}\langle(\lambda - A_B)z, z\rangle \geq \|B^* z\|^2$, $z \in D(A)$, imply that

$$\|(I - A_B)^{-\beta}x\|^2 \lesssim \|(I - A)^{-\beta}x\|^2 \leq \frac{C_\tau}{c_\tau}\int_0^\tau \|B^* T_B(t)x\|^2\,dt \leq \frac{C_\tau}{c_\tau}\int_0^\tau \operatorname{Re}\langle T_B(t)(\lambda - A_B)x, T_B(t)x\rangle\,dt.$$

Since $D(I - A_B^*) = D(A) = D(I - A_B)$, Theorem 4.2 gives $D((I - A_B^*)^\beta) = D((I - A_B)^\beta)$, and in particular $(I - A_B^*)^\beta(I - A_B)^{-\beta} \in \mathcal{L}(X)$. Hence if $\lambda \in \mathbb{C}_+$ and $x \in D((-A_B)^{1+2\beta})$ are arbitrary, the above estimate and contractivity of $(T_B(t))_{t\geq 0}$ imply that

$$\begin{aligned}
\|x\|^2 &\lesssim \frac{C_\tau}{c_\tau}\int_0^\tau \operatorname{Re}\langle T_B(t)(\lambda - A_B)(I - A_B)^\beta x, T_B(t)(I - A_B)^\beta x\rangle\,dt \\
&= \frac{C_\tau}{c_\tau}\int_0^\tau \operatorname{Re}\langle (I - A_B^*)^\beta (I - A_B)^{-\beta} T_B(t)(\lambda - A_B)(I - A_B)^{2\beta}x, T_B(t)x\rangle\,dt \\
&\leq \frac{\tau C_\tau}{c_\tau}\|(I - A_B^*)^\beta(I - A_B)^{-\beta}\|\,\|(\lambda - A_B)(I - A_B)^{2\beta}x\|\,\|x\|.
\end{aligned}$$

Since $\mathbb{C}_+ \subseteq \rho(A_B)$ we in particular obtain

$$\sup_{0<\operatorname{Re}\lambda<1} \|(\lambda - A_B)^{-1}(I - A_B)^{-2\beta}\| < \infty.$$

Thus $\|(\lambda - A_B)^{-1}\| \lesssim 1 + |\lambda|^{2\beta}$ for $0 < \operatorname{Re}\lambda < 1$ by [Latushkin and Shvydkoy 2001, Lemma 3.2]. In particular, the inequality $\|(\lambda - A_B)^{-1}\| \geq 1/\operatorname{dist}(\lambda, \sigma(A_B))$ implies that $i\mathbb{R} \subseteq \rho(A_B)$ and $\|(is - A_B)^{-1}\| \lesssim 1 + |s|^{2\beta}$ for $s \in \mathbb{R}$. Finally, for $\beta \in (0, 1]$, the estimate (4-3) follows from Theorem 2.7, and for $\beta = 0$ the claim follows from the Gearhart–Prüss theorem. □

As shown in the following proposition, nonuniform observability of (B^*, A) can also be characterised in terms of the orbits of the semigroup $(T_B(t))_{t\geq 0}$.

Proposition 4.5. *Let A be skew-adjoint and $B \in \mathcal{L}(U, X)$. The pair (B^*, A) is nonuniformly observable with parameters $\beta \in [0, 1]$, $\tau > 0$ if and only if*

$$\|(I - A)^{-\beta} x\|^2 \lesssim \|x\|^2 - \|T_B(\tau)x\|^2, \quad x \in X. \tag{4-4}$$

In particular, if (4-4) *holds for some $\beta \in [0, 1]$ and $\tau > 0$, then $i\mathbb{R} \subseteq \rho(A_B)$ and $\|(is - A_B)^{-1}\| \lesssim 1 + |s|^{2\beta}$ for $s \in \mathbb{R}$.*

Proof. Fix $\beta \in [0, 1]$ and $\tau > 0$. As in the proof of Lemma 4.3, we have $\|(I - A)^{-\beta} x\| \lesssim \|(I - A_B)^{-\beta} x\| \lesssim \|(I - A)^{-\beta} x\|$ for all $x \in X$ by Theorem 4.2. For every $x \in D(A) = D(A_B)$ we have

$$\begin{aligned} 2\int_0^\tau \|B^* T_B(t)x\|^2 \, dt &= 2\int_0^\tau \operatorname{Re}\langle(-A + BB^*)T_B(t)x, T_B(t)x\rangle \, dt \\ &= -\int_0^\tau \frac{d}{dt}\|T_B(t)x\|^2 \, dt = \|x\|^2 - \|T_B(\tau)x\|^2. \end{aligned}$$

Thus (4-4) is equivalent to nonuniform observability of the pair (B^*, A_B) with parameters β and τ, which in turn is equivalent to nonuniform observability of (B^*, A) with parameters β and τ by Lemma 4.3(b). If (4-4) holds, then nonuniform observability of (B^*, A) and Theorem 4.4 imply that $i\mathbb{R} \subseteq \rho(A_B)$ and $\|(is - A_B)^{-1}\| \lesssim 1 + |s|^{2\beta}$ for $s \in \mathbb{R}$. □

Note that by Theorem 4.2 the norm $\|(I - A)^{-\beta} x\|$ on the left-hand side of (4-4) can be replaced by $\|(I - A_B)^{-\beta} x\|$, or by $\|(-A)^{-\beta} x\|$ if $0 \in \rho(A)$. Estimates similar to (4-4) have been used in the literature in order to prove polynomial decay rates for $(T_B(t))_{t \geq 0}$ based on discrete-time iterations, especially for damped wave equations [Russell 1975] and coupled partial differential equations [Rauch et al. 2005; Duyckaerts 2007]. In particular, in the special case $\beta = \frac{1}{2}$ condition (4-4) is equivalent to the observability estimate [Duyckaerts 2007, equation (39)]. Thus Theorem 4.4 improves and generalises the stability result in [loc. cit., Section 5] in the case where A is skew-adjoint. Finally, if A generates a contraction semigroup and $B \in \mathcal{L}(U, X)$, then nonuniform observability of (B^*, A) with parameters $\beta \in [0, 1]$ and $\tau > 0$ implies (4-4).

4B. ***Time-domain conditions for second-order problems.*** In this section we study nonuniform observability for second-order systems of the form

$$\ddot{w}(t) + Lw(t) + DD^*\dot{w}(t) = 0, \quad t \geq 0. \tag{4-5}$$

Throughout the section, L, D, A and B are as in Section 2B. In the proofs of our results we also make use of the operator $|A_{\text{diag}}| : D(|A_{\text{diag}}|) \subseteq X \to X$ defined by

$$|A_{\text{diag}}| = \begin{pmatrix} L^{1/2} & 0 \\ 0 & L^{1/2} \end{pmatrix}, \quad D(|A_{\text{diag}}|) = D(A). \tag{4-6}$$

For second-order systems the concept of nonuniform observability in Definition 4.1 has the following alternative characterisation.

Proposition 4.6. *Let L, D, A and B be as in Section 2B. The pair (B^*, A) is nonuniformly observable with parameter $\beta \in [0, 1]$ and $\tau > 0$ if and only if*

$$\|L^{(1-\beta)/2}w_0\|_H^2 + \|L^{-\beta/2}w_1\|_H^2 \lesssim \int_0^\tau \|D^*\dot{w}(t)\|_U^2\,dt,$$

where w is the (classical) solution of

$$\ddot{w}(t) + Lw(t) = 0, \qquad w(0) = w_0 \in H_1, \quad \dot{w}(0) = w_1 \in H_{1/2}.$$

Proof. Fix $\beta \in [0, 1]$ and $\tau > 0$. Since $0 \in \rho(A)$, the norm $\|(I - A)^{-\beta}x\|$ in (4-1) can be replaced by $\|(-A)^{-\beta}x\|$. If $|A_{\mathrm{diag}}|$ is defined as in (4-6), then for $x = (x_1, x_2) \in X = H_{1/2} \times H$ we have

$$\|-A^{-1}x\|_X^2 = \|L^{-1}x_2\|_{H_{1/2}}^2 + \|x_1\|_H^2 = \||A_{\mathrm{diag}}|^{-1}x\|_X^2.$$

Thus Theorem 4.2 implies that $\|(-A)^{-\beta}x\| \lesssim \||A_{\mathrm{diag}}|^{-\beta}x\| \lesssim \|(-A)^{-\beta}x\|$ for all $x \in X$, and hence

$$\|(-A)^{-\beta}x\|^2 \lesssim \|L^{(1-\beta)/2}x_1\|_H^2 + \|L^{-\beta/2}x_2\|_H^2 \lesssim \|(-A)^{-\beta}x\|^2$$

for all $x = (x_1, x_2) \in X$. The claims now follow from the fact that for $x = (w_0, w_1) \in D(A) = H_1 \times H_{1/2}$ we have $T(t)x \in D(A)$ and $B^*T(t)x = D^*\dot{w}(t)$ for all $t \geq 0$. □

We conclude this section by studying the damped second-order equation (4-5) for damping operators $D \in \mathcal{L}(U, H)$ satisfying

$$\|L^{-\alpha/2}w\| \lesssim \|D^*w\| \lesssim \|L^{-\alpha/2}w\|, \quad w \in H, \tag{4-7}$$

for some $\alpha \in (0, 1]$. Nonuniform stability of such equations was studied in [Liu and Zhang 2015], and in [Dell'Oro and Pata 2021] in a slightly more general setting. The assumptions on D are satisfied in particular for the damping operator $D = L^{-\alpha/2}$ in the wave and beam equations in [loc. cit., Section 15], as well as for the damped Rayleigh plate studied in [Liu and Zhang 2015, Section 3]. We shall show that such damping implies nonuniform observability in the sense of Definition 4.1. In particular, the following proposition reproduces the result of [loc. cit., Theorem 2.1] for a symmetric damping operator of the form DD^* and for $\alpha \in (0, 1]$. The degree of stability was shown in [loc. cit., Section 3] to be optimal for a class of systems with a diagonal L.

Proposition 4.7. *Let L, D, A and B be as in Section 2B with $D \in \mathcal{L}(U, H)$ such that* (4-7) *holds for some constant $\alpha \in (0, 1]$. Then the pair (B^*, A) is nonuniformly observable with parameter $\beta = \alpha$ and for any $\tau > (\pi + 2\pi^3)\|L^{-1/2}\|^{-1}$. Moreover, the semigroup $(T_B(t))_{t\geq 0}$ generated by A_B is polynomially stable and there exists a constant $C > 0$ such that*

$$\|T_B(t)x\| \leq \frac{C}{t^{1/(2\alpha)}}\|A_Bx\|, \quad x \in D(A_B),\ t > 0.$$

Proof. We begin by showing that if we define $(0, I) \in \mathcal{L}(X, H)$, then the pair $((0, I), A)$ is exactly observable for any $\tau > (\pi + 2\pi^3)\|L^{-1/2}\|^{-1}$. To prove this, let $\delta_0 = \|L^{-1/2}\|$. Then Lemma 3.8 shows that every nontrivial (s, δ_0)-wavepacket x of A has the form $x = (w, i\,\mathrm{sign}(s)L^{1/2}w)$, where w is a $(|s|, \delta_0)$-wavepacket of $L^{1/2}$, and for such x we have

$$\|(0, I)x\|_H = \|L^{1/2}w\|_H = \frac{1}{\sqrt{2}}\|x\|_X.$$

Since $\|(0, I)\| = 1$, it follows from [Miller 2012, Corollary 2.17] that the pair $((0, I), A)$ is exactly observable for $\tau > (\pi + 2\pi^3)\|L^{-1/2}\|^{-1}$.

If $|A_{\text{diag}}|$ is defined as in (4-6), then $|A_{\text{diag}}|^{-1}$ commutes with A, and thus the same is true for $|A_{\text{diag}}|^{-\alpha}$. As in the proof of Proposition 4.6 we have $\|(-A)^{-\alpha}x\| \lesssim \||A_{\text{diag}}|^{-\alpha}x\| \lesssim \|(-A)^{-\alpha}x\|$ for all $x \in X$. We may write $B^* = (0, D^*) = (0, D^*L^{\alpha/2})|A_{\text{diag}}|^{-\alpha}$, where the operator $D^*L^{\alpha/2}$ is bounded below by assumption. Thus, for any fixed $\tau > (\pi + 2\pi^3)\|L^{-1/2}\|^{-1}$ and for all $x \in D(A)$, exact observability of $((0, I), A)$ implies that

$$\int_0^\tau \|B^*T(t)x\|_U^2\, dt \gtrsim \int_0^\tau \|(0, I)T(t)|A_{\text{diag}}|^{-\alpha}x\|_H^2\, dt \gtrsim \||A_{\text{diag}}|^{-\alpha}x\|_X^2 \gtrsim \|(-A)^{-\alpha}x\|_X^2. \tag{4-8}$$

Theorem 4.2 now implies that the pair (B^*, A) is nonuniformly observable with parameter $\beta = \alpha$ and with the chosen $\tau > (\pi + 2\pi^3)\|L^{-1/2}\|^{-1}$. Since A is skew-adjoint, the remaining claims follow from Theorem 4.4. □

5. Optimality of the decay rates

In this section we investigate the optimality of our nonuniform decay estimates for the damped semigroup $(T_B(t))_{t\geq 0}$. In particular, we present lower bounds for $\|T_B(\,\cdot\,)A_B^{-1}\|$, which in turn impose a restriction on the growth of $N^{-1}(t)$ as $t \to \infty$ in estimate (3-1). Our results will allow us to show that our resolvent estimates and the resulting nonuniform decay rates are optimal or near-optimal in several situations of interest, including various PDE models to be explored in Section 6. As we shall see in Section 6A3 below, however, there are also situations of interest in which our techniques fail to produce sharp results and, in particular, the resolvent estimates obtained by means of nonuniform Hautus tests or wavepacket conditions are necessarily suboptimal.

Our first result of this section provides a lower bound for the resolvent norm $\|(is - A_B)^{-1}\|$ near eigenvalues of A. Here A is assumed to be skew-adjoint, but it need not have compact resolvent. In this section we define $B_s := (B^*P_s)^* \in \mathcal{L}(U, X)$, where $P_s := \chi_{\{s\}}(-iA)$ is the orthogonal projection onto $\operatorname{Ker}(is - A)$. Note that $\operatorname{Ran}(B_s) \subseteq \operatorname{Ker}(is - A)$ and hence we subsequently consider B_s as an operator from U into $\operatorname{Ker}(is - A)$. If $\operatorname{Ran}(B_s) = \operatorname{Ker}(is - A)$, we write $B_s^+ \in \mathcal{L}(\operatorname{Ker}(is - A), U)$ for the *Moore–Penrose pseudoinverse* of B_s. If $\dim \operatorname{Ker}(is - A) = 1$ and $B_s \neq 0$, then $\|B_s^+\| = \|B_s\|^{-1}$.

Proposition 5.1. *Let A and B satisfy Assumption 2.1 and suppose that A is skew-adjoint. Suppose, in addition, that* $i\mathbb{R} \subseteq \rho(A_B)$ *and let* $N : \mathbb{R} \to (0, \infty)$ *be a function such that* $\|(is - A_B)^{-1}\| \leq N(s)$ *for all* $s \in \mathbb{R}$. *Then* $\operatorname{Ran}(B_s) = \operatorname{Ker}(is - A)$ *for all* $s \in \mathbb{R}$, *and* $N(s) \geq \|B_s^+\|^2$ *for all* $s \in \mathbb{R}$ *such that* $is \in \sigma_p(A)$.

Proof. Fix $is \in \sigma_p(A)$ and let $y \in \operatorname{Ker}(is - A)$ be arbitrary. Then $\langle y, z\rangle_X = \langle y, P_s z\rangle_X$ for all $z \in X$. Hence if $x \in D(A_B)$ is such that $(is - A_B)x = y$, then

$$\langle y, z\rangle_X = \langle (is - A_{-1})x, P_s z\rangle_{X_{-1},X_1} + \langle BB^*x, P_s z\rangle_{X_{-1},X_1}$$

for all $z \in X$. It is straightforward to show that the first term on the right-hand side is zero, so by definition of B_s we have $\langle y, z\rangle_X = \langle B_s B^*x, z\rangle_X$ for all $z \in X$. Thus $B_s B^*x = y$. Since $y \in \operatorname{Ker}(is - A)$ was arbitrary, we deduce that $\operatorname{Ran}(B_s) = \operatorname{Ker}(is - A)$, and in particular the Moore–Penrose pseudoinverse

$B_s^+ \in \mathcal{L}(\mathrm{Ker}(is - A), U)$ of B_s is well defined. Now $\|B_s^+ y\| = \min\{\|u\| : u \in U \text{ and } B_s u = y\}$, so by the identity $B_s B^* x = y$ and Lemma 2.11 we have

$$\|B_s^+ y\|^2 \le \|B^* x\|^2 = \|B^*(is - A_B)^{-1} y\|^2 \le N(s)\|y\|^2.$$

This holds for all $y \in \mathrm{Ker}(is - A)$, so $\|B_s^+\|^2 \le N(s)$. □

Remark 5.2. If the skew-adjoint operator A in Proposition 5.1 has pure point spectrum and the eigenvalues of A are uniformly separated (but not necessarily simple), so that the spectral gap

$$d_{\mathrm{gap}} := \inf\{|s - s'| : is, is' \in \sigma(A),\ s \neq s'\}$$

is strictly positive, then the norms $\|B_s^+\|$ can be used to construct functions δ and γ for which Theorem 3.5 provides the optimal rate of resolvent growth. Indeed, if we choose a constant $\delta(s) \equiv \delta := d_{\mathrm{gap}}/4 > 0$, then all nontrivial $(s, \delta(s))$-wavepackets of A are eigenvectors corresponding to the unique eigenvalue is' in the interval $i(s - \delta, s + \delta)$. If $B_{s'}$ maps surjectively onto $\mathrm{Ker}(is' - A)$ (which is in fact necessary for is' to be an element of the resolvent set $\rho(A_B)$), then for every $x \in \mathrm{Ker}(is' - A)$ we have

$$\|B^* x\| = \|B_{s'}^* x\| \ge \|B_{s'}^+\|^{-1}\|x\|.$$

The wavepacket condition (3-4) is therefore satisfied for every bounded function γ such that $\gamma(s) \equiv \|B_{s'}^+\|^{-1}$ whenever $s \in (s' - \delta, s' + \delta)$ and $is' \in \sigma(A)$. Theorem 3.5 then implies that $\|(is - A_B)^{-1}\| \lesssim \gamma(s)^{-2}$, and by Proposition 5.1 this estimate is sharp in the sense that $N(s') \ge \gamma(s')^{-2}$ whenever $is' \in \sigma(A)$ and N is as in (3-1).

As Proposition 5.1 provides us with a lower bound for the resolvent of A_B, we proceed by showing that such a bound implies a lower bound for orbits of $(T_B(t))_{t\ge 0}$. This will be done in a more general context in anticipation of possible applications elsewhere. It was shown in [Batty and Duyckaerts 2008, Proposition 1.3] that one cannot in general hope for a better rate of decay than that given in Theorem 2.7. The following new result is a consequence of [loc. cit., Proposition 1.3]. More specifically, it is a variant of a claim made in [Batty et al. 2016, Theorem 1.1] and in the discussion following [Arendt et al. 2011, Theorem 4.4.14], and it gives a sharp optimality statement of the same type but which, crucially, is applicable as soon as one has a lower bound for the resolvent along a (possibly unknown) unbounded sequence of points on the imaginary axis. The proof uses the same ideas as that of [Batty et al. 2016, Corollary 6.11].

Proposition 5.3. *Let X be a Banach space and let $(T(t))_{t\ge 0}$ be a bounded semigroup on X whose generator A satisfies $i\mathbb{R} \subseteq \rho(A)$. Suppose that $N : \mathbb{R}_+ \to (0, \infty)$ is a continuous nondecreasing function such that $N(s) \to \infty$ as $s \to \infty$ and*

$$\limsup_{|s|\to\infty} \frac{\|(is - A)^{-1}\|}{N(|s|)} > 0. \tag{5-1}$$

Then there exists $c > 0$ such that

$$\limsup_{t\to\infty} N^{-1}(ct)\|T(t)A^{-1}\| > 0, \tag{5-2}$$

and if N has positive increase then (5-2) *holds for all $c > 0$.*

Proof. Consider the continuous nondecreasing function $n : \mathbb{R}_+ \to (0, \infty)$ defined by

$$n(t) = \sup_{\tau \geq t} \|T(\tau)A^{-1}\|, \quad t \geq 0,$$

and let n^{-1} denote any right-inverse of n. Note that n takes strictly positive values since by (5-1) the semigroup $(T(t))_{t\geq 0}$ cannot be nilpotent, and that $n(t) \to 0$ as $t \to \infty$ by [Batty and Duyckaerts 2008, Theorem 1.1]. Furthermore, by (5-1) and [loc. cit., Proposition 1.3] we may find a constant $c > 0$ and an increasing sequence $(s_k)_{k\in\mathbb{N}}$ of positive numbers such that $s_k \to \infty$ as $k \to \infty$ and $N(s_k) < cn^{-1}((2s_k)^{-1})$ for all $k \in \mathbb{N}$. Let $t_k = n^{-1}((2s_k)^{-1})$ for $k \in \mathbb{N}$. Then $t_k \to \infty$ as $k \to \infty$ because N is assumed to be unbounded, and we have $s_k = (2n(t_k))^{-1}$, $k \in \mathbb{N}$. Now $N(N^{-1}(ct_k)) = ct_k > N(s_k)$ and hence $N^{-1}(ct_k) > (2n(t_k))^{-1}$ for all $k \in \mathbb{N}$. Letting $K = \sup_{t\geq 0} \|T(t)\|$, it follows that

$$\frac{1}{2N^{-1}(ct_k)} \leq n(t_k) \leq K\|T(t_k)A^{-1}\|, \quad k \in \mathbb{N},$$

which establishes (5-2). If N has positive increase then by [Rozendaal et al. 2019, Proposition 2.2] we have $N^{-1}(t) \asymp N^{-1}(ct)$ as $t \to \infty$ for all $c > 0$, which immediately yields the second statement. □

Remark 5.4. If N is not assumed to have positive increase then it is possible for (5-1) to be satisfied but for (5-2) to hold only for certain values of $c > 0$. We refer the interested reader to the discussion following [Rozendaal et al. 2019, Remark 3.3] for an example of a contraction semigroup on a Hilbert space such that (5-1) holds for $N(s) = \log(s)$, $s \geq 2$, and $\|T(t)A^{-1}\| = O(e^{-t/2})$ as $t \to \infty$. In particular, (5-2) does not hold for any $c \in \left(0, \frac{1}{2}\right)$.

The considerations above lead to the following statement, which is the main result of this section. It is an immediate consequence of Propositions 5.1 and 5.3, both of which are applicable under more general assumptions. The result provides lower bounds for orbits of $(T_B(t))_{t\geq 0}$ under an assumption on the action of B^* on eigenvectors of A associated with imaginary eigenvalues $is_k \in \sigma_p(A)$. These lower bounds will allow us to show in Section 6B below that the nonuniform decay rates we obtain from our observability conditions are optimal (or near-optimal) in several concrete situations of interest.

Theorem 5.5. *Let A and B satisfy Assumption 2.1 and suppose that A is skew-adjoint and that $i\mathbb{R} \subseteq \rho(A_B)$. If there exist a sequence $(s_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$, $|s_k| \to \infty$ as $k \to \infty$ and a continuous nondecreasing function $N_0 : \mathbb{R}_+ \to (0, \infty)$ of positive increase such that $\|B^+_{s_k}\|^2 \geq N_0(|s_k|)$ for all $k \in \mathbb{N}$, then*

$$\limsup_{t\to\infty} N_0^{-1}(t)\|T_B(t)A_B^{-1}\| > 0.$$

Consequently, if (3-1) *holds then there exists a sequence $(t_k)_{k\in\mathbb{N}} \subseteq (0, \infty)$ with $t_k \to \infty$ as $k \to \infty$ such that $N^{-1}(t_k) \lesssim N_0^{-1}(t_k)$ for all $k \in \mathbb{N}$.*

We finish this section with a result of independent interest, offering an asymptotic estimate for a collection of eigenvalues of A_B under a uniform spectral gap condition of the type discussed in Remark 5.2.

Proposition 5.6. *Let A be skew-adjoint and suppose that $B \in \mathcal{L}(U, X)$ is compact. Suppose further that $\sigma(A) = \sigma_p(A)$ and that this set is infinite, that $\dim \operatorname{Ker}(is - A) = 1$ for every $is \in \sigma(A)$, and that $d_{\mathrm{gap}} > 0$. Then there exist a family $(\lambda_s)_{is\in\sigma_p(A)}$ and $s_0 \geq 0$ such that $\lambda_s \in \sigma(A_B)$ for $|s| \geq s_0$ and $|\lambda_s - (is - \|B_s\|^2)| = o(\|B_s\|^2)$ as $|s| \to \infty$.*

Proof. First, we note that

$$\{\lambda \in \mathbb{C}_- : \operatorname{Ker}(I + B^*(\lambda - A)^{-1}B) \neq \{0\}\} \subseteq \sigma_p(A_B).$$

Indeed, if $\lambda \in \mathbb{C}_-$ and $u \in U \setminus \{0\}$ are such that $B^*(\lambda - A)^{-1}Bu = -u$, then $(\lambda - A_B)(\lambda - A)^{-1}Bu = 0$. Since $(\lambda - A)^{-1}Bu \neq 0$ (otherwise $u = -B^*(\lambda - A)^{-1}Bu = 0$), we conclude that $\lambda \in \sigma_p(A_B)$. This reduces our problem to finding suitable points $\lambda \in \mathbb{C}_-$ with $\operatorname{Ker}(I + B^*(\lambda - A)^{-1}B) \neq \{0\}$.

Our assumptions on A and compactness of B imply that $\|B_s\| = \|P_s B\| \to 0$ as $|s| \to \infty$. Fix $is \in \sigma_p(A)$ such that $|s| \geq 9\|B\|^2$ and $\|B_s\|^2 \leq d_{\mathrm{gap}}$. By Proposition 5.1, B_s maps surjectively onto $\operatorname{Ker}(is - A)$, and therefore $B_s \neq 0$. Let

$$F_s(\lambda) = (\lambda - is)(I + B^*(\lambda - A)^{-1}B).$$

Note that for $\lambda \in \rho(A)$ we have $\operatorname{Ker}(I + B^*(\lambda - A)^{-1}B) \neq \{0\}$ if and only if $\operatorname{Ker}(F_s(\lambda)) \neq \{0\}$. Our aim is to apply Rouché's theorem for operator-valued functions [Gohberg and Sigal 1971, Theorem 2.2]. We have $F_s(\lambda) = G_s(\lambda) + H_s(\lambda)$ with

$$G_s(\lambda) = \lambda - is + B_s^* B_s, \qquad H_s(\lambda) = (\lambda - is)B^*(\lambda - A)^{-1}B - B_s^* B_s.$$

Since $B_s^* B_s$ is a rank-1 operator and $\dim X > 1$, $G_s(\lambda)$ is boundedly invertible if and only if $\lambda \notin \{is - \|B_s\|^2, is\}$. Let $r_s = \|B_s\|^2/2$ and define the closed disk $\Omega_s = \{\lambda \in \mathbb{C} : |\lambda - (is - \|B_s\|^2)| \leq r_s\} \subseteq \mathbb{C}_-$ and $\Gamma_s = \partial\Omega_s$. Then $G_s(\lambda)$ is boundedly invertible for all $\lambda \in \Omega_s \setminus \{is - \|B_s\|^2\}$, and for all $\lambda \in \Gamma_s$ we have

$$\|G_s(\lambda)^{-1}\| = \frac{1}{\operatorname{dist}(\lambda, \{is - \|B_s\|^2, is\})} = \frac{1}{r_s}.$$

Let $J_s = \{s' \in \mathbb{R} : |s' - s| \leq |s|/2\}$. For every $s' \in \mathbb{R} \setminus J_s$ and every $\lambda \in \Omega_s$ we have

$$|\lambda - is'| \geq |is' - is| - |\lambda - is| \geq \frac{|s|}{2} - \frac{3}{2}\|B_s\|^2 \geq \frac{|s|}{3},$$

where the last inequality follows from the condition $|s| \geq 9\|B\|^2$. Hence, for every $\lambda \in \Omega_s$,

$$\|B^*(\lambda - A)^{-1}\chi_{\mathbb{R}\setminus J_s}(-iA)B\| \leq \|B^*\| \sup_{|s'-s|>|s|/2} \frac{1}{|\lambda - is'|}\|B\| \leq \frac{3\|B\|^2}{|s|}.$$

Thus, for every $u \in U$ with $\|u\| \leq 1$, by the Cauchy–Schwarz inequality, the uniform spectral gap assumption and Bessel's identity, we see that

$$\begin{aligned}
\frac{\|H_s(\lambda)u\|}{|\lambda - is|} &\leq \|B^*(\lambda - A)^{-1}\chi_{\mathbb{R}\setminus J_s}(-iA)Bu\| + \left\| B^*(\lambda - A)^{-1}\chi_{J_s}(-iA)Bu - \frac{B_s^* B_s u}{\lambda - is} \right\| \\
&\leq \frac{3\|B\|^2}{|s|} + \left\| \sum_{is' \in (\sigma_p(A) \cap iJ_s)\setminus\{is\}} \frac{1}{\lambda - is'} B_{s'}^* B_{s'} u \right\| \\
&\leq \frac{3\|B\|^2}{|s|} + \sup_{|s'|\geq|s|/2} \|B_{s'}^*\| \left(2\sum_{j=1}^{\infty} \frac{1}{d_{\mathrm{gap}}^2 j^2} \right)^{1/2} \left(\sum_{is' \in \sigma_p(A)} \|B_{s'}u\|^2 \right)^{1/2} \\
&\leq \frac{3\|B\|^2}{|s|} + \frac{\pi\|B\|}{\sqrt{3}d_{\mathrm{gap}}} \sup_{|s'|\geq|s|/2} \|B_{s'}\|.
\end{aligned}$$

Thus $\|H_s(\lambda)\| \le q_s|\lambda - is|$ for some $q_s \ge 0$ satisfying $q_s \to 0$ as $|s| \to \infty$. Then, for $|s|$ large enough and $\lambda \in \Gamma_s$,

$$\|G_s(\lambda)^{-1}H_s(\lambda)\| \le \frac{q_s|\lambda - is|}{r_s} \le 3q_s < 1.$$

Rouché's theorem [Gohberg and Sigal 1971, Theorem 2.2] now implies that for every $is \in \sigma_p(A)$ with $|s|$ sufficiently large there exists $\lambda_s \in \Omega_s$ such that $\mathrm{Ker}(F(\lambda_s)) \neq \{0\}$, and the proof is complete. □

Observe that if A and B are as in Proposition 5.6 and if $i\mathbb{R} \subseteq \rho(A_B)$, then the result implies that

$$\liminf_{|s|\to\infty} \|B_s\|^2 \|(is - A_B)^{-1}\| > 0.$$

Then using Proposition 5.3 as in Theorem 5.5, we obtain a lower bound for $\|T_B(\,\cdot\,)A_B^{-1}\|$ along a sequence $(t_k)_{k\in\mathbb{N}} \subseteq (0,\infty)$ with $t_k \to \infty$ as $k \to \infty$. We omit a precise formulation of the corresponding statement since it is completely analogous to Theorem 5.5.

6. Nonuniform stability of damped partial differential equations

In this section we apply our general results to several concrete partial differential equations of different types. In particular, we consider damped wave equations on one- and two-dimensional spatial domains, a one-dimensional fractional Klein–Gordon equation, and a damped Euler–Bernoulli beam equation. We also refer to a recent article [Su et al. 2020] for an application of Theorem 3.5 in the study of a coupled PDE system describing the dynamics of linearised water waves.

6A. *Wave equations on two-dimensional domains.* In this section we consider wave equations on bounded simply connected domains $\Omega \subseteq \mathbb{R}^2$ which are either convex or have sufficiently regular (say C^2) boundary to ensure that the domain of the Dirichlet Laplacian on Ω is included in $H^2(\Omega)$. The wave equation with viscous damping and Dirichlet boundary conditions is given by

$$w_{tt}(\xi,t) - \Delta w(\xi,t) + b(\xi)^2 w_t(\xi,t) = 0, \quad \xi \in \Omega,\ t > 0, \tag{6-1a}$$

$$w(\xi,t) = 0, \quad \xi \in \partial\Omega,\ t > 0, \tag{6-1b}$$

$$w(\,\cdot\,,0) = w_0(\,\cdot\,) \in H^2(\Omega) \cap H_0^1(\Omega), \quad w_t(\,\cdot\,,0) = w_1(\,\cdot\,) \in H_0^1(\Omega). \tag{6-1c}$$

Here $b \in L^\infty(\Omega)$ is the nonnegative damping coefficient. It is well known that the geometry of Ω and the region where $b(\,\cdot\,) > 0$ have great impact on the asymptotic properties of the wave equation. In the framework of Section 2B we set $H = L^2(\Omega)$, $L = -\Delta$ with domain $H_1 = H^2(\Omega) \cap H_0^1(\Omega)$, and define $U = L^2(\Omega)$ and $D \in \mathcal{L}(L^2(\Omega))$ by $Du = bu$ for all $u \in L^2(\Omega)$. Since $D \in \mathcal{L}(U, H)$, the function μ_0 in Section 3B can be chosen to be bounded.

6A1. *Exact observability of the Schrödinger group.* In order to apply Proposition 3.10 to the damped wave equation (6-1) we need to understand the observability properties of the Schrödinger group on Ω. Of particular interest here is the case of exact observability of the Schrödinger group, which corresponds to (3-13) being satisfied for constant functions M_0 and m_0. In such cases Proposition 3.10 immediately yields the resolvent bound $\|(is - A_B)^{-1}\| \lesssim 1 + s^2$, $s \in \mathbb{R}$, so by Theorem 2.7 (and Remark 2.8) classical solutions of the corresponding abstract Cauchy problem decay like (and in fact faster than) $t^{-1/2}$ as $t \to \infty$. This was first proved in [Anantharaman and Léautaud 2014], but we mention that, similarly to [Joly and Laurent 2020, Appendix B], Proposition 3.10 also allows us to deal with the much more general situation where (3-13) is satisfied for functions M_0 and m_0 which satisfy suitable lower bounds but need not be constant. We take advantage of this added generality in Section 6A2 below.

The study of energy decay of damped waves via observability conditions has a long history [Slemrod 1974; Russell 1975; Benchimol 1977; Lebeau 1996; Ammari and Tucsnak 2001; Burq and Hitrik 2007; Cavalcanti et al. 2019; Letrouit and Sun 2023; Laurent and Léautaud 2021], and in particular it predates the resolvent approach. It is not surprising, therefore, that there is a rich literature on exact observability of the Schrödinger group, giving many concrete examples to which our abstract theory may be applied. For instance, if Ω is a rectangle then it follows from a classical result due to [Jaffard 1990] that the Schrödinger group corresponding to our system is exactly observable for every nonnegative $b \in L^\infty(\Omega)$ such that $\operatorname{ess\,sup}_{\xi\in\omega} b(\xi) > 0$ for some nonempty open set $\omega \subseteq \Omega$; see [Burq and Zworski 2019] for an even stronger result on the torus. Similarly, it follows from [Burq and Zworski 2004, Theorem 9] that if Ω is the Bunimovich stadium then the corresponding Schrödinger group is exactly observable provided the damping b has strictly positive essential infimum on a neighbourhood of one of the sides of the rectangle meeting a half-disk and also at one point on the opposite side. This allows us to recover under a slightly weaker assumption the decay rate obtained in [Burq and Hitrik 2007, Theorem 1.1]. Finally, if Ω is a disk then by [Anantharaman et al. 2016, Theorem 1.2] the Schrödinger group is exactly observable whenever $\operatorname{ess\,sup}_{\xi\in\omega} b(\xi) > 0$ for some open subset ω of $\overline{\Omega}$ such that $\omega \cap \partial\Omega \neq \varnothing$. In fact, this condition is also necessary for exact observability, as can be seen by considering so-called *whispering gallery modes*. We thus recover the decay rate for classical solutions obtained in [Anantharaman et al. 2016, Remark 1.7]. Further examples of when the Schrödinger group is exactly observable, including also higher-dimensional situations, may be found in [Anantharaman and Léautaud 2014, Section 2A]. We point out in passing that there is also scope to apply directly the wavepacket result Theorem 3.9, which underlies Proposition 3.10. One case in which this is possible is if one knows that $\operatorname{ess\,sup}_{\xi\in\omega} b(\xi) > 0$ for some open set $\omega \subseteq \Omega$ such that $\|w\|_{L^2(\omega)} \geq c\|w\|_{L^2(\Omega)}$ for some constant $c > 0$ and all eigenfunctions w of the Dirichlet Laplacian on Ω. This would allow us to take γ_0 to be constant in Theorem 3.9, provided we know how to choose δ_0 in such a way that the $(s, \delta_0(s))$-wavepackets of $(-\Delta)^{1/2}$ are eigenfunctions associated with a single eigenvalue of Δ. The appropriate lower bound is obtained in [Hassell et al. 2009] in the case where Ω is a polygonal region and ω contains a neighbourhood of each of the vertices of Ω, and in fact these assumptions can be relaxed somewhat; see [Hassell et al. 2009, Remark 4]. Choosing an appropriate δ_0, however, requires detailed information on the distribution of the eigenvalues of the Dirichlet Laplacian on Ω, which imposes a rather severe restriction on the domains Ω for which this approach is likely to bear fruit.

6A2. *Large damping away from a submanifold.* We consider the damped Klein–Gordon equation on the square $\Omega = (0, 1)^2$. This is a slight variant of (6-1) in which Δ is replaced by $\Delta - m$ for some $m > 0$. Furthermore, we view Ω as the 2-torus $\mathbb{T}^2$ by imposing periodic rather than Dirichlet boundary conditions, thus allowing us to use the results of [Burq and Zuily 2016]. We apply our abstract results, setting

$H = L^2(\mathbb{T}^2)$ and $L = -\Delta + m$ with domain $H_1 = H^2(\mathbb{T}^2)$ in the framework of Section 2B, in order to derive resolvent estimates under the assumption that the damping coefficient b satisfies a certain type of lower bound away from a proper submanifold Σ of $\mathbb{T}^2$. A typical example would be for Σ to be a circle of the form $\Sigma = \{(\xi_1, \xi_2) \in \Omega : \xi_1 \in (0, 1)\}$ for some fixed $\xi_2 \in (0, 1)$, but the results in [Burq and Zuily 2016] also apply in a much more general setting than this. The following result is a simple extension of [loc. cit., Corollary 1.3] in our special case. The distance referred to here is the geodesic distance on the manifold $\mathbb{T}^2$.

Corollary 6.1. *Let $r : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing function satisfying $r(s) > 0$ for all $s > 0$, and suppose that $b(\xi)^2 \geq r(\mathrm{dist}(\xi, \Sigma))$ for all $\xi \in \mathbb{T}^2$. Then $i\mathbb{R} \subseteq \rho(A_B)$ and there exist $\varepsilon \in (0, 1)$ and $s_0 > 0$ such that*

$$\|(is - A_B)^{-1}\| \lesssim r(\varepsilon|s|^{-1/2})^{-1}, \quad |s| \geq s_0.$$

Proof. The inclusion $i\mathbb{R} \subseteq \rho(A_B)$ may be obtained for instance by following the argument used in the proof of [Anantharaman and Léautaud 2014, Lemma 4.2]. Note in particular that the origin is removed from the spectrum as a result of the shift we apply to the Laplacian. We now prove the resolvent estimate. Given $\varepsilon \in (0, 1)$ and $s \in \mathbb{R} \setminus \{0\}$ let $\omega_{\varepsilon,s} = \{\xi \in \mathbb{T}^2 : \mathrm{dist}(\xi, \Sigma) < \varepsilon|s|^{-1/2}\}$. By [Burq and Zuily 2016, Theorem 1.1] (but see also [Sogge 1988]) there exists $s_0 > m$ such that

$$\|w\|_{L^2(\omega_{\varepsilon,s})} \lesssim \varepsilon^{1/2}(|s|^{-1}\|(s^2 - L)w\|_{L^2(\mathbb{T}^2)} + \|w\|_{L^2(\mathbb{T}^2)}) \tag{6-2}$$

for all $w \in H^2(\mathbb{T}^2)$, $\varepsilon \in (0, 1)$ and $s \in \mathbb{R}$ with $|s| \geq s_0$. By assumption we have $b(\xi)^2 \geq r(\varepsilon|s|^{-1/2})$ for all $\xi \in \mathbb{T}^2 \setminus \omega_{\varepsilon,s}$. Thus if we let $m_\varepsilon(s) = r(\varepsilon|s|^{-1/2})^{-1}$ for $\varepsilon \in (0, 1)$ and $|s| \geq s_0$, then

$$m_\varepsilon(s)\|bw\|^2_{L^2(\mathbb{T}^2)} \geq m_\varepsilon(s)\|bw\|^2_{L^2(\mathbb{T}^2 \setminus \omega_{\varepsilon,s})} \geq \|w\|^2_{L^2(\mathbb{T}^2)} - \|w\|^2_{L^2(\omega_{\varepsilon,s})},$$

and hence by (6-2) and an application of Young's inequality we may choose $\varepsilon \in (0, 1)$ sufficiently small to ensure that

$$\|w\|^2_{L^2(\mathbb{T}^2)} \lesssim |s|^{-2}\|(s^2 - L)w\|^2_{L^2(\mathbb{T}^2)} + m_\varepsilon(s)\|bw\|^2_{L^2(\mathbb{T}^2)}$$

for all $w \in H^2(\mathbb{T}^2)$ and all $s \in \mathbb{R}$ such that $|s| \geq s_0$. The result now follows from Proposition 3.10 and Remark 3.7. □

We may use Corollary 6.1 to study the asymptotic behaviour of solutions of the damped Klein–Gordon equation. In particular, if $r(s) = cs^{2\kappa}$ for some constants $c, \kappa > 0$ then Corollary 6.1 yields the estimate $\|(is - A_B)^{-1}\| \lesssim 1 + |s|^\kappa$ for $s \in \mathbb{R}$, and hence by Theorem 2.7 any classical solution decays at the rate $t^{-1/\kappa}$. Note that this is worse than the rate obtained under additional assumptions in [Léautaud and Lerner 2017; Datchev and Kleinhenz 2020] for the classical damped wave equation (6-1), which formally corresponds to the choice $m = 0$ in our setting. On the other hand, it is stated in [Burq and Zuily 2016, Remark 1.5] that in general the rate $t^{-1/\kappa}$ cannot be improved. The main value of Corollary 6.1 lies in the fact that it leads to interesting nonpolynomial resolvent estimates whenever the function r providing the lower bound is chosen appropriately.

6A3. *Suboptimality of the observability and wavepacket conditions.* In this section we discuss certain natural limitations of our results in Section 3, and in particular describe situations where the nonuniform decay rates obtained by our methods are suboptimal. As shown in [Burq and Hitrik 2007; Anantharaman and Léautaud 2014; Léautaud and Lerner 2017; Datchev and Kleinhenz 2020; Sun 2023] in the case of multidimensional wave equations with viscous damping, rates of nonuniform decay are dependent not only on the location of the damping but also on the smoothness of the damping coefficient b. By studying the damped wave equation (6-1) on a square $\Omega = (0,1)^2$ we can illustrate that the resolvent growth rates in Sections 3 and 4 are inherently suboptimal due to the fact that our observability concepts — the nonuniform Hautus test, the wavepacket condition, the observability of the Schrödinger group and the nonuniform observability — are unable to detect the degree of smoothness of the damping coefficient b.

For this purpose, let $\omega = \big(0, \frac{1}{2}\big) \times (0,1)$. For any arbitrarily small $\varepsilon \in \big(0, \frac{1}{2}\big)$ we may as in [Burq and Hitrik 2007, Section 3] define a smooth nonnegative damping coefficient b_ε such that $\operatorname{supp} b_\varepsilon \subseteq \omega$, $\|b_\varepsilon\|_{L^\infty} \le 1$, and $\|(is - A_{B_\varepsilon})^{-1}\| \lesssim 1 + |s|^{1+\varepsilon}$, $s \in \mathbb{R}$, where $B_\varepsilon \in \mathcal{L}(L^2(\Omega), X)$ is the damping operator associated with b_ε. Now consider the damping coefficient $b_\chi = \chi_\omega$, and denote the damping operator associated with this function by $B_\chi \in \mathcal{L}(L^2(\Omega), X)$. For this damping coefficient the optimal order of resolvent growth is known to be $1 + |s|^{3/2}$ [Stahn 2017; Anantharaman and Léautaud 2014], and in particular $\limsup_{|s|\to\infty} |s|^{-3/2} \|(is - A_{B_\chi})^{-1}\| > 0$. However, since $b_\chi(\xi) \ge b_\varepsilon(\xi)$ for all $\xi \in \Omega$, we clearly have

$$\|B_\chi^* x\| \ge \|B_\varepsilon^* x\|, \quad x \in X.$$

Hence the nonuniform Hautus test (3-2), the wavepacket condition (3-4), observability of the Schrödinger group (3-13), or nonuniform observability (4-1) for the pair (B_ε^*, A) immediately implies the same property for the pair (B_χ^*, A) with the same parameters. In particular, any resolvent estimate of the form $\|(is - A_{B_\varepsilon})^{-1}\| \le N(s)$, $s \in \mathbb{R}$, obtained from Theorem 3.2, Theorem 3.5, Proposition 3.10 or Theorem 4.4 also implies that $\|(is - A_{B_\chi})^{-1}\| \le N(s)$ for $s \in \mathbb{R}$. However, by [Anantharaman and Léautaud 2014, Proposition B.1] we then also have $\limsup_{|s|\to\infty} |s|^{3/2} N(s) > 0$. This means that $N(s)$ is a suboptimal upper bound for $\|(is - A_{B_\varepsilon})^{-1}\|$ as $|s| \to \infty$.

Comparing the rates of nonuniform decay of (6-1) with the two damping profiles b_ε and b_χ also shows that in the second part of Theorem 3.2 it is in general impossible to choose functions M and m satisfying $M + m \lesssim N$. To see this, let M_ε and m_ε be functions M and m corresponding to the damping b_ε. Then the inequality $b_\chi \ge b_\varepsilon$ implies that (B_χ^*, A), too, satisfies the Hautus test for the same functions M_ε and m_ε, and by Theorem 3.2 we have $\|(is - A_{B_\chi})^{-1}\| \lesssim M_\varepsilon(s) + m_\varepsilon(s)$, $s \in \mathbb{R}$. However, since the optimal order of resolvent growth for the damping b_χ is $|s|^{3/2}$, the conclusion cannot be true unless

$$\limsup_{|s|\to\infty} |s|^{3/2} (M_\varepsilon(s) + m_\varepsilon(s)) > 0.$$

Thus $M_\varepsilon + m_\varepsilon$ provides a strictly worse resolvent bound than the estimate $\|(is - A_{B_\varepsilon})^{-1}\| \lesssim 1 + |s|^{1+\varepsilon}$, $s \in \mathbb{R}$, obtained in [Burq and Hitrik 2007, Section 3].

Finally, comparison of the damping coefficients b_ε and b_χ further shows that a dissipative perturbation of a generator of a polynomially stable semigroup can strictly worsen the rate of decay. Indeed, since $b_\chi \ge b_\varepsilon$ by construction, the "additional damping" of the difference $b_\Delta = b_\chi - b_\varepsilon \ge 0$ increases the asymptotic rate of resolvent growth as $|s| \to \infty$ from at most $|s|^{1+\varepsilon}$ to $|s|^{3/2}$. In terms of the semigroup generators this means that A_{B_ε} has a strictly slower asymptotic resolvent growth than A_{B_χ} even though A_{B_χ} is a dissipative perturbation of A_{B_ε}.

6B. *Damped wave equations on one-dimensional domains.*

6B1. *Damping at a single interior point.* In this section we consider the one-dimensional wave equation with pointwise damping studied in [Ammari and Tucsnak 2001, Section 5.1]; see also [Rzepnicki and Schnaubelt 2018] for a closely related problem on the stability of two serially connected strings. Our arguments rely essentially on ideas from [Ammari and Tucsnak 2001]. Given an irrational number $\xi_0 \in (0, 1)$, let us consider the problem

$$w_{tt}(\xi, t) - w_{\xi\xi}(\xi, t) + w_t(t, \xi_0)\delta_{\xi_0}(\xi) = 0, \quad \xi \in (0, 1),\ t > 0, \tag{6-3a}$$

$$w(0, t) = 0, \quad w(1, t) = 0,\ t > 0, \tag{6-3b}$$

$$w(\,\cdot\,, 0) = w_0(\,\cdot\,) \in H^2(0, 1) \cap H_0^1(0, 1), \quad w_t(\,\cdot\,, 0) = w_1(\,\cdot\,) \in H_0^1(0, 1). \tag{6-3c}$$

As shown in [Ammari and Tucsnak 2001, Section 5.1], the system (6-3) satisfies the assumptions in Section 2B with $H = L^2(0, 1)$, $L = -\partial_{\xi\xi}$ with domain $H_1 = H^2(0, 1) \cap H_0^1(0, 1)$, and L has positive square root with domain $H_{1/2} = H_0^1(0, 1)$. The damping operator D is given by $Du = \delta_{\xi_0} u$ for all $u \in U = \mathbb{C}$, where δ_{ξ_0} is the Dirac delta distribution at $\xi = \xi_0$, and we indeed have $D \in \mathcal{L}(\mathbb{C}, H_{-1/2})$ and $D^* \in \mathcal{L}(H_{1/2}, \mathbb{C})$, where $H_{-1/2} = H^{-1}(0, 1)$ and $H_{1/2} = H_0^1(0, 1)$. In order to describe the domain $D(A_B)$, note that $A_{-1}^{-1} B = (-L^{-1}\delta_{\xi_0}, 0) = (z, 0)$, where $z \in H_0^1(0, 1)$ is the solution of the differential equation $z'' = \delta_{\xi_0}$ with boundary conditions $z(0) = z(1) = 0$ in $H^{-1}(0, 1)$. We thus have

$$z(\xi) = \begin{cases} \xi(1 - \xi_0), & 0 < \xi \le \xi_0, \\ \xi_0(1 - \xi), & \xi_0 < \xi \le 1. \end{cases}$$

Since $D(A_B) = \{x \in X_B : A_{-1}x - BB^*x \in X\}$ by Remark 2.4, we deduce that (cf. [Ammari and Tucsnak 2001, Section 5.1])

$$D(A_B) = \{(u + z(\,\cdot\,)v(\xi_0), v) : u \in H^2(0, 1) \cap H_0^1(0, 1),\ v \in H_0^1(0, 1)\},$$

and therefore classical solutions of (6-3) correspond to initial conditions

$$w_0 = w_{00} + z(\,\cdot\,)w_1(\xi_0), \quad w_{00} \in H^2(0, 1) \cap H_0^1(0, 1), \quad w_1 \in H_0^1(0, 1). \tag{6-4}$$

Since the eigenvalues $\lambda_n^2 = n^2\pi^2$, $n \in \mathbb{N}$, and corresponding normalised eigenfunctions $\phi_n(\,\cdot\,) = \sqrt{2}\sin(n\pi\,\cdot\,)$ of L are known explicitly, we may use the wavepacket condition in Theorem 3.9 to analyse the stability properties of the damped system (6-3). Indeed, the eigenvalues $\lambda_n = n\pi$, $n \in \mathbb{N}$, of $L^{1/2}$ have a uniform gap, so we may choose $\delta(s) \equiv \pi/4$. The nontrivial $(s, \delta(s))$-wavepackets of $L^{1/2}$ are then simply multiples of the eigenfunctions ϕ_n for $n \in \mathbb{N}$ such that $n\pi \in (s - \pi/4, s + \pi/4)$. For any $n \in \mathbb{N}$ we have

$$|D^*\phi_n| = |\phi_n(\xi_0)| = \sqrt{2}|\sin(n\pi\xi_0)|.$$

In order to determine the rate of resolvent growth we need to estimate the coefficients $|D^*\phi_n|$ from below. This certainly requires ξ_0 to be an irrational number, but in fact we shall need to assume more, namely that ξ_0 is *badly approximable* by rationals. It is known, for instance, that given any $\varepsilon > 0$ almost every irrational $\xi_0 \in (0, 1)$ has the property that

$$\min_{m\in\mathbb{N}} \left|\xi_0 - \frac{m}{n}\right| \geq \frac{1}{n^2 \log(n)^{1+\varepsilon}} \tag{6-5}$$

for all sufficiently large $n \geq 2$, while simultaneously for almost every irrational $\xi_0 \in (0, 1)$ there exist rationals m/n with arbitrarily large values of $n \geq 2$ such that

$$\left|\xi_0 - \frac{m}{n}\right| \leq \frac{1}{n^2 \log(n)}; \tag{6-6}$$

see for instance [Khinchin 1964, Theorem 32]. A rather special class of irrationals $\xi_0 \in (0, 1)$ is the set of irrationals that have *constant type*. These are commonly defined to be those irrational numbers which have uniformly bounded coefficients in their partial fractions expansions. Irrationals of constant type include all irrational *quadratic numbers*, that is to say irrational solutions of quadratic equations with integer coefficients. As shown in [Lang 1966, Chapter II, Theorem 6], an irrational number $\xi_0 \in (0, 1)$ has constant type if and only if there is a constant $c_{\xi_0} > 0$ such that

$$\min_{m\in\mathbb{N}} \left|\xi_0 - \frac{m}{n}\right| \geq \frac{c_{\xi_0}}{n^2}, \quad n \in \mathbb{N}. \tag{6-7}$$

It follows from the Dirichlet approximation theorem [Lang 1966, Chapter II, Theorem 1] that for any irrational number $\xi_0 \in (0, 1)$ there exist rationals m/n with arbitrarily large values of $n \in \mathbb{N}$ such that

$$\left|\xi_0 - \frac{m}{n}\right| \leq \frac{1}{n^2}. \tag{6-8}$$

The following result yields (essentially) sharp rates of decay for the energy of our damped system for irrational numbers $\xi_0 \in (0, 1)$ of different nature.

Corollary 6.2. *Let w be the* (*classical*) *solution of* (6-3) *corresponding to initial conditions as in* (6-4).

(a) *Fix $\varepsilon > 0$. For almost every irrational number $\xi_0 \in (0, 1)$ there exists $C_\varepsilon > 0$ such that*

$$\|(w(\cdot\,, t), w_t(\cdot\,, t))\|_{H^1\times L^2} \leq C_\varepsilon \frac{\log(t)^{1+\varepsilon}}{t^{1/2}} \|(w_{00}, w_1)\|_{H^2\times H^1}, \quad t \geq 2. \tag{6-9}$$

Moreover, the rate is almost optimal in the sense that if $r : \mathbb{R}_+ \to (0, \infty)$ is any function such that $r(t) = o(t^{-1/2}\log(t))$ as $t \to \infty$, then there exist w_0, w_1 as in (6-4) *for which $r(t)^{-1}\|(w(\cdot\,, t), w_t(\cdot\,, t))\|_{H^1\times L^2}$ is unbounded as $t \to \infty$.*

(b) *If $\xi_0 \in (0, 1)$ is an irrational number of constant type then there exists $C > 0$ such that*

$$\|(w(\cdot\,, t), w_t(\cdot\,, t))\|_{H^1\times L^2} \leq \frac{C}{t^{1/2}} \|(w_{00}, w_1)\|_{H^2\times H^1}, \quad t \geq 1.$$

Moreover, the rate is optimal in the sense that if $r : \mathbb{R}_+ \to (0, \infty)$ *is any function such that* $r(t) = o(t^{-1/2})$ *as* $t \to \infty$, *then there exist* w_0, w_1 *as in* (6-4) *for which* $r(t)^{-1}\|(w(\cdot\,, t), w_t(\cdot\,, t))\|_{H^1\times L^2}$ *is unbounded as* $t \to \infty$.

Proof. The form of the estimates follows from Theorem 2.7 and the property that for initial conditions as in (6-4) we have

$$\|A_B(w_0, w_1)\|_X^2 = \|A(w_{00}, w_1)\|_X^2 = \|w_{00}''\|_{L^2}^2 + \|w_1\|_{H^1}^2.$$

In order to prove (a), we will use Theorem 3.9. As shown in [Ammari and Tucsnak 2001, Lemma 5.3], we have $|s|\,\|D^*((1+is)^2 + L_{-1})^{-1}D\| \lesssim 1,\ s \in \mathbb{R}$. To verify the wavepacket condition, let ξ_0 be such that (6-5) holds. For a given $n \geq 2$, choose $m \in \mathbb{N}$ in such a way that $C_n \in \mathbb{R}$ defined by

$$\xi_0 = \frac{m}{n} + \frac{C_n}{n^2\log(n)^{1+\varepsilon}}$$

has minimal absolute value. By (6-5) we have $1 \leq |C_n| \leq n\log(n)^{1+\varepsilon}/2$ for all sufficiently large $n \geq 2$, and since $2r/\pi \leq \sin(r) \leq r$ for $0 \leq r \leq \pi/2$ it follows that

$$|D^*\phi_n| = \sqrt{2}|\sin(n\pi\xi_0)| = \sqrt{2}\left|\sin\left(\frac{C_n\pi}{n\log(n)^{1+\varepsilon}}\right)\right| \geq \frac{2\sqrt{2}}{n\log(n)^{1+\varepsilon}}$$

for all sufficiently large $n \geq 2$. Thus by Theorem 3.9 we have $\|(is - A_B)^{-1}\| \lesssim s^2\log(|s|)^{2+2\varepsilon},\ |s| \geq 2$, and hence (6-9) follows from Theorem 2.7; see also [Batty et al. 2016, Theorem 1.3].

In order to prove the optimality statement, note that by (6-6) there exist infinitely many $n \geq 2$ for which $|C_n| \leq \log(n)^{\varepsilon}$ and therefore also

$$|D^*\phi_n| = \sqrt{2}\left|\sin\left(\frac{C_n\pi}{n\log(n)^{1+\varepsilon}}\right)\right| \leq \frac{\sqrt{2}\pi}{n\log(n)}.$$

Now Proposition 5.1 shows that

$$\limsup_{|s|\to\infty} \frac{\|(is - A_B)^{-1}\|}{|s|^2\log(|s|)^2} > 0,$$

and it follows from Proposition 5.3 that

$$\limsup_{t\to\infty} \frac{\log(t)}{t^{-1/2}}\|T_B(t)A_B^{-1}\| > 0.$$

Now the optimality statement follows from a simple application of the uniform boundedness principle.

The argument for part (b) is entirely analogous and slightly simpler. It uses (6-7) and (6-8) in place of (6-5) and (6-6), respectively. □

6B2. *Weak damping.* In this section we consider a *weakly damped* wave equation on (0, 1), namely

$$w_{tt}(\xi, t) - w_{\xi\xi}(\xi, t) + b(\xi)\int_0^1 b(r)w_t(r, t)\,dr = 0, \quad \xi \in (0, 1),\ t > 0, \tag{6-10a}$$

$$w(0, t) = 0, \quad w(1, t) = 0, \quad t > 0, \tag{6-10b}$$

$$w(\cdot\,, 0) = w_0(\cdot) \in H^2(0, 1) \cap H_0^1(0, 1), \quad w_t(\cdot\,, 0) = w_1(\cdot) \in H_0^1(0, 1), \tag{6-10c}$$

where $b \in L^2(0,1;\mathbb{R})$ is the damping coefficient. The wave equation has the form considered in Section 2B with $H = L^2(0,1)$, $L = -\partial_{\xi\xi}$ with domain $H_1 = H^2(0,1) \cap H_0^1(0,1)$, and L has positive square root with domain $H_{1/2} = H_0^1(0,1)$. Moreover, $U = \mathbb{C}$ and $D \in \mathcal{L}(\mathbb{C}, H)$ is the rank-1 operator defined by $Du = bu$ for all $u \in \mathbb{C}$.

The operator L is the same as in Section 6B1. Hence if we define $\delta(s) \equiv \pi/4$ then the nontrivial $(s, \delta(s))$-wavepackets of $L^{1/2}$ are multiples of the normalised eigenfunctions ϕ_n for $n \in \mathbb{N}$ such that $n\pi \in (s - \pi/4, s + \pi/4)$. For any $n \in \mathbb{N}$ we have

$$|D^*\phi_n| = \sqrt{2}\left|\int_0^1 b(\xi)\sin(n\pi\xi)\,d\xi\right|.$$

For a large class of functions b these Fourier sine series coefficients have explicit expressions. In order to have $i\mathbb{R} \subseteq \rho(A_B)$ we require that $D^*\phi_n \neq 0$ for all $n \in \mathbb{N}$, and the rate at which $|D^*\phi_n|$ decays to zero as $n \to \infty$ determines the rate of resolvent growth. In the following we summarise the conclusions of Theorem 3.5 for a class of dampings.

Corollary 6.3. *Assume that* $|D^*\phi_n| \gtrsim f(n\pi)$, $n \in \mathbb{N}$, *for a continuous strictly decreasing function* $f : \mathbb{R}_+ \to (0,\infty)$ *such that* $f(\cdot)^{-1}$ *has positive increase. Then there exist* $C, t_0 > 0$ *such that for all* $w_0 \in H^2(0,1) \cap H_0^1(0,1)$ *and* $w_1 \in H_0^1(0,1)$ *the (classical) solution* w *of* (6-10) *satisfies*

$$\|(w(\cdot,t), w_t(\cdot,t))\|_{H^1\times L^2} \le \frac{C}{N^{-1}(t)}\|(w_0,w_1)\|_{H^2\times H^1}, \quad t \ge t_0, \tag{6-11}$$

where N^{-1} *is the inverse function of* $N(\cdot) := f(\cdot)^{-2}$. *Moreover, if there exists an increasing sequence* $(n_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ *such that* $|D^*\phi_{n_k}| \lesssim f(n_k\pi)$ *for all* $k \in \mathbb{N}$, *then the decay rate is optimal in the sense of Theorem 5.5.*

Proof. If $|D^*\phi_n| \gtrsim f(n\pi)$, $n \in \mathbb{N}$, then the wavepacket condition in (3-11) is satisfied for $\delta_0 = \pi/4$ and $\gamma_0(s) = f(s + \pi/4)$. Moreover, since $D \in \mathcal{L}(\mathbb{C}, H)$, we have $|s|\,\|D^*((1+is)^2 + L)^{-1}D\| \lesssim 1$, $s \in \mathbb{R}$. Thus Theorem 3.9 implies that $\|(is - A_B)^{-1}\| \lesssim f(|s| + \pi/4)^{-2}$, $s \in \mathbb{R}$, and Theorem 2.7 yields (6-11) with the function N_0 defined by $N_0(s) = f(s+\pi/4)^{-2}$ for $s > 0$. The claim now follows from the fact that $N^{-1} = N_0^{-1} + \pi/4$. □

For the particular damping functions b defined by $b(\xi) = 1 - \xi$, $b(\xi) = \xi^2(1-\xi)$ and $b(\xi) = \chi_{(0,\xi_0)}(\xi)$, where $\xi_0 \in (0,1)$ is an irrational of constant type, the optimal decay rates are given by (writing $b_n = D^*\phi_n$ for brevity)

$$b(\xi) = 1 - \xi, \qquad b_n = \frac{\sqrt{2}}{n\pi}, \qquad N^{-1}(t)^{-1} \lesssim t^{-1/2}, \tag{6-12a}$$

$$b(\xi) = \xi^2(1-\xi), \quad b_n = \frac{2\sqrt{2}(2(-1)^n - 1)}{n^3\pi^3}, \quad N^{-1}(t)^{-1} \lesssim t^{-1/6}, \tag{6-12b}$$

$$b(\xi) = \chi_{(0,\xi_0)}(\xi), \quad b_n = \frac{\sqrt{2}(1 - \cos(n\pi\xi_0))}{n\pi}, \quad N^{-1}(t)^{-1} \lesssim t^{-1/6}. \tag{6-12c}$$

The required upper and lower bounds for $|D^*\phi_n|$ in the third example follow by arguments similar to those used in the proof of Corollary 6.2, once again using (6-7) and (6-8). Optimality in all three examples is a consequence of Theorem 5.5.

Remark 6.4. The above discussion implies that the Fourier sine series coefficients $b_n = D^*\phi_n$ of the damping b determine the resolvent growth and thus the rate of energy decay in (6-10). So it is natural to try to relate the energy decay to the properties of b and $(b_n)_{n\in\mathbb{N}}$ directly. However, it is difficult to give a succinct answer here without specifying a precise class of functions b. First note that since $b \in L^2(0,1)$, we have $(b_n)_{n\in\mathbb{N}} \in \ell^2$. On the other hand, the results in [Nazarov 1997] show that for any $(c_n)_{n\in\mathbb{N}} \in \ell^2$ with $c_n \geq 0$ there exists $b \in C[0,1]$ such that $|b_n| \geq c_n$ for all $n \in \mathbb{N}$, and thus any rate of decay that can be achieved with a damping function $b \in L^2(0,1)$ can also be achieved with a more regular function $b \in C[0,1]$. However, imposing further regularity properties on b, such as Hölder-type conditions, changes the situation substantially.

In general, finer estimates for decay of $(b_n)_{n\in\mathbb{N}}$ depend heavily on the modulus of continuity (or the integral modulus of continuity) of b, and conversely for $(b_n)_{n\in\mathbb{N}}$ close in a sense to being monotone one may infer regularity properties of b from the sequence $(b_n)_{n\in\mathbb{N}}$; see for instance [Edwards 1979, Chapter 7], [Zygmund 2002, Chapter 5], and [Dyachenko et al. 2019].

Note finally that any polynomial rate of decay $t^{-\alpha}$ with $\alpha \in (0,1)$ can be achieved by choosing the damping function $b \in L^2(0,1)$ such that $b_n = n^{-1/(2\alpha)}$ for $n \in \mathbb{N}$. Moreover, by [Nazarov 1997] the same scale of polynomial rates can be realised by means of continuous damping functions. It would be interesting to consider similar statements about other scales of decay rates, for instance of regularly varying functions, but we do not pursue this here.

6C. ***A damped fractional Klein–Gordon equation.*** In this example we consider a "fractional Klein–Gordon equation" with viscous damping studied in [Malhi and Stanislavova 2020]; see also [Green 2020]. For a fixed $\alpha \in (0,1]$ this system has the form

$$w_{tt}(\xi,t) + (-\partial_{\xi\xi})^{\alpha} w(\xi,t) + m w(\xi,t) + b(\xi)^2 w_t(\xi,t) = 0, \quad \xi \in \mathbb{R},\ t > 0, \tag{6-13}$$

$$w(\,\cdot\,,0) = w_0(\,\cdot\,) \in H^{2\alpha}(\mathbb{R}), \quad w_t(\,\cdot\,,0) = w_1(\,\cdot\,) \in H^{\alpha}(\mathbb{R}), \tag{6-14}$$

where $m > 0$ and $b \in L^\infty(\mathbb{R})$ is the nonnegative damping coefficient. We assume that $\operatorname{ess\,inf}_{\xi\in\omega} b(\xi) > 0$ for some nonempty open set $\omega \subseteq \mathbb{R}$ which is invariant under translation by 2π.

Polynomial stability of this equation was studied, e.g., in [Malhi and Stanislavova 2020]. In the following proposition we use the wavepacket condition (3-11) to derive the same resolvent estimate under the above assumptions on b (strictly weaker conditions on the damping were also considered recently in [Green 2020]). The fractional Klein–Gordon equation is again of the form studied in Section 2B, now with $H = U = L^2(\mathbb{R})$, $L = (-\partial_{\xi\xi})^{\alpha} + m > 0$ with domain $H_1 = H^{2\alpha}(\mathbb{R})$ and $H_{1/2} = H^{\alpha}(\mathbb{R})$. The damping operator $D \in \mathcal{L}(L^2(\mathbb{R}))$ is the multiplication operator defined by $Du = bu$ for all $u \in L^2(\mathbb{R})$.

Proposition 6.5. *Let* $0 < \alpha < 1$. *There exists* $C > 0$ *such that for every* $w_0 \in H^{2\alpha}(\mathbb{R})$ *and* $w_1 \in H^{\alpha}(\mathbb{R})$ *the solution* w *of the fractional Klein–Gordon equation satisfies*

$$\|(w(\,\cdot\,,t), w_t(\,\cdot\,,t))\|_{H^{\alpha}\times L^2} \leq \frac{C}{t^{\alpha/(2-2\alpha)}} \|(w_0,w_1)\|_{H^{2\alpha}\times H^{\alpha}}, \quad t > 0.$$

Proof. Let us begin by showing that the classical Klein–Gordon equation corresponding to $\alpha = 1$ is exponentially stable. Due to the properties of the damping coefficients we may choose a smooth and

2π-periodic function b_1 such that $0 \le b_1 \le b$ on $\mathbb{R}$ and $\inf_{\xi\in\omega_1} b_1(\xi) > 0$ for a nonempty open set $\omega_1 \subseteq \omega$. By [Burq and Joly 2016, Theorem 1.2] the Klein–Gordon equation with damping coefficient b_1 is exponentially stable. If we define $D_1 \in \mathcal{L}(L^2(\mathbb{R}))$ so that $D_1 u = b_1 u$ for all $u \in L^2(\mathbb{R})$, and define $B_1 = \left(\begin{smallmatrix} 0 \\ D_1 \end{smallmatrix}\right)$, then (B_1^*, A) is exactly observable, and by [Miller 2012, Corollary 2.17] the pair $(\mathcal{B}_1^*, A)$ satisfies the wavepacket condition (3-4) for constant functions $\delta(s) \equiv \delta > 0$ and $\gamma(s) \equiv \gamma > 0$. However, since $b(\xi) \ge b_1(\xi)$ for all $\xi \in \mathbb{R}$ we see that also (B^*, A) satisfies the wavepacket condition for the same functions δ and γ.

Let us temporarily write L_α for the operator $(-\partial_{\xi\xi})^\alpha + m,\ 0 < \alpha \le 1$, accepting that this entails a minor abuse of notation. Since $\sigma(L_\alpha) \subseteq [m, \infty)$ for $0 < \alpha \le 1$, we obtain from Lemma 3.8 that

$$\|D^* w\|_U \ge \gamma_1 \|w\|_H \tag{6-15}$$

for all (s, δ_1)-wavepackets w of $L_\alpha^{1/2}$, where $\delta_1, \gamma_1 > 0$ are suitable constants.

For $0 < \alpha \le 1$ and any bounded function $\delta_0 : \mathbb{R}_+ \to (0, \infty)$ the $(s, \delta_0(s))$-wavepackets of $L_\alpha^{1/2}$ are precisely the elements of $\operatorname{Ran}(\chi_{I_{s,\delta_0(s)}}(L_\alpha^{1/2}))$, where $I_{s,\delta_0(s)} = (s - \delta_0(s), s + \delta_0(s))$. Using the spectral theorem we see that if $I \subseteq [\sqrt{m}, \infty)$ is a bounded interval then $\operatorname{Ran}(\chi_I(L_\alpha^{1/2})) = \operatorname{Ran}(\chi_{J_\alpha}(L_1^{1/2}))$, where $J_\alpha = ((I^2 - m)^{1/\alpha} + m)^{1/2}$. Now fix $\alpha \in (0, 1)$ and let $\delta_0(s) = c(1 + s^{\alpha^{-1}-1}),\ s \ge 0$, where $c > 0$ is a constant. Straightforward estimates show that the images of the intervals $I_{s,\delta_0} \cap [\sqrt{m}, \infty)$ under the map $I \mapsto J_\alpha$ have length bounded by some constant multiple of c. It follows that (6-15) holds also for all $(s, \delta_0(s))$-wavepackets w of $L_\alpha^{1/2}$ provided that c is sufficiently small. (Here the form of the function δ_0 can either be guessed or alternatively derived by considering the images of constant-width intervals under the inverse of the map $I \mapsto J_\alpha$.) Moreover, since $D \in \mathcal{L}(L^2(\Omega))$, we have $|s|\,\|D^*((1+is)^2+L)^{-1}D\| \lesssim 1$, $s \in \mathbb{R}$. Thus we deduce from Theorem 3.9 that $\|(is - A_B)^{-1}\| \lesssim 1 + |s|^{2(\alpha^{-1}-1)^{-1}}$ for $s \in \mathbb{R}$. The claim now follows directly from Theorem 2.7. ☐

6D. ***A weakly damped beam equation.*** In this section we consider the stability of the following Euler–Bernoulli beam equation with weak damping:

$$w_{tt}(\xi, t) + w_{\xi\xi\xi\xi}(\xi, t) + b(\xi)\int_0^1 b(r) w_t(r, t)\,dr = 0, \quad \xi \in (0, 1),\ t > 0, \tag{6-16}$$

$$w(0, t) = 0, \quad w_{\xi\xi}(0, t) = 0,\ t > 0, \tag{6-17}$$

$$w(1, t) = 0, \quad w_{\xi\xi}(1, t) = 0,\ t > 0, \tag{6-18}$$

$$w(\,\cdot\,, 0) = w_0(\,\cdot\,) \in H^4(0, 1) \cap H_0^1(0, 1), \tag{6-19}$$

$$w_t(\,\cdot\,, 0) = w_1(\,\cdot\,) \in H^2(0, 1) \cap H_0^1(0, 1), \tag{6-20}$$

where $b \in L^2(0, 1; \mathbb{R})$ is the damping coefficient. The boundary conditions describe a situation in which the beam is *simply supported*.

The beam equation fits into the framework of Section 2B with the choices $H = L^2(0, 1)$ and

$$L = \partial_{\xi\xi\xi\xi}, \quad H_1 = \{x \in H^4(0, 1) : x(0) = x''(0) = x(1) = x''(1) = 0\}.$$

The operator L is invertible and positive and its positive square root is given by $L^{1/2} = -\partial_{\xi\xi}$ with domain $H_{1/2} = H^2(0, 1) \cap H_0^1(0, 1)$. The eigenvalues and normalised eigenfunctions of $L^{1/2}$ are given by

$\lambda_n = n^2\pi^2$ and $\phi_n(\,\cdot\,) = \sqrt{2}\sin(n\pi\,\cdot\,)$, respectively, for $n \in \mathbb{N}$. As in Section 6B2, $U = \mathbb{C}$ and $D \in \mathcal{L}(\mathbb{C}, H)$ is the rank-1 operator defined by $Du = bu$ for all $u \in \mathbb{C}$.

Our aim is to study the asymptotic behaviour of the solutions of the damped beam equation using the wavepacket condition in Theorem 3.9. Since the eigenvalues $\lambda_n = n^2\pi^2$, $n \in \mathbb{N}$, have a uniform gap, we may choose $\delta(s) \equiv \pi^2/4$. The nontrivial $(s, \delta(s))$-wavepackets of $L^{1/2}$ are then multiples of the eigenfunctions ϕ_n for $n \in \mathbb{N}$ such that $n^2\pi^2 \in (s - \pi^2/4, s + \pi^2/4)$. For any $n \in \mathbb{N}$ we have

$$|D^*\phi_n| = \sqrt{2}\left|\int_0^1 b(\xi)\sin(n\pi\xi)\,d\xi\right|.$$

These Fourier sine series coefficients are identical to the ones in Section 6B2. However, the locations of the eigenvalues of A now result in a slower rate of resolvent growth than in the case of the wave equation. In order to have $i\mathbb{R} \subseteq \rho(A_B)$ it is again necessary that $D^*\phi_n \neq 0$ for all $n \in \mathbb{N}$. However, since the gaps between the eigenvalues $n^2\pi^2$ of $L^{1/2}$ grow without bound as $n \to \infty$, the same damping has a greater relative effect for the beam equation than for the wave equation.

Corollary 6.6. *Assume that $|D^*\phi_n| \gtrsim f(n^2\pi^2)$ for a continuous strictly decreasing function $f : \mathbb{R}_+ \to (0, \infty)$ such that $f(\,\cdot\,)^{-1}$ has positive increase. Then there exist $C, t_0 > 0$ such that for every $w_0 \in H_1$ and $w_1 \in H_{1/2}$ the (classical) solution of the weakly damped beam equation satisfies*

$$\|(w(\,\cdot\,,t), w_t(\,\cdot\,,t))\|_{H^2\times L^2} \le \frac{C}{N^{-1}(t)}\|(w_0, w_1)\|_{H^4\times H^2}, \quad t \ge t_0,$$

where N^{-1} is the inverse function of $N(\,\cdot\,) := f(\,\cdot\,)^{-2}$. Moreover, if there exists an increasing sequence $(n_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ such that $|D^*\phi_{n_k}| \lesssim f(n_k\pi)$ for all $k \in \mathbb{N}$, then the decay rate is optimal in the sense of Theorem 5.5.

The coefficients $|D^*\phi_n|$ for the functions b defined by $b(\xi) = 1 - \xi$, $b(\xi) = \xi^2(1-\xi)$ and $b(\xi) = \chi_{(0,\xi_0)}(\xi)$ (with $\xi_0 \in (0, 1)$ an irrational number of constant type) are presented in (6-12), and for these functions Corollary 6.6 implies the asymptotic rates t^{-1}, $t^{-1/3}$ and $t^{-1/3}$ as $t \to \infty$, respectively. Note finally that Remark 6.4 applies also in the setting of this section.

References

[Ammari and Nicaise 2015] K. Ammari and S. Nicaise, *Stabilization of elastic systems by collocated feedback*, Lecture Notes in Math. **2124**, Springer, 2015. MR Zbl

[Ammari and Tucsnak 2001] K. Ammari and M. Tucsnak, "Stabilization of second order evolution equations by a class of unbounded feedbacks", *ESAIM Control Optim. Calc. Var.* **6** (2001), 361–386. MR Zbl

[Ammari et al. 2017] K. Ammari, A. Bchatnia, and K. El Mufti, "Non-uniform decay of the energy of some dissipative evolution systems", *Z. Anal. Anwend.* **36**:2 (2017), 239–251. MR Zbl

[Anantharaman and Léautaud 2014] N. Anantharaman and M. Léautaud, "Sharp polynomial decay rates for the damped wave equation on the torus", *Anal. PDE* **7**:1 (2014), 159–214. MR Zbl

[Anantharaman et al. 2016] N. Anantharaman, M. Léautaud, and F. Macià, "Wigner measures and observability for the Schrödinger equation on the disk", *Invent. Math.* **206**:2 (2016), 485–599. MR Zbl

[Arendt et al. 2011] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace transforms and Cauchy problems*, 2nd ed., Monogr. Math. **96**, Birkhäuser, Basel, 2011. MR Zbl

[Batty and Duyckaerts 2008] C. J. K. Batty and T. Duyckaerts, "Non-uniform stability for bounded semi-groups on Banach spaces", *J. Evol. Equ.* **8**:4 (2008), 765–780. MR Zbl

[Batty et al. 2016] C. J. K. Batty, R. Chill, and Y. Tomilov, "Fine scales of decay of operator semigroups", *J. Eur. Math. Soc.* **18**:4 (2016), 853–929. MR Zbl

[Benchimol 1977] C. D. Benchimol, "Feedback stabilizability in Hilbert spaces", *Appl. Math. Optim.* **4**:3 (1977), 225–248. MR Zbl

[Borichev and Tomilov 2010] A. Borichev and Y. Tomilov, "Optimal polynomial decay of functions and operator semigroups", *Math. Ann.* **347**:2 (2010), 455–478. MR Zbl

[Burq and Hitrik 2007] N. Burq and M. Hitrik, "Energy decay for damped wave equations on partially rectangular domains", *Math. Res. Lett.* **14**:1 (2007), 35–47. MR Zbl

[Burq and Joly 2016] N. Burq and R. Joly, "Exponential decay for the damped wave equation in unbounded domains", *Commun. Contemp. Math.* **18**:6 (2016), art. id. 1650012. MR Zbl

[Burq and Zuily 2016] N. Burq and C. Zuily, "Concentration of Laplace eigenfunctions and stabilization of weakly damped wave equation", *Comm. Math. Phys.* **345**:3 (2016), 1055–1076. MR Zbl

[Burq and Zworski 2004] N. Burq and M. Zworski, "Geometric control in the presence of a black box", *J. Amer. Math. Soc.* **17**:2 (2004), 443–471. MR Zbl

[Burq and Zworski 2019] N. Burq and M. Zworski, "Rough controls for Schrödinger operators on 2-tori", *Ann. H. Lebesgue* **2** (2019), 331–347. MR Zbl

[Cavalcanti et al. 2019] M. M. Cavalcanti, T. F. Ma, P. Marín-Rubio, and P. N. Seminario-Huertas, "Dynamics of Riemann waves with sharp measure-controlled damping", preprint, 2019. arXiv 1908.04814

[Chen et al. 1991] G. Chen, S. A. Fulling, F. J. Narcowich, and S. Sun, "Exponential decay of energy of evolution equations with locally distributed damping", *SIAM J. Appl. Math.* **51**:1 (1991), 266–301. MR Zbl

[Chill et al. 2020] R. Chill, D. Seifert, and Y. Tomilov, "Semi-uniform stability of operator semigroups and energy decay of damped waves", *Philos. Trans. Roy. Soc. A* **378**:2185 (2020), art. id. 20190614. MR

[Curtain and Weiss 2006] R. F. Curtain and G. Weiss, "Exponential stabilization of well-posed systems by colocated feedback", *SIAM J. Control Optim.* **45**:1 (2006), 273–297. MR Zbl

[Curtain and Weiss 2019] R. F. Curtain and G. Weiss, "Strong stabilization of (almost) impedance passive systems by static output feedback", *Math. Control Relat. Fields* **9**:4 (2019), 643–671. MR Zbl

[Datchev and Kleinhenz 2020] K. Datchev and P. Kleinhenz, "Sharp polynomial decay rates for the damped wave equation with Hölder-like damping", *Proc. Amer. Math. Soc.* **148**:8 (2020), 3417–3425. MR Zbl

[Debruyne and Seifert 2019] G. Debruyne and D. Seifert, "Optimality of the quantified Ingham–Karamata theorem for operator semigroups with general resolvent growth", *Arch. Math.* (*Basel*) **113**:6 (2019), 617–627. MR Zbl

[Dell'Oro and Pata 2021] F. Dell'Oro and V. Pata, "Second order linear evolution equations with general dissipation", *Appl. Math. Optim.* **83**:3 (2021), 1877–1917. MR Zbl

[Duyckaerts 2007] T. Duyckaerts, "Optimal decay rates of the energy of a hyperbolic-parabolic system coupled by an interface", *Asymptot. Anal.* **51**:1 (2007), 17–45. MR Zbl

[Dyachenko et al. 2019] M. I. Dyachenko, A. B. Mukanov, and S. Y. Tikhonov, "Smoothness of functions and Fourier coefficients", *Mat. Sb.* **210**:7 (2019), 94–119. In Russian; translated in *Sb. Math.* **210**:7 (2019), 994–1018. MR Zbl

[Edwards 1979] R. E. Edwards, *Fourier series: a modern introduction, I*, 2nd ed., Grad. Texts in Math. **64**, Springer, 1979. MR Zbl

[Gohberg and Sigal 1971] I. C. Gohberg and E. I. Sigal, "An operator generalization of the logarithmic residue theorem and the theorem of Rouché", *Mat. Sb.* (*N.S.*) **84(126)** (1971), 607–629. In Russian; translated in *Math. USSR-Sb.* **13** (1971), 603–625. MR Zbl

[Green 2020] W. Green, "On the energy decay rate of the fractional wave equation on $\mathbb{R}$ with relatively dense damping", *Proc. Amer. Math. Soc.* **148**:11 (2020), 4745–4753. MR Zbl

[Guo and Luo 2002] B.-Z. Guo and Y.-H. Luo, "Controllability and stability of a second-order hyperbolic system with collocated sensor/actuator", *Systems Control Lett.* **46**:1 (2002), 45–65. MR Zbl

[Hassell et al. 2009] A. Hassell, L. Hillairet, and J. Marzuola, "Eigenfunction concentration for polygonal billiards", *Comm. Partial Differential Equations* **34**:4-6 (2009), 475–485. MR Zbl

[Jaffard 1990] S. Jaffard, "Contrôle interne exact des vibrations d'une plaque rectangulaire", *Port. Math.* **47**:4 (1990), 423–429. MR Zbl

[Joly and Laurent 2020] R. Joly and C. Laurent, "Decay of semilinear damped wave equations: cases without geometric control condition", *Ann. H. Lebesgue* **3** (2020), 1241–1289. MR Zbl

[Kato 1961] T. Kato, "A generalization of the Heinz inequality", *Proc. Japan Acad.* **37** (1961), 305–308. MR Zbl

[Khinchin 1964] A. Y. Khinchin, *Continued fractions*, Univ. Chicago Press, 1964. MR Zbl

[Lang 1966] S. Lang, *Introduction to Diophantine approximations*, Addison-Wesley, Reading, MA, 1966. MR Zbl

[Lasiecka and Triggiani 1981] I. Lasiecka and R. Triggiani, "A cosine operator approach to modeling $L_2(0, T; L_2(\Gamma))$: boundary input hyperbolic equations", *Appl. Math. Optim.* **7**:1 (1981), 35–93. MR Zbl

[Lasiecka and Triggiani 2000] I. Lasiecka and R. Triggiani, *Control theory for partial differential equations: continuous and approximation theories, II: Abstract hyperbolic-like systems over a finite time horizon*, Encycl. Math. Appl. **75**, Cambridge Univ. Press, 2000. MR Zbl

[Lasiecka and Triggiani 2003] I. Lasiecka and R. Triggiani, "$L_2(\Sigma)$-regularity of the boundary to boundary operator B^*L for hyperbolic and Petrowski PDEs", *Abstr. Appl. Anal.* **2003**:19 (2003), 1061–1139. MR Zbl

[Latushkin and Shvydkoy 2001] Y. Latushkin and R. Shvydkoy, "Hyperbolicity of semigroups and Fourier multipliers", pp. 341–363 in *Systems, approximation, singular integral operators, and related topics* (Bordeaux, France, 2000), edited by A. A. Borichev and N. K. Nikolski, Oper. Theory Adv. Appl. **129**, Birkhäuser, Basel, 2001. MR Zbl

[Laurent and Léautaud 2021] C. Laurent and M. Léautaud, "Logarithmic decay for linear damped hypoelliptic wave and Schrödinger equations", *SIAM J. Control Optim.* **59**:3 (2021), 1881–1902. MR Zbl

[Léautaud and Lerner 2017] M. Léautaud and N. Lerner, "Energy decay for a locally undamped wave equation", *Ann. Fac. Sci. Toulouse Math.* (6) **26**:1 (2017), 157–205. MR Zbl

[Lebeau 1996] G. Lebeau, "Équation des ondes amorties", pp. 73–109 in *Algebraic and geometric methods in mathematical physics* (Crimea, Ukraine, 1993), edited by A. Boutet de Monvel and V. Marchenko, Math. Phys. Stud. **19**, Kluwer Acad., Dordrecht, 1996. MR Zbl

[Letrouit and Sun 2023] C. Letrouit and C. Sun, "Observability of Baouendi–Grushin-type equations through resolvent estimates", *J. Inst. Math. Jussieu* **22**:2 (2023), 541–579. MR Zbl

[Liu and Rao 2005] Z. Liu and B. Rao, "Characterization of polynomial decay rate for the solution of linear evolution equation", *Z. Angew. Math. Phys.* **56**:4 (2005), 630–644. MR Zbl

[Liu and Zhang 2015] Z. Liu and Q. Zhang, "A note on the polynomial stability of a weakly damped elastic abstract system", *Z. Angew. Math. Phys.* **66**:4 (2015), 1799–1804. MR Zbl

[Malhi and Stanislavova 2020] S. Malhi and M. Stanislavova, "On the energy decay rates for the 1D damped fractional Klein–Gordon equation", *Math. Nachr.* **293**:2 (2020), 363–375. MR Zbl

[Miller 2012] L. Miller, "Resolvent conditions for the control of unitary groups and their approximations", *J. Spectr. Theory* **2**:1 (2012), 1–55. MR Zbl

[Nazarov 1997] F. L. Nazarov, "The Bang solution of the coefficient problem", *Algebra i Analiz* **9**:2 (1997), 272–287. In Russian; translated in *St. Petersburg Math. J.* **9**:2 (1998), 407–419. MR Zbl

[Oostveen 2000] J. Oostveen, *Strongly stabilizable distributed parameter systems*, Frontiers Appl. Math. **20**, Soc. Indust. Appl. Math., Philadelphia, PA, 2000. MR Zbl

[Paunonen 2017] L. Paunonen, "Robust controllers for regular linear systems with infinite-dimensional exosystems", *SIAM J. Control Optim.* **55**:3 (2017), 1567–1597. MR Zbl

[Ramdani et al. 2005] K. Ramdani, T. Takahashi, G. Tenenbaum, and M. Tucsnak, "A spectral approach for the exact observability of infinite-dimensional systems with skew-adjoint generator", *J. Funct. Anal.* **226**:1 (2005), 193–229. MR Zbl

[Rauch et al. 2005] J. Rauch, X. Zhang, and E. Zuazua, "Polynomial decay for a hyperbolic-parabolic coupled system", *J. Math. Pures Appl.* (9) **84**:4 (2005), 407–470. MR Zbl

[Rozendaal et al. 2019] J. Rozendaal, D. Seifert, and R. Stahn, "Optimal rates of decay for operator semigroups on Hilbert spaces", *Adv. Math.* **346** (2019), 359–388. MR Zbl

[Russell 1975] D. L. Russell, "Decay rates for weakly damped systems in Hilbert space obtained with control-theoretic methods", *J. Differential Equations* **19**:2 (1975), 344–370. MR Zbl

[Rzepnicki and Schnaubelt 2018] Ł. Rzepnicki and R. Schnaubelt, "Polynomial stability for a system of coupled strings", *Bull. Lond. Math. Soc.* **50**:6 (2018), 1117–1136. MR Zbl

[Salamon 1987] D. Salamon, "Infinite-dimensional linear systems with unbounded control and observation: a functional analytic approach", *Trans. Amer. Math. Soc.* **300**:2 (1987), 383–431. MR Zbl

[Slemrod 1974] M. Slemrod, "A note on complete controllability and stabilizability for linear control systems in Hilbert space", *SIAM J. Control* **12** (1974), 500–508. MR Zbl

[Sogge 1988] C. D. Sogge, "Concerning the L^p norm of spectral clusters for second-order elliptic operators on compact manifolds", *J. Funct. Anal.* **77**:1 (1988), 123–138. MR Zbl

[Staffans 2002] O. J. Staffans, "Passive and conservative continuous-time impedance and scattering systems, I: Well-posed systems", *Math. Control Signals Systems* **15**:4 (2002), 291–315. MR Zbl

[Stahn 2017] R. Stahn, "Optimal decay rate for the wave equation on a square with constant damping on a strip", *Z. Angew. Math. Phys.* **68**:2 (2017), art. id. 36. MR Zbl

[Su et al. 2020] P. Su, M. Tucsnak, and G. Weiss, "Stabilizability properties of a linearized water waves system", *Systems Control Lett.* **139** (2020), art. id. 104672. MR Zbl

[Sun 2023] C. Sun, "Sharp decay rate for the damped wave equation with convex-shaped damping", *Int. Math. Res. Not. IMRN* **2023**:7 (2023), 5905–5973. MR Zbl

[Tucsnak and Weiss 2009] M. Tucsnak and G. Weiss, *Observation and control for operator semigroups*, Birkhäuser, Basel, 2009. MR Zbl

[Tucsnak and Weiss 2014] M. Tucsnak and G. Weiss, "Well-posed systems: the LTI case and beyond", *Automatica J. IFAC* **50**:7 (2014), 1757–1779. MR Zbl

[Weiss 2003] G. Weiss, "Optimal control of systems with a unitary semigroup and with colocated control and observation", *Systems Control Lett.* **48**:3-4 (2003), 329–340. MR Zbl

[Weiss and Tucsnak 2003] G. Weiss and M. Tucsnak, "How to get a conservative well-posed linear system out of thin air, I: Well-posedness and energy balance", *ESAIM Control Optim. Calc. Var.* **9** (2003), 247–274. MR Zbl

[Zygmund 2002] A. Zygmund, *Trigonometric series, I, II*, 3rd ed., Cambridge Univ. Press, 2002. MR Zbl

Received 20 Dec 2019. Revised 15 Sep 2021. Accepted 19 Nov 2021.

RALPH CHILL: ralph.chill@tu-dresden.de
Institut für Analysis, Fakultät für Mathematik, TU Dresden, Dresden, Germany

LASSI PAUNONEN: lassi.paunonen@tuni.fi
Mathematics, Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland

DAVID SEIFERT: david.seifert@ncl.ac.uk
School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne, United Kingdom

REINHARD STAHN: reinhardstahn@t-online.de
Institut für Analysis, Fakultät für Mathematik, TU Dresden, Dresden, Germany

YURI TOMILOV: ytomilov@impan.pl
Institute of Mathematics, Polish Academy of Sciences, Warsaw, Poland

ANALYSIS AND PDE
Vol. 16 (2023), No. 5, pp. 1133–1203
DOI: 10.2140/apde.2023.16.1133

LONG TIME SOLUTIONS FOR QUASILINEAR HAMILTONIAN PERTURBATIONS OF SCHRÖDINGER AND KLEIN–GORDON EQUATIONS ON TORI

ROBERTO FEOLA, BENOÎT GRÉBERT AND FELICE IANDOLI

We consider quasilinear, Hamiltonian perturbations of the cubic Schrödinger and of the cubic (derivative) Klein–Gordon equations on the d-dimensional torus. If $\epsilon \ll 1$ is the size of the initial datum, we prove that the lifespan of solutions is *strictly* larger than the local existence time ϵ^{-2}. More precisely, concerning the Schrödinger equation we show that the lifespan is at least of order $O(\epsilon^{-4})$, and in the Klein–Gordon case we prove that the solutions exist at least for a time of order $O(\epsilon^{-8/3^-})$ as soon as $d \geq 3$. Regarding the Klein–Gordon equation, our result presents novelties also in the case of semilinear perturbations: we show that the lifespan is at least of order $O(\epsilon^{-10/3^-})$, improving, for cubic nonlinearities and $d \geq 4$, the general results of Delort (*J. Anal. Math.* **107** (2009), 161–194) and Fang and Zhang (*J. Differential Equations* **249**:1 (2010), 151–179).

1.	Introduction	1133
2.	Small divisors	1144
3.	Functional setting	1150
4.	Paradifferential formulation of the problems	1163
5.	Approximately symplectic maps	1170
6.	Diagonalization	1178
7.	Energy estimates	1186
8.	Proof of the main results	1198
	Acknowledgements	1200
	References	1201

1. Introduction

This paper is concerned with the study of the lifespan of solutions of two classes of quasilinear, Hamiltonian equations on the d-dimensional torus $\mathbb{T}^d := (\mathbb{R}/2\pi\mathbb{Z})^d$, $d \geq 1$. We study quasilinear perturbations of the Schrödinger and Klein–Gordon equations.

The Schrödinger equation we consider is

$$\begin{cases} \mathrm{i}\partial_t u + \Delta u - V * u + [\Delta(h(|u|^2))]h'(|u|^2)u - |u|^2 u = 0, \\ u(0,x) = u_0(x), \end{cases} \tag{NLS}$$

Felice Iandoli was supported by ERC grant ANADEL 757996. Roberto Feola and Benoît Grébert were supported by the Centre Henri Lebesgue ANR-11-LABX- 0020-01 and by ANR-15-CE40-0001-02 "BEKAM" of the ANR.
MSC2020: 37K45, 35S50, 35B35, 35B45, 35L05, 35Q55.
Keywords: quasilinear equations, paradifferential calculus, energy estimates, small divisors.

where $\mathbb{C} \ni u := u(t,x),\ x \in \mathbb{T}^d,\ d \geq 1,\ V(x)$ is a real-valued potential even with respect to x, $h(x)$ is a function in $C^\infty(\mathbb{R};\mathbb{R})$ such that $h(x) = O(x^2)$ as $x \to 0$. The initial datum u_0 has small size and belongs to the Sobolev space $H^s(\mathbb{T}^d)$ (see (3-2)) with $s \gg 1$.

We examine also the Klein–Gordon equation

$$\begin{cases} \partial_{tt}\psi - \Delta\psi + m\psi + f(\psi) + g(\psi) = 0, \\ \psi(0,x) = \psi_0, \\ \partial_t\psi(0,x) = \psi_1, \end{cases} \tag{KG}$$

where $\mathbb{R} \ni \psi := \psi(t,x),\ x \in \mathbb{T}^d,\ d \geq 1$ and $m > 0$. The initial data (ψ_0, ψ_1) have small size and belong to the Sobolev space $H^s(\mathbb{T}^d) \times H^{s-1}(\mathbb{T}^d)$ for some $s \gg 1$. The nonlinearity $f(\psi)$ has the form

$$f(\psi) := -\sum_{j=1}^{d} \partial_{x_j}(\partial_{\psi_{x_j}} F(\psi, \nabla\psi)) + (\partial_\psi F)(\psi, \nabla\psi), \tag{1-1}$$

where $F(y_0, y_1, \dots, y_d) \in C^\infty(\mathbb{R}^{d+1}, \mathbb{R})$, and has a zero of order at least 5 at the origin. The nonlinear term $g(\psi)$ has the form

$$g(\psi) = (\partial_{y_0} G)(\psi, \Lambda_{\mathrm{KG}}^{1/2}\psi) + \Lambda_{\mathrm{KG}}^{1/2}(\partial_{y_1} G)(\psi, \Lambda_{\mathrm{KG}}^{1/2}\psi), \tag{1-2}$$

where $G(y_0, y_1) \in C^\infty(\mathbb{R}^2;\mathbb{R})$ is a homogeneous polynomial of degree 4 and Λ_{KG} is the operator

$$\Lambda_{\mathrm{KG}} := (-\Delta + m)^{1/2}, \tag{1-3}$$

defined by linearity as

$$\Lambda_{\mathrm{KG}} e^{\mathrm{i} j\cdot x} = \Lambda_{\mathrm{KG}}(j) e^{\mathrm{i} j\cdot x}, \qquad \Lambda_{\mathrm{KG}}(j) = \sqrt{|j|^2 + m} \quad \text{for all } j \in \mathbb{Z}^d. \tag{1-4}$$

***Historical introduction for* (NLS).** Quasilinear Schrödinger equations of the specific form (NLS) appear in many domains of physics like plasma physics and fluid mechanics [Litvak and Sergeev 1978; Porkolab and Goldman 1976], quantum mechanics [Hasse 1980], and condensed matter theory [Makhankov and Fedyanin 1984]. They are also important in the study of Kelvin waves in the superfluid turbulence [Laurie et al. 2010]. Equations of the form (NLS) posed in the *Euclidean space* have received the attention of many mathematicians. The first result, concerning the local well-posedness, is due to Poppenberg [2001] in the one-dimensional case. This has been generalized by Colin [2002] to any dimension. A more general class of equations is considered in the pioneering work by Kenig, Ponce and Vega [Kenig et al. 2004]. These results of local well-posedness have been recently optimized, in terms of regularity of the initial condition, by Marzuola, Metcalfe and Tataru [Marzuola et al. 2021]. Existence of standing waves has been studied in [Colin 2003; Colin and Jeanjean 2004]. The global well-posedness was established by de Bouard, Hayashi and Saut [de Bouard et al. 1997] in dimensions 2 and 3 for small data. This proof is based on dispersive estimates and the energy method. New ideas have been introduced in studying the global well-posedness for other quasilinear equations on the *Euclidean space*. Here the aforementioned tools are combined with *normal form* reductions. We quote [Ionescu and Pusateri 2015; 2018] for the water-waves equation in two dimensions.

Very little is known when (NLS) is posed on a *compact manifold*. The first local well-posedness results on the circle are given in the work by Baldi, Haus and Montalto [Baldi et al. 2018] and in [Feola and Iandoli 2019]. Recently these results have been generalized to the case of tori of any dimension in [Feola and Iandoli 2022]. Except these local existence results, nothing is known concerning the long time behavior of the solutions. The problem of *global existence/blow-up* is completely open. In the aforementioned paper [de Bouard et al. 1997] they use the dispersive character of the flow of the linear Schrödinger equation. This property is not present on compact manifolds: the solutions of the linear Schrödinger equation do not decay when the time goes to infinity. However in the one-dimensional case in [Feola and Iandoli 2020; 2021] it is proven that *small* solutions of quasilinear Schrödinger equations exist for long, but finite, times. In these works two of us exploit the fact that quasilinear Schrödinger equations may be reduced to constant coefficients through a *paracomposition* generated by a diffeomorphism of the circle. This powerful tool has been used for the same purpose by other authors in the context of water-waves equations, firstly by Berti and Delort [2018] in a *nonresonant* regime, and secondly by Berti, Feola and Pusateri [Berti et al. 2023; 2021b] and Berti, Feola and Franzoi [Berti et al. 2021a] in the *resonant* case. We also mention that this feature has been used in other contexts for the same equations; for instance Feola and Procesi [2015] proved the existence of a large set of *quasiperiodic* (and hence globally defined) solutions when the problem is posed on the circle. This "reduction to constant coefficients" is a peculiarity of one-dimensional problems; in higher dimensions new ideas have to be introduced. For quasilinear equations on tori of dimension 2 we quote the paper about long-time solutions for water-waves problem in [Ionescu and Pusateri 2019], where a different *normal form* analysis was presented.

***Historical introduction for* (KG).** The local existence for (KG) is classical and we refer to [Kato 1975]. Many analyses have been done for *global/long time* existence.

When the equation is posed on the Euclidean space we have global existence for small and localized data in [Delort 2016; Stingo 2018]; here the authors use dispersive estimates on the linear flow combined with quasilinear normal forms.

For (KG) on compact manifolds we quote [Delort 2012; 2015] on $\mathbb{S}^d$ and [Delort and Szeftel 2004] on $\mathbb{T}^d$. The results obtained, in terms of length of the lifespan of solutions, are stronger in the case of the spheres. More precisely, in the case of spheres the authors show the following: if m in (KG) is chosen outside of a set of zero Lebesgue measure, then for any natural number N, any initial condition of size ϵ (small depending on N) produces a solution whose lifespan is at least of magnitude ϵ^{-N}. In the case of tori in [Delort and Szeftel 2004] they consider a quasilinear equation, vanishing quadratically at the origin and they prove that the lifespan of solutions is of order ϵ^{-2} if the initial condition has size ϵ small enough. The differences between the two results are due to the different behaviors of the eigenvalues of the square root of the Laplace–Beltrami operator on $\mathbb{S}^d$ and $\mathbb{T}^d$. The difficulty on the tori is a consequence of the fact that the set of differences of eigenvalues of $\sqrt{-\Delta_{\mathbb{T}^d}}$ is dense in $\mathbb{R}$ if $d \geq 2$; this does not happen in the case of spheres. A more general set of manifolds where this does not happen is the Zoll manifolds; in this case we quote Delort and Szeftel [2006] and Bambusi, Delort, Grébert and Szeftel [Bambusi et al. 2007] for semilinear Klein–Gordon equations. For semilinear Klein–Gordon equations on tori we have the results of [Delort 2009; Fang and Zhang 2010]. In [Delort 2009] the author proves that if the nonlinearity is vanishing at order $k+1$ at zero then any initial datum of small size ϵ produces a solution whose lifespan is at least of magnitude $\epsilon^{-k(1+2/d)}$, up to a logarithmic loss. In [Fang and Zhang 2010] the authors obtain a time $O(\epsilon^{-k(3/2)^-})$. We improve these results; see Theorems 4 and 3, when $k=2$.

Statement of the main results. The aim of this paper is to prove, in the spirit of [Ionescu and Pusateri 2019], that we may go beyond the trivial time of existence, given by the local well-posedness theorem, which is ϵ^{-2} since we are considering equations vanishing cubically at the origin and initial conditions of size ϵ.

In order to state our main theorem for (NLS) we need to make some hypotheses on the potential V. We consider potentials having the form

$$\begin{gathered} V(x)=(2\pi)^{-d/2}\sum_{\xi\in\mathbb{Z}^d}\widehat{V}(\xi)e^{\mathrm{i}\xi\cdot x},\\ \widehat{V}(\xi)=\frac{x_\xi}{4(1+|\xi|^2)^{m/2}},\quad x_\xi\in\left[-\tfrac12,\tfrac12\right]\subset\mathbb{R},\ m\in\mathbb{N},\ m>\tfrac12 d. \end{gathered}\tag{1-5}$$

We endow the set $\mathcal{O}:=\left[-\frac12,\frac12\right]^{\mathbb{Z}^d}$ with the standard probability measure on product spaces. This choice of the function defining the convolution potential is standard [Faou and Grébert 2013; Bambusi and Grébert 2006]; essentially one needs that the Fourier coefficients decay at a certain rate and that the function $V(x)$ depends on some free parameters x_ξ. Our main theorem is the following.

Theorem 1 (long-time existence for NLS). *Consider* (NLS) *with $d\geq 2$. There exists $\mathcal{N}\subset\mathcal{O}$ having zero Lebesgue measure such that if x_ξ in* (1-5) *is in $\mathcal{O}\setminus\mathcal{N}$, we have the following. There exists $s_0=s_0(d,m)\gg 1$ such that for any $s\geq s_0$ there are constants $c_0>0$ and $\epsilon_0>0$ such that for any $0<\epsilon\leq\epsilon_0$ we have the following. If $\|u_0\|_{H^s}<\frac14\epsilon$, there exists a unique solution of the Cauchy problem* (NLS) *such that*

$$u(t,x)\in C^0([0,T);H^s(\mathbb{T}^d)),\qquad \sup_{t\in[0,T)}\|u(t,\cdot)\|_{H^s}\leq\epsilon,\qquad T\geq c_0\epsilon^{-4}.\tag{1-6}$$

In the one-dimensional case we do not need any external parameter and we shall prove the following theorem.

Theorem 2. *Consider* (NLS) *with $V\equiv 0$ and $d=1$. There exists $s_0\gg 1$ such that for any $s\geq s_0$ there are constants $c_0>0$ and $\epsilon_0>0$ such that for any $0<\epsilon\leq\epsilon_0$ we have the following. If $\|u_0\|_{H^s}<\frac14\epsilon$, there exists a unique solution of the Cauchy problem* (NLS) *such that*

$$u(t,x)\in C^0([0,T);H^s(\mathbb{T}^d)),\qquad \sup_{t\in[0,T)}\|u(t,\cdot)\|_{H^s}\leq\epsilon,\qquad T\geq c_0\epsilon^{-4}.\tag{1-7}$$

These are, to the best of our knowledge, the firsts results of this kind for quasilinear Schrödinger equations posed on compact manifolds of dimension greater than 1.

Our main theorem regarding the problem (KG) is the following.

Theorem 3 (long-time existence for KG). *Consider* (KG) *with $d\geq 2$. There exists $\mathcal{N}\subset[1,2]$ having zero Lebesgue measure such that if $m\in[1,2]\setminus\mathcal{N}$ we have the following. There exists $s_0=s_0(d)\gg 1$ such that for any $s\geq s_0$ the following holds. For any $\delta>0$ there exists $\epsilon_0=\epsilon_0(s,m,\delta)>0$ such that for any $0<\epsilon\leq\epsilon_0$ and any initial data $(\psi_0,\psi_1)\in H^{s+1/2}(\mathbb{T}^d)\times H^{s-1/2}(\mathbb{T}^d)$ such that*

$$\|\psi_0\|_{H^{s+1/2}}+\|\psi_1\|_{H^{s-1/2}}\leq\tfrac{1}{32}\epsilon,$$

there exists a unique solution of the Cauchy problem (KG) *such that*

$$\begin{aligned}&\psi(t,x)\in C^0([0,T);H^{s+1/2}(\mathbb{T}^d))\cap C^1([0,T);H^{s-1/2}(\mathbb{T}^d)),\\&\sup_{t\in[0,T)}(\|\psi(t,\cdot)\|_{H^{s+1/2}}+\|\partial_t\psi(t,\cdot)\|_{H^{s-1/2}})\leq\epsilon,\quad T\geq\epsilon^{-\mathtt{a}+\delta},\end{aligned}\tag{1-8}$$

where $\mathtt{a}=3$ *if* $d=2$ *and* $\mathtt{a}=\frac{8}{3}$ *if* $d\geq 3$.

The time of existence in (1-8) is intimately connected with the lower bounds on the four waves interactions given in Section 2B. More precisely the time of existence is larger then $\epsilon^{-2-2/\beta}$ with β given in Proposition 2.2. This is the reason for the difference between the result in $d=2$ (where $\beta=2^+$) and $d\geq 3$ (where $\beta=3^+$). We do not know if this result is sharp; this is an open problem. Despite this fact, Theorem 2 improves the general result in [Delort 2009; Fang and Zhang 2010] in the particular case of cubic nonlinearities in the following sense. First of all we can consider more general equations containing derivatives in the nonlinearity (with “small” quasilinear term). Furthermore, adapting our proof to the semilinear case (i.e., when $f=0$ in (KG) and (1-1) and G in (1-2) does not depend on y_1), we obtain the better time of existence $\epsilon^{-10/3^-}$ for any $d\geq 4$. Indeed, in this case, the time of existence is $\epsilon^{-2-4/\beta}$ with β as above. This is the content of the next theorem.

Theorem 4. *Consider* (KG) *with* $f=0$ *and* g *independent of* y_1. *Then the result of Theorem 3 holds true*, *replacing* $\mathtt{a}=3$ *and* $\mathtt{a}=\frac{8}{3}$ *with* $\mathtt{a}=4$ *and* $\mathtt{a}=\frac{10}{3}$ *respectively.*

Comments on the results. We begin by discussing the (NLS) case. Our method covers also more general cubic terms. For instance we could replace the term $|u|^2u$ with $g(|u|^2)u$, where $g(\cdot)$ is any analytic function vanishing at the origin and having a primitive $G'=g$. We preferred not to write the paper in the most general case since the nonlinearity $|u|^2u$ is a good representative for the aforementioned class and allows us to avoid complicating the notation further. We also remark that we consider a class of potentials V more general than the one we used in [Feola and Iandoli 2020; 2021] and more similar to the one used in [Bambusi and Grébert 2006] in a semilinear context.

Secondly, we remark that, beside the mathematical interest, it would be very interesting, from a physical point of view, to be able to deal with the case $h(\tau)\sim\tau$ with $\tau\sim 0$. Indeed, for instance, if we choose $h(\tau)=\sqrt{1+\tau}-1$, the respective equation (NLS) models the self-channeling of a high-power, ultra-short laser pulse in matter; see [Borovskiĭ and Galkin 1993]. Unfortunately we need in our estimates $h(\tau)\sim\tau^{1+\sigma}$ with $\sigma>0$. More precisely we need the purely quasilinear part of the equation $[\Delta(h(|u|^2))]h'(|u|^2)u$ to be *smaller* ($O(\epsilon^{3+4\sigma})$, $\epsilon\ll 1$) than the semilinear one ($O(\epsilon^3)$). At present we are not able to perform a normal form analysis which is able to reduce the size of the purely quasilinear part. Whence, if such a quasilinear term were $O(\epsilon^3)$, then the time of existence we are able to obtain would not be better than $O(\epsilon^{-2})$. Since h has to be smooth, this leads to $h(\tau)\sim\tau^2,\ \tau\sim 0$.

Also in the (KG) case we are not able to deal with the interesting case of cubic quasilinear term. This is the reason why we require that the nonlinearity f in (1-1) has a zero of order at least 4 at the origin.

We introduce the following notation: given $j_1,\dots,j_p\in\mathbb{R}^+,\ p\geq 2$, we define

$$\max\nolimits_i\{j_1,\dots,j_p\}=i\text{-th largest among } j_1,\dots,j_p.\tag{1-9}$$

We use normal forms (the same strategy is used for (NLS) as well) and therefore small divisors problems arise. The small divisors, coming from the four waves interaction, are of the form

$$\Lambda_{\mathrm{KG}}(\xi-\eta-\zeta)-\Lambda_{\mathrm{KG}}(\eta)+\Lambda_{\mathrm{KG}}(\zeta)-\Lambda_{\mathrm{KG}}(\xi), \tag{1-10}$$

with Λ_{KG} defined in (1-4). In this case we prove the lower bound (see (1-9))

$$\begin{aligned}|\Lambda_{\mathrm{KG}}(\xi-\eta-\zeta)-\Lambda_{\mathrm{KG}}(\eta)+\Lambda_{\mathrm{KG}}(\zeta)-\Lambda_{\mathrm{KG}}(\xi)|&\\ \gtrsim \max{}_2\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}^{-N_0}\max\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}^{-\beta}&\end{aligned} \tag{1-11}$$

for almost any value of the mass m in the interval $[1,2]$ and where β is any real number in the open interval $(3,4)$. The second factor in the right-hand side of the above inequality represents a loss of derivatives when dividing by the quantity (1-10) which may be transformed in a loss of length of the lifespan through partition of frequencies. This is an extra difficulty, compared with the (NLS) case (for which lower bounds without loss have been proved in [Faou and Grébert 2013]), which makes the problem challenging already in a semilinear setting. The estimate (1-11) with $\beta\in(3,4)$ has been already obtained in [Fang and Zhang 2010]. We provide here a different and simpler proof, in the particular case of four waves interaction, which does not use the theory of subanalytic functions. We also quote Bernier, Faou and Grébert [Bernier et al. 2020] who use a control of the small divisors involving only the largest index (and not $\max_2$ as in (1-11)). They obtained, in the semilinear case, the control of the Sobolev norm for a time $T\sim\epsilon^{-\mathtt{a}}$, with $\mathtt{a}$ arbitrarily large, but assuming that the initial datum satisfies $\|\psi_0\|_{H^{s'+1/2}}+\|\psi_1\|_{H^{s'-1/2}}<c_0\epsilon$ for some $s'\equiv s'(\mathtt{a})>s$, i.e., allowing a loss of regularity.

Ideas of the proof. In our proof we shall use a *quasilinear normal forms/modified energies* approach; this seems to be the only successful one in order to improve the time of existence implied by the local theory. We recall, indeed, that on $\mathbb{T}^d$ the dispersive character of the solutions is absent. Moreover, the lack of conservation laws and the quasilinear nature of the equation prevent the use of *semilinear* techniques as done by Bambusi and Grébert [2006] and Bambusi, Delort, Grébert and Szeftel [Bambusi et al. 2007].

The most important feature of (NLS) and (KG), for our purposes, is their Hamiltonian structure. This property guarantees some key cancellations in the *energy estimates* that will be explained later on in this introduction.

Equation (NLS) may be indeed rewritten as

$$\partial_t u=-\mathrm{i}\nabla_{\bar u}\mathscr{H}_{\mathrm{NLS}}(u,\bar u)=\mathrm{i}(\Delta u-V*u-p(u)),$$

where $\nabla_{\bar u}:=\frac{1}{2}(\nabla_{\mathrm{Re}(u)}+\mathrm{i}\nabla_{\mathrm{Im}(u)})$, ∇ denotes the L^2-gradient, and the Hamiltonian function $\mathscr{H}_{\mathrm{NLS}}$ and the nonlinearity p are

$$\begin{gathered}\mathscr{H}_{\mathrm{NLS}}(u,\bar u):=\int_{\mathbb{T}^d}|\nabla u|^2+(V*u)\bar u+P(u,\nabla u)\,dx,\\ P(u,\nabla u):=\tfrac{1}{2}(|\nabla(h(|u|^2))|^2+|u|^4),\qquad p(u):=(\partial_{\bar u}P)(u,\nabla u)-\sum_{j=1}^{d}\partial_{x_j}(\partial_{\bar u_{x_j}}P)(u,\nabla u).\end{gathered} \tag{1-12}$$

Equation (KG) is Hamiltonian as well. Thanks to (1-1), (1-2) we have that (KG) can be written as

$$\begin{cases} \partial_t \psi = \partial_\phi \mathscr{H}_{\mathrm{KG}}(\psi,\phi) = \phi, \\ \partial_t \phi = -\partial_\psi \mathscr{H}_{\mathrm{KG}}(\psi,\phi) = -\Lambda_{\mathrm{KG}}^2 \psi - f(\psi) - g(\psi), \end{cases} \tag{1-13}$$

where $\mathscr{H}_{\mathrm{KG}}(\psi,\phi)$ is the Hamiltonian

$$\mathscr{H}_{\mathrm{KG}}(\psi,\phi) = \int_{\mathbb{T}^d} \tfrac{1}{2}\phi^2 + \tfrac{1}{2}(\Lambda_{\mathrm{KG}}^2\psi)\psi + F(\psi,\nabla\psi) + G(\psi,\Lambda_{\mathrm{KG}}^{1/2}\psi)\,dx. \tag{1-14}$$

We describe below our strategy in the case of the (NLS) equation. The strategy for (KG) is similar.

In [Feola and Iandoli 2022] we proved an energy estimate, without any assumption of smallness on the initial condition, for a more general class of equations. This energy estimate, on (NLS) with small initial datum, would read

$$E(t) - E(0) \lesssim \int_0^t \|u(\tau,\cdot)\|_{H^s}^2 E(\tau)\,d\tau, \tag{1-15}$$

where $E(t) \sim \|u(t,\cdot)\|_{H^s}^2$. An estimate of this kind implies, by a standard bootstrap argument, that the lifespan of the solutions is of order at least $O(\epsilon^{-2})$, where ϵ is the size of the initial condition. To increase the time to $O(\epsilon^{-4})$ one would like to show the improved inequality

$$E(t) - E(0) \lesssim \int_0^t \|u(\tau,\cdot)\|_{H^s}^4 E(\tau)\,d\tau. \tag{1-16}$$

Our main goal is to obtain such an estimate.

Paralinearization of (NLS). The first step is the paralinearization, à la [Bony 1981], of the equation as a system of the variables $(u,\bar{u})$; see Proposition 4.2. We rewrite (NLS) as a system of the form (compare with (4-12))

$$\partial_t U = -\mathrm{i}E\big((-\Delta + V*)U + \mathscr{A}_2(U)U + \mathscr{A}_1(U)U\big) + X_{H_4}(U) + R(U), \quad E := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad U := \begin{bmatrix} u \\ \bar{u} \end{bmatrix},$$

where $\mathscr{A}_2(U)$ is a 2×2 self-adjoint matrix of *paradifferential operators* of order 2 (see (4-11), (4-10)), $\mathscr{A}_1(U)$ is a self-adjoint, diagonal matrix of paradifferential operators of order 1 (see (4-12), (4-10)). This algebraic configuration of the matrices (in particular the fact that $\mathscr{A}_1(U)$ is diagonal) is a consequence of the Hamiltonian structure of the equation. The summand X_{H_4} is the cubic term (coming from the paralinearization of $|u|^2u$, see (4-13)) and $\|R(U)\|_{H^s}$ is bounded from above by $\|U\|_{H^s}^7$ for s large enough.

Both the matrices $\mathscr{A}_2(U)$ and $\mathscr{A}_1(U)$ vanish when U goes to 0. Since we assume that the function h, appearing in (NLS), vanishes quadratically at zero, as a consequence of (4-10), we have

$$\|\mathscr{A}_2(U)\|_{\mathscr{L}(H^s;H^{s-2})}, \|\mathscr{A}_1(U)\|_{\mathscr{L}(H^s;H^{s-1})} \lesssim \|U\|_{H^s}^6,$$

where by $\mathscr{L}(X;Y)$ we denote the space of linear operators from X to Y. We also remark that the summand X_{H_4} is a *Hamiltonian vector field* with Hamiltonian function $H_4(u) = \int_{\mathbb{T}^d} |u|^4\,dx$.

Diagonalization of the second-order operator. The matrix of paradifferential operators $\mathscr{A}_2(U)$ is not diagonal; therefore the first step, in order to be able to get at least the weak estimate (1-15), is to diagonalize the system at the maximum order. This is possible since, because of the smallness assumption, the operator $E(-\Delta+\mathscr{A}_2(U))$ is locally elliptic. In Section 6A1 we introduce a new unknown $W=\Phi_{\mathrm{NLS}}(U)U$, where $\Phi_{\mathrm{NLS}}(U)$ is a parametrix built from the matrix of the eigenvectors of $E(-\Delta+\mathscr{A}_2(U))$; see (6-4), (6-2). The system in the new coordinates reads

$$\partial_t W=-\mathrm{i}E\big((-\Delta+V*)U+\mathscr{A}_2^{(1)}(U)W+\mathscr{A}_1^{(1)}(U)W\big)+X_{H_4}(W)+R^{(1)}(U),$$

where both $\mathscr{A}_2^{(1)}(U)$, $\mathscr{A}_1^{(1)}(U)$ are diagonal, see (6-11), and where $\|R^{(1)}(U)\|_{H^s}\lesssim\|U\|_{H^s}^7$ for s large enough. We note also that the cubic vector field X_{H_4} remains the same because the map $\Phi_{\mathrm{NLS}}(U)$ is equal to the identity plus a term vanishing at order 6 at zero; see (6-5).

Diagonalization of the cubic vector field. In the second step, in Section 6A2, we diagonalize the cubic vector field X_{H_4}. It is fundamental for our purposes to preserve the Hamiltonian structure of this cubic vector field in this diagonalization procedure. In view of this we perform a (approximately) *symplectic* change of coordinates generated from the Hamiltonian in (5-3) and (5-2) (note that this is not the case for the diagonalization at order 2). Actually the simplecticity of this change of coordinates is one of the most delicate points in our paper. The entire Section 5 is devoted to this. This diagonalization is implemented in order to simplify a *low-high* frequencies analysis. More precisely we prove that the cubic vector field may be conjugated to a diagonal one modulo a smoothing remainder. The diagonal part shall cancel out in the energy estimate due to a symmetrization argument based on its Hamiltonian character. As a consequence the time of existence shall be completely determined by the smoothing reminder. Since this remainder is smoothing, the contribution coming from high frequencies is already “small”; therefore the normal form analysis involves only the *low modes*. This will be explained later on in this introduction.

We explain the result of this diagonalization. We define a new variable $Z=\Phi_{\mathscr{B}_{\mathrm{NLS}}}(W)$, see (6-20), and we obtain the new diagonal system (compare with (6-22))

$$\partial_t Z=-\mathrm{i}E\big((-\Delta+V*)Z+\mathscr{A}_2^{(1)}(U)Z+\mathscr{A}_1^{(1)}(U)Z\big)+X_{\mathrm{H}_4}(Z)+R_5^{(2)}(U),\tag{1-17}$$

where the new vector field $X_{\mathrm{H}_4}(Z)$ is still Hamiltonian, with Hamiltonian function defined in (6-25), and it is equal to a skew-selfadjoint and diagonal matrix of bounded paradifferential operators modulo smoothing reminders; see (6-23). Here $R_5^{(2)}(U)$ satisfies the quintic estimates (6-24).

Introduction of the energy norm. Once we achieve the diagonalization of the system, we introduce an *energy norm* which is equivalent to the Sobolev one. Assume for simplicity $s=2n$, with n a natural number. Thanks to the smallness condition on the initial datum, we prove in Section 7A1 that

$$\|(-\Delta\mathbb{1}+\mathscr{A}_2(U)+\mathscr{A}_1(U))^{s/2}f\|_{L^2}\sim\|f\|_{H^s}$$

for any function f in $H^s(\mathbb{T}^d)$. Therefore by setting[1]

$$Z_n:=[E(-\Delta\mathbb{1}+\mathscr{A}_2(U)+\mathscr{A}_1(U))]^{s/2}Z,$$

[1]To be precise, the definition of $Z_n=(z_n,\bar{z}_n)$ in 7A1 is slightly different than the one presented here, but they coincide modulo smoothing corrections. For simplicity of notation, and in order to avoid technicalities, in this introduction we presented it in this way.

we are reduced to studying the L^2 norm of the function Z_n. This is done in Lemma 7.2. Since the system is now diagonalized, we write the scalar equation, see Lemma 7.3, solved by z_n

$$\partial_t z_n = -\mathrm{i} T_{\mathscr{L}} z_n - \mathrm{i} V * z_n - \Delta^n X^{+}_{\mathtt{H}_4}(Z) + R_n(U),$$

where we denote by $T_{\mathscr{L}}$ the element on the diagonal of the self-adjoint operator $-\Delta \mathbb{1} + \mathscr{A}_2(U) + \mathscr{A}_1(U)$; see (7-1), (3-6). $X^{+}_{\mathtt{H}_4}(Z)$ is the first component of the Hamiltonian vector field $X_{\mathtt{H}_4}(Z)$ and $R_n(U)$ is a bounded remainder satisfying the quintic estimate (7-12).

Cancellations and normal forms. At this point, still in Lemma 7.3, we split the Hamiltonian vector field $X^{+}_{\mathtt{H}_4} = X^{+,\mathrm{res}}_{\mathtt{H}_4} + X^{+,\perp}_{\mathtt{H}_4}$, where $X^{+,\mathrm{res}}_{\mathtt{H}_4}$ is the *resonant* part; see (3-84) and (3-83). The first important fact, which is an effect of the Hamiltonian- and Gauge-preserving structure, is that the resonant term $\Delta^n X^{+,\mathrm{res}}_{\mathtt{H}_4}$ does not give any contribution to the energy estimates. This key cancellation may be interpreted as a consequence of the fact that the *super actions*

$$I_p := \sum_{j \in \mathbb{Z}^d, |j| = p} |\hat{z}(j)|^2, \quad p \in \mathbb{N}, \; Z := \begin{bmatrix} z \\ \bar{z} \end{bmatrix},$$

where $\hat{z}$ is defined in (3-1), are prime integrals of the resonant Hamiltonian vector field $X^{+,\mathrm{res}}_{\mathtt{H}_4}(Z)$ in the spirit[2] of [Faou et al. 2013]. This is the content of Lemma 7.4, more specifically (7-16).

We are left with the study of the term $-\Delta^n X^{+,\perp}_{\mathtt{H}_4}$. In Lemma 7.3 we prove $-\Delta^n X^{+,\perp}_{\mathtt{H}_4} = B^{(1)}_n(Z) + B^{(2)}_n(Z)$, where $B^{(1)}_n(Z)$ does not contribute to energy estimates and $B^{(2)}_n(Z)$ is smoothing, gaining one space derivative; see (7-11) and Lemma 3.7. The cancellation for $B^{(1)}_n(Z)$ is again a consequence of the Hamiltonian structure and it is proven in Lemma 7.4, more specifically (7-17).

Summarizing we obtain the energy estimate (see (3-3))

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|z_n(t)\|^2_{L^2} = \mathrm{Re}(-\mathrm{i} T_{\mathscr{L}} z_n, z_n)_{L^2} + \mathrm{Re}(-\mathrm{i} V * z_n, z_n)_{L^2} \quad & (1\text{-}18)\\
+ \mathrm{Re}(R_n(U), z_n)_{L^2} \quad & (1\text{-}19)\\
+ \mathrm{Re}(-\Delta^n X^{+,\mathrm{res}}_{\mathtt{H}_4}(Z), z_n)_{L^2} \quad & (1\text{-}20)\\
+ \mathrm{Re}(B^{(1)}_n(Z), z_n)_{L^2} \quad & (1\text{-}21)\\
+ \mathrm{Re}(B^{(2)}_n(Z), z_n)_{L^2}. \quad & (1\text{-}22)
\end{aligned}$$

The right-hand side in (1-18) equals zero because $\mathrm{i} T_{\mathscr{L}}$ is skew-self-adjoint and the Fourier coefficients of V in (1-5) are real-valued. The term (1-19) is bounded from above by $\|z_n\|_{L^2}\|U\|^5_{H^s}$; (1-20) equals zero thanks to (7-16); the summand (1-21) equals zero as well because of (7-17). Setting $E(t) = \|z_n(t)\|^2_{L^2}$, the only term which is still not good in order to obtain an estimate of the form (1-16) is (1-22).

In order to improve the time of existence we need to reduce the size of this new term (1-22) by means of *normal forms/integration by parts*. Our aim is to prove that

$$\int_0^t \mathrm{Re}(B^{(2)}_n(Z), z_n)_{L^2}(\sigma)\, d\sigma \lesssim \epsilon^2 \tag{1-23}$$

[2] More generally, this cancellation can be viewed as a consequence of the commutation of the linear flow with the resonant part of the nonlinear perturbation which is a key of the Birkhoff normal form theory; see for instance [Grébert 2007].

as long as $t \lesssim \epsilon^{-4}$ and $\|z_n\|_{L^2} \lesssim \epsilon$. The thesis follows from this fact by using a classical bootstrap argument. Let us set $\mathscr{B}_{\mathrm{NLS}}(\sigma) := \mathrm{Re}(B_n^{(2)}(Z), z_n)_{L^2}(\sigma)$. The term $\mathscr{B}_{\mathrm{NLS}}$ may be expressed as (see Proposition 7.5)

$$\mathscr{B}_{\mathrm{NLS}} \sim \sum_{\xi,\eta,\zeta\in\mathscr{R}^c} \langle\xi\rangle^{2n} b(\xi,\eta,\zeta)\hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}(\zeta)\hat{\bar{z}}(-\xi); \tag{1-24}$$

the sum is restricted to the set of *nonresonant indexes*, see (3-83), and the coefficients satisfy

$$|b(\xi,\eta,\zeta)| \lesssim \frac{\langle\xi\rangle^{2n}}{\max_1(\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle)},$$

where the constant depends on $\max_2(\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle)$ and where we have defined the Japanese bracket $\langle\xi\rangle := \sqrt{1+|\xi|^2}$ for $\xi\in\mathbb{R}^d$. We fix $N\in\mathbb{R}^+$ and we let $\mathscr{B}_{\mathrm{NLS}} := \mathscr{B}_{\mathrm{NLS},\le N} + \mathscr{B}_{\mathrm{NLS},>N}$, where $\mathscr{B}_{\mathrm{NLS},\le N}$ is as in (1-24) with the sum restricted to those indexes such that $\max_1(\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle)\le N$. It is easy to show (see Lemma 7.7) that $\int_0^t \|\mathscr{B}_{\mathrm{NLS},>N}(\sigma)\|_{H^s}\,d\sigma \lesssim tN^{-1}\|z\|_{H^s}^4$. This is due to the fact that the coefficients $b(\xi,\eta,\zeta)$ are decaying. Let us analyze the contribution given by $\mathscr{B}_{\mathrm{NLS},\le N}$.

We define the operator Λ_{NLS} as the Fourier multiplier acting on periodic functions as

$$\Lambda_{\mathrm{NLS}} e^{\mathrm{i}\xi\cdot x} = \Lambda_{\mathrm{NLS}}(\xi) e^{\mathrm{i}\xi\cdot x}, \quad \Lambda_{\mathrm{NLS}}\in\mathbb{R},\ \Lambda_{\mathrm{NLS}}(\xi) := |\xi|^2 + \widehat{V}(\xi),\ \xi\in\mathbb{Z}^d, \tag{1-25}$$

where $\widehat{V}(\xi)$ are the *real* Fourier coefficients of the convolution potential $V(x)$ given in (1-5). Recalling (1-17), we have

$$\partial_t \hat{z}(\xi) = -\mathrm{i}\Lambda_{\mathrm{NLS}}(\xi)\hat{z}(\xi) + \widehat{Q}(\xi),$$

where $Q := -\mathrm{i}T_\Sigma z + X^+_{\mathtt{H}_4}(z) + R_5^{(2)}$, with T_Σ a paradifferential operator (see (3-6)) with symbol Σ, which is real, of order 2 and homogeneity 6 in z, and $R_5^{(2)}$ is a quintic reminder. We set $\hat{g}(\xi) := e^{\mathrm{i}t\Lambda_{\mathrm{NLS}}(\xi)}\hat{z}(\xi)$ and we obtain

$$\begin{aligned}\int_0^t & \mathscr{B}_{\mathrm{NLS},\le N}(\sigma)\,d\sigma \\ &\sim \int_0^t \sum_{\xi,\eta,\zeta\in\mathscr{R}^c} \mathtt{1}_{\{\max\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}\le N\}} b(\xi,\eta,\zeta) e^{-\mathrm{i}\sigma\omega_{\mathrm{NLS}}(\xi,\eta,\zeta)} \hat{g}(\xi-\eta-\zeta)\hat{\bar{g}}(\eta)\hat{g}(\zeta)\hat{\bar{g}}(-\xi)\langle\xi\rangle^{2n}\,d\sigma,\end{aligned}$$

with ω_{NLS} defined in (2-1). Integrating by parts in σ, we obtain

$$\begin{aligned}\int_0^t \mathscr{B}_{\mathrm{NLS},\le N}(\sigma)\,d\sigma \sim &\int_0^t (\mathscr{T}_<[z,\bar{z},z], T_{\langle\xi\rangle^{2n}}(\partial_t+\mathrm{i}\Lambda_{\mathrm{NLS}})z)_{L^2}\,d\sigma \\ &+ \int_0^t (\mathscr{T}_<[(\partial_t+\mathrm{i}\Lambda_{\mathrm{NLS}})z,\bar{z},z], T_{\langle\xi\rangle^{2n}}z)_{L^2}\,d\sigma \\ &+ \int_0^t (\mathscr{T}_<[z,\bar{z},(\partial_t+\mathrm{i}\Lambda_{\mathrm{NLS}})z], T_{\langle\xi\rangle^{2n}}z)_{L^2}\,d\sigma \\ &+ \int_0^t (\mathscr{T}_<[z,\overline{(\partial_t+\mathrm{i}\Lambda_{\mathrm{NLS}})z},z], T_{\langle\xi\rangle^{2n}}z)_{L^2}\,d\sigma + O(\|u\|_{H^s}^4),\end{aligned} \tag{1-26}$$

where $\mathscr{T}_<(z_1,z_2,z_3)$ is the multilinear form whose Fourier coefficient is

$$\widehat{\mathscr{T}_<}(\xi) = \frac{1}{(2\pi)^d}\sum_{\eta,\zeta\in\mathbb{Z}^d} \mathtt{t}_<(\xi,\eta,\zeta)\hat{z}_1(\xi-\eta-\zeta)\hat{z}_2(\eta)\hat{z}_3(\zeta), \quad \mathtt{t}_<(\xi,\eta,\zeta) = \frac{-1}{\mathrm{i}\omega_{\mathrm{NLS}}(\xi,\eta,\zeta)} b(\xi,\eta,\zeta).$$

The denominators $\omega_{\rm NLS}$ are never dangerous since we have very good lower bounds on them; see Proposition 2.1 (see also Lemma 7.7). Let us consider, for instance, the first term in the right-hand side of (1-26). We have

$$\begin{aligned}&\int_0^t (\mathscr{T}_<[z,\bar z,z], T_{\langle\xi\rangle^{2n}}(\partial_t+\mathrm{i}\Lambda_{\rm NLS})z)_{L^2}(\sigma)\,d\sigma\\&=\int_0^t (T_{\langle\xi\rangle^2}\mathscr{T}_<[z,\bar z,z], -T_{\langle\xi\rangle^{2n-2}}\mathrm{i}T_\Sigma z)_{L^2}(\sigma)\,d\sigma+\int_0^t \big(\mathscr{T}_<[z,\bar z,z], T_{\langle\xi\rangle^{2n}}(X^+_{\mathrm{H}^{(4)}_{\rm NLS}}(Z)+R_5^{(2,+)}(U))\big)_{L^2}(\sigma)\,d\sigma.\end{aligned}$$

The first term may be estimated by the Cauchy–Schwarz inequality obtaining

$$\int_0^t \|\mathscr{T}_<(z,\bar z,z)\|_{H^2}(\sigma)\|T_\Sigma z\|_{H^{s-2}}(\sigma)\,d\sigma. \tag{1-27}$$

Since Σ is a symbol of order 2 and homogeneity 6, the second factor is bounded from above by ϵ^6 as soon as $\|z(\sigma)\|_{H^s}\lesssim\epsilon$. Since $\mathscr{T}_<$ is supported on frequencies lower than N, the $\langle\xi\rangle^2$ symbol of H^2 norm, multiplied by the coefficients $b(\xi,\eta,\zeta)$ of the first term in (1-27), provides a factor N (see Lemma 7.7 for details); since it has homogeneity 4, we have also a factor ϵ^4 as soon as $\|z(\sigma)\|_{H^s}\lesssim\epsilon$. We eventually bound (1-27) by $tN\epsilon^{10}$. Analogously, the second term in (1-27) may be bounded from above by $t\epsilon^6$.

Recalling the contribution given by $\mathscr{B}_{{\rm NLS},>N}$, we can bound $\int_0^t\mathscr{B}_{\rm NLS}(\sigma)\,d\sigma$ from above by $t(\epsilon^4N^{-1}+\epsilon^{10}N+\epsilon^6)+\epsilon^4$. Choosing $N=\epsilon^{-2}$ we immediately note that the last quantity stays of size ϵ^2 as soon as $t\lesssim\epsilon^{-4}$.

As said before the strategy for (KG) is similar except for the control of the small divisors (1-11).

We summarize the plan concerning (KG) focusing on the main differences with respect to (NLS). In Section 4B we paralinearize the equation obtaining, after passing to the complex variables (4-24), the system of equations of order 1 (4-44). In Section 6B we diagonalize the system: the operator of order 1 is treated in Proposition 6.11 and that of order zero in Proposition 6.13. As done for (NLS), we diagonalize the operator of order zero paying attention to preserve its Hamiltonian structure. We consider the function z solving (6-48) and we define the new variable $z_n:=\langle D\rangle^n z$, where $\langle D\rangle$ is the Fourier multiplier having symbol $\langle\xi\rangle$. We want to bound the L^2-norm of the variable z_n, which solves (7-41). The evolution of the L^2-norm is studied in Proposition 7.10. From this proposition we understand that, in order to improve the energy estimates, we need to perform a normal form on the nonresonant term $\mathscr{B}$ in (7-55), which has coefficients decaying as in (7-56). We proceed as done in the (NLS) case. We fix $N\in\mathbb{R}^+$ and we split $\mathscr{B}$ in two pieces, one supported for frequencies such that $\max_1\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}\le N$ and the other for $\max_1\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}>N$. The contribution to the energy estimate of the second one is $tN^{-1}\epsilon^4$. Again in this point we exploit the smoothing property in (7-55). We focus on the part of $\mathscr{B}$ coming from small frequencies. We perform in Proposition 7.12 an integration by parts in the same spirit as done in the (NLS) case; see (7-74). When integrating by parts, the small denominators $\omega_{\rm KG}(\xi,\eta,\zeta)$ appear. In this case we do not have nice bounds as in the (NLS) case, indeed we only know that $|\omega_{\rm KG}(\xi,\eta,\zeta)|\gtrsim\max_1\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}^{-\beta}$, where β is bigger than 3 in dimension $d\ge3$ and it is bigger than 2 in dimension $d=2$. Hence such divisors give an extra factor N^β in the energy estimates (recall that we are dealing with the case of small frequencies $\le N$). After the integration by parts one has to use (7-39). The term $\tilde{a}_2^+(x,\xi)\Lambda_{\mathrm{KG}}(\xi)$ therein has homogeneity 3 and order 1, so that its contribution to the energy estimates in (7-75) is $tN^\beta\epsilon^7$. Indeed the unboundedness of Λ_{KG} is compensated for by the coefficients of $\mathscr{B}$, which gain one derivative. The vector field $X^+_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z)$ has homogeneity 3 and has no loss of derivatives, so that its contribution to (7-76) is $tN^{\beta-1}\epsilon^6$ (the "-1" comes from the coefficients of $\mathscr{B}$). The contribution of the remainder in (7-39) is negligible. We have one last term which is the one coming from the boundary term of the integration by parts which is bounded by $N^{\beta-1}\epsilon^4$. Summarizing we have obtained (compare with (7-63))

$$\left|\int_0^t \mathscr{B}(\sigma)\,d\sigma\right| \lesssim t(\epsilon^7N^\beta+\epsilon^6N^{\beta-1}+\epsilon^4N^{-1})+\epsilon^4N^{\beta-1},$$

where the term $\epsilon^4N^{-1}t$ is coming from the high frequencies of $\mathscr{B}$. Choosing $N:=\epsilon^{-2/\beta}$ we note that the right-hand side of the above inequality is controlled by ϵ^2 as soon as $t\lesssim\epsilon^{-2(1+1/\beta)}$, which is the time announced just after the statement of Theorem 3.

We explain the role of the parameter $\mathtt{a}$ in Theorem 4. In the semilinear case we have $f=0$ and g independent of y_1 in (KG), so there are no derivatives in the nonlinearity. When we pass to the system of order 1 in (4-44), one has $\mathscr{A}_1\equiv 0$ and that the cubic term $X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(U)$ may be decomposed as a paradifferential operator of order -1 plus a trilinear reminder whose coefficients have the better (compared to the quasilinear case (7-56)) decay $\max_1(\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle)^{-2}$ (see Remark 4.6). We perform the integration by parts as in the quasilinear case. Here we do not have the contribution coming from $\tilde{a}_2^+(x,\xi)\Lambda_{\mathrm{KG}}(\xi)$ (because this term equals zero in the semilinear case), which was ϵ^7N^β. Moreover the contribution of the cubic semilinear term is $\epsilon^6N^{\beta-2}$ (instead of $\epsilon^6N^{\beta-1}$ as before), thanks to the better decay of the coefficients in the cubic reminder. The high-frequency part is also smaller and it gives $N^{-2}\epsilon^6$, instead of $N^{-1}\epsilon^6$. One eventually obtains $\left|\int_0^t\mathscr{B}(\sigma)\,d\sigma\right|\lesssim t(\epsilon^6N^{\beta-2}+\epsilon^4N^{-2})+\epsilon^4N^{\beta-2}$. If one chooses $N=\epsilon^{-2/\beta}$ one can bound the previous quantity as soon as $t\lesssim\epsilon^{-(2+4/\beta)}$, which means $t\lesssim\epsilon^{-10/3^-}$ when $d\geq 3$ and $t\lesssim\epsilon^{-4^-}$ if $d=2$.

2. Small divisors

As pointed out in the Introduction the proofs our main theorems are based on a normal form approach. As a consequence we shall deal with small divisors problems. This section is devoted to establishing suitable lower bounds for generic (in a probabilistic way) choices of the parameters (x_ξ in (1-5) for (NLS) and m in (1-4) for (KG)), except for indices for which the small divisor is identically zero.

2A. *Nonresonance conditions for* (NLS). In the following proposition we give lower bounds for the small divisors arising from the normal form for (NLS).

Proposition 2.1. *Consider the phase $\omega_{\mathrm{NLS}}(\xi,\eta,\zeta)$ defined as*

$$\omega_{\mathrm{NLS}}(\xi,\eta,\zeta):=\Lambda_{\mathrm{NLS}}(\xi-\eta-\zeta)-\Lambda_{\mathrm{NLS}}(\eta)+\Lambda_{\mathrm{NLS}}(\zeta)-\Lambda_{\mathrm{NLS}}(\xi),\quad (\xi,\eta,\zeta)\in\mathbb{Z}^{3d},\tag{2-1}$$

where Λ_{NLS} is in (1-25) *and the potential V is in* (1-5). *We have the following*:

(i) *Let $d \geq 2$. There exists $\mathscr{N} \subset \mathscr{O}$ with zero Lebesgue measure such that, for any $(x_i)_{i\in\mathbb{Z}^d} \in \mathscr{O} \setminus \mathscr{N}$, there exist $\gamma > 0$, $N_0 := N_0(d, m) > 0$ ($m > \frac{1}{2}d$ see* (1-5)) *such that for any $(\xi, \eta, \zeta) \notin \mathscr{R}$ (see* (3-83)) *one has*

$$|\omega_{\mathrm{NLS}}(\xi, \eta, \zeta)| \geq \gamma \max_2\{|\xi - \eta - \zeta|, |\eta|, |\zeta|\}^{-N_0}. \tag{2-2}$$

(ii) *Let $d = 1$ and assume that $V \equiv 0$. Then one has $|\omega_{\mathrm{NLS}}(\xi, \eta, \zeta)| \gtrsim 1$ unless*

$$\xi = \zeta, \quad \eta = \xi - \eta - \zeta \qquad or \qquad \xi = \xi - \eta - \zeta, \quad \eta = \zeta, \quad \xi, \eta, \zeta \in \mathbb{Z}. \tag{2-3}$$

Proof. Item (i) follows by Proposition 2.8 in [Faou and Grébert 2013]. Item (ii) is classical. ☐

2B. *Nonresonance conditions for* (KG). Recall the symbol $\Lambda_{\mathrm{KG}}(j)$ in (1-4). We shall prove the following important proposition.

Proposition 2.2 (nonresonance conditions). *Consider the phase $\omega^{\vec{\sigma}}_{\mathrm{KG}}(\xi, \eta, \zeta)$ defined as*

$$\omega^{\vec{\sigma}}_{\mathrm{KG}}(\xi, \eta, \zeta) := \sigma_1\Lambda_{\mathrm{KG}}(\xi - \eta - \zeta) + \sigma_2\Lambda_{\mathrm{KG}}(\eta) + \sigma_3\Lambda_{\mathrm{KG}}(\zeta) - \Lambda_{\mathrm{KG}}(\xi), \quad (\xi, \eta, \zeta) \in \mathbb{Z}^{3d}, \tag{2-4}$$

where $\vec{\sigma} := (\sigma_1, \sigma_2, \sigma_3) \in \{\pm\}^3$ and Λ_{KG} is in (1-4). *Let $0 < \sigma \ll 1$ and set $\beta := 2 + \sigma$ if $d = 2$, and $\beta := 3 + \sigma$ if $d \geq 3$. There exists $\mathscr{C}_\beta \subset [1, 2]$ with Lebesgue measure 1 such that, for any $m \in \mathscr{C}_\beta$, there exist $\gamma > 0$, $N_0 := N_0(d, m) > 0$ such that for any $(\xi, \eta, \zeta) \notin \mathscr{R}$ (see* (3-83)) *one has*

$$|\omega^{\vec{\sigma}}_{\mathrm{KG}}(\xi, \eta, \zeta)| \geq \gamma \max_2\{|\xi - \eta - \zeta|, |\eta|, |\zeta|\}^{-N_0} \max\{|\xi - \eta - \zeta|, |\eta|, |\zeta|\}^{-\beta}. \tag{2-5}$$

The case $d = 2$ follows by Theorem 2.1.1 in [Delort 2009]; the rest of this subsection is devoted to the proof of Proposition 2.2 in the case $d \geq 3$. Throughout this subsection, in order to lighten the notation, we shall write $\Lambda_{\mathrm{KG}}(j) \rightsquigarrow \Lambda_j$ for any $j \in \mathbb{Z}^d$ and $d \geq 3$. The main ingredient is the following.

Proposition 2.3. *Let $4 > \beta > 3$. There exist $\alpha > 0$ and $\mathscr{C}_\beta \subset [1, 2]$ a set of Lebesgue measure 1 and for $m \in \mathscr{C}_\beta$ there exists $\kappa(m) > 0$ such that*

$$|\sigma_1\Lambda_{j_1} + \sigma_2\Lambda_{j_2} + \sigma_3\Lambda_{j_3} + \sigma_4\Lambda_{j_4}| \geq \frac{\kappa(m)}{|j_3|^\alpha |j_1|^\beta} \tag{2-6}$$

for all $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in \{-1, 1\}$, $j_1, j_2, j_3, j_4 \in \mathbb{Z}^d$ satisfying $|j_1| \geq |j_2| \geq |j_3| \geq |j_4|$ and $\sigma_1 j_1 + \sigma_2 j_2 + \sigma_3 j_3 + \sigma_4 j_4 = 0$, except when $\sigma_1 = \sigma_4 = -\sigma_2 = -\sigma_3$ and $|j_1| = |j_2| \geq |j_3| = |j_4|$.

The Proposition 2.3 implies Proposition 2.2. Its proof is done in three steps.

<u>Step 1</u>: control with respect to the highest index.

Lemma 2.4. *There exist $\nu > 0$ and $\mathscr{M}_\nu \subset [1, 2]$ a set of Lebesgue measure 1 and for $m \in \mathscr{M}_\nu$ there exists $\gamma(m) > 0$ such that*

$$|\sigma_1\Lambda_{j_1} + \sigma_2\Lambda_{j_2} + \sigma_3\Lambda_{j_3} + \sigma_4\Lambda_{j_4}| \geq \gamma(m)|j_1|^{-\nu} \tag{2-7}$$

for all $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in \{-1, 1\}$, $j_1, j_2, j_3, j_4 \in \mathbb{Z}^d$ satisfying $|j_1| \geq |j_2| \geq |j_3| \geq |j_4|$, except when $\sigma_1 = \sigma_4 = -\sigma_2 = -\sigma_3$ and $|j_1| = |j_2| \geq |j_3| = |j_4|$.

The proof of this lemma is standard and follows the line of Theorem 6.5 in [Bambusi 2003]; see also [Bambusi and Grébert 2006; Eliasson et al. 2016]. We briefly repeat the steps.

Let us assume that $j_1, j_2, j_3, j_4 \in \mathbb{Z}^d$ satisfy $|j_1| > |j_2| > |j_3| > |j_4|$. First of all, by reasoning as in Lemma 3.2 in [Eliasson et al. 2016], one can deduce the following.

Lemma 2.5. *Consider the matrix D whose entry at place (p, q) is given by $(d^p/dm^p)\Lambda_{j_q}$, $p, q = 1, \dots, 4$. The modulus of the determinant of D is bounded from below: one has $|\det(D)| \geq C|j_1|^{-\mu}$, where $C > 0$ and $\mu > 0$ are universal constants.*

From Lemma 3.3 in [Eliasson et al. 2016] we learn:

Lemma 2.6. *Let $u^{(1)}, \dots, u^{(4)}$ be four independent vectors in $\mathbb{R}^4$ with $\|u^{(i)}\|_{\ell^1} \leq 1$. Let $w \in \mathbb{R}^4$ be an arbitrary vector. Then there exist $i \in [1, \dots, 4]$ such that $|u^{(i)} \cdot w| \geq C\|w\|_{\ell^1} \det(u^{(1)}, \dots, u^{(4)})$.*

Let us define

$$\psi_{\mathrm{KG}}(m) = \sigma_1\Lambda_{j_1}(m) + \sigma_2\Lambda_{j_2}(m) + \sigma_3\Lambda_{j_3}(m) + \sigma_4\Lambda_{j_4}(m).$$

Combining Lemmas 2.5 and 2.6 we deduce the following.

Corollary 2.7. *For any $m \in [1, 2]$ there exists an index $i \in [1, \dots, 4]$ such that*

$$\left|\frac{d^i\psi_{\mathrm{KG}}}{dm^i}(m)\right| \geq C|j_1|^{-\mu}.$$

Now we need the following result (see Lemma B.1 in [Eliasson 2002]):

Lemma 2.8. *Let $\mathfrak{g}(x)$ be a C^{n+1}-smooth function on the segment $[1, 2]$ such that*

$$|\mathfrak{g}'|_{C^n} = \beta \quad \textit{and} \quad \max_{1\leq k\leq n}\min_x |\partial^k\mathfrak{g}(x)| = \sigma.$$

Then

$$\operatorname{meas}(\{x : |\mathfrak{g}(x)| \leq \rho\}) \leq C_n\Big(\frac{\beta}{\sigma} + 1\Big)\Big(\frac{\rho}{\sigma}\Big)^{1/n}.$$

Define

$$\mathscr{E}_j(\kappa) := \{m \in [1, 2] : |\sigma_1\Lambda_{j_1} + \sigma_2\Lambda_{j_2} + \sigma_3\Lambda_{j_3} + \sigma_4\Lambda_{j_4}| \leq \kappa|j_1|^{-\nu}\}.$$

By combining Corollary 2.7 and Lemma 2.8 we get

$$\operatorname{meas}(\mathscr{E}_j(\kappa)) \leq C|j_1|^{\mu}(\kappa|j_1|^{\mu-\nu})^{1/4} \leq C\kappa^{1/4}|j_1|^{(5\mu-\nu)/4}. \tag{2-8}$$

Define

$$\mathscr{E}(\kappa) = \bigcup_{|j_1|>|j_2|>|j_3|>|j_4|} \mathscr{E}_j(\kappa),$$

and set $\nu = 5\mu + 4(4d+1)$. Then (2-8) implies $\operatorname{meas}(\mathscr{E}(\kappa)) \leq C\kappa^{1/4}$. Then taking $m \in \bigcup_{\kappa>0}([1, 2]\setminus\mathscr{E}(\kappa))$ we obtain (2-7) for any $|j_1| > |j_2| > |j_3| > |j_4|$. Furthermore $\bigcup_{\kappa>0}([1, 2]\setminus\mathscr{E}(\kappa))$ has measure 1. Now if for instance $|j_1| = |j_2|$ then we are left with a small divisor of the type $|2\Lambda_{j_1} + \sigma_3\Lambda_{j_3} + \sigma_4\Lambda_{j_4}|$ or $|\Lambda_{j_3} + \sigma_4\Lambda_{j_4}|$, i.e., involving two or three frequencies. So following the same line we can also manage this case.

<u>Step 2</u>: control with respect to the third-highest index. In this step we show that small divisors can be controlled by a smaller power of $|j_1|$, even if it means transferring part of the weight to $|j_3|$.

Proposition 2.9. *Let $4 > \beta > 3$. There exists $\mathcal{N}_\beta \subset [1, 2]$ a set of Lebesgue measure 1 and for $m \in \mathcal{N}_\beta$ there exists $\kappa(m) > 0$ such that*

$$|\Lambda_{j_1} - \Lambda_{j_2} + \sigma_3 \Lambda_{j_3} + \sigma_4 \Lambda_{j_4}| \geq \frac{\kappa(m)}{|j_3|^{2d+6} |j_1|^\beta}$$

for all $\sigma_3, \sigma_4 \in \{-1, +1\}$, for all $j_1, j_2, j_3, j_4 \in \mathbb{Z}^d$ satisfying $|j_1| > |j_2| \geq |j_3| > |j_4|$, the momentum condition $j_1 - j_2 + \sigma_3 j_3 + \sigma_4 j_4 = 0$ and

$$|j_1| \geq J(\kappa, |j_3|) := \left(\frac{C}{\kappa}\right)^{1/(4-\beta)} |j_3|^{(2d+11)/(4-\beta)},$$

where C is a universal constant.

We begin with two elementary lemmas:

Lemma 2.10. *Let $\sigma = \pm 1$, $j, k \in \mathbb{Z}^d$, with $|j| > |k| > 0$ and $|j| \geq 8$, and $[1, 2] \ni m \mapsto \mathfrak{g}(m)$ a C^1 function satisfying $|\mathfrak{g}'(m)| \leq 1/(10|j|^3)$ for $m \in [1, 2]$. For all $\kappa > 0$ there exists $\mathcal{D} \equiv \mathcal{D}(j, k, \sigma, \kappa, \mathfrak{g}) \subset [1, 2]$ such that for $m \in \mathcal{D}$*

$$|\Lambda_j + \sigma \Lambda_k - \mathfrak{g}(m)| \geq \kappa$$

and

$$\operatorname{meas}([1, 2] \setminus \mathcal{D}) \leq 10\kappa |j|^3.$$

Proof. Let $f(m) = \Lambda_j + \sigma \Lambda_k - \mathfrak{g}(m)$. In the case $\sigma = -1$, which is the worst, we have

$$f'(m) = \frac{1}{2}\left(\frac{1}{\sqrt{|j|^2+m}} - \frac{1}{\sqrt{|k|^2+m}}\right) - \mathfrak{g}'(m) = \frac{|k|^2 - |j|^2}{2(\sqrt{|j|^2+m} + \sqrt{|k|^2+m})\sqrt{|j|^2+m}\sqrt{|k|^2+m}} - \mathfrak{g}'(m).$$

We want to estimate $|f'(m)|$ from above. By using that $4(|j|^2+2)^{3/2} \leq 5|j|^3$ for $|j| \geq 8$ we get

$$|f'(m)| \geq \frac{1}{5|j|^3} - \frac{1}{10|j|^3} \geq \frac{1}{10|j|^3}.$$

In the case $\sigma = 1$, the same bound holds true. Then we conclude by a standard argument that

$$\operatorname{meas}\{m \in [1, 2] : |f(m)| \leq \kappa\} \leq 10\kappa |j|^3. \qquad \square$$

Lemma 2.11. *Let $j, k \in \mathbb{Z}^d$, with $|j| \geq |k|$ and $|j - k| \leq |j|^{1/2}$. Then*

$$\Lambda_j - \Lambda_k = \frac{(j, j-k)}{|j|} + \mathfrak{g}(|j|, |j-k|, (j-k, j), m) + O\left(\frac{|j-k|^5}{|j|^4}\right) \tag{2-9}$$

for some explicit rational function $\mathfrak{g}$.

Furthermore

$$|\partial_m \mathfrak{g}(|j|, |j-k|, (j, j-k), m)| \leq \frac{1}{2|j|^{3/2}}, \tag{2-10}$$

$$|\mathfrak{g}(|j|, |j-k|, (j, j-k), m)| \leq \frac{3|j-k|^2}{|j|} \tag{2-11}$$

uniformly with respect to $j, k \in \mathbb{Z}^d$ with $|j| \geq |k|$, $|j-k| \leq |j|^{1/2}$ and $|j|$ large enough.

Proof. By Taylor expansion we have for $|j|$ large

$$\Lambda_j = |j|\left(1+\frac{m}{|j|^2}\right)^{1/2} = |j| + \frac{m}{2|j|} - \frac{m^2}{8|j|^3} + O\left(\frac{1}{|j|^5}\right)$$

and

$$\begin{aligned}\Lambda_k &= |j|\left(1+\frac{2(k-j,j)+|j-k|^2+m}{|j|^2}\right)^{1/2}\\ &= |j| + \frac{2(k-j,j)+|j-k|^2+m}{2|j|} - \frac{(2(k-j,j)+|j-k|^2+m)^2}{8|j|^3}\\ &\qquad + \frac{3}{48}\frac{(2(k-j,j)+|j-k|^2+m)^3}{|j|^5} - \frac{15}{16}\frac{1}{4!}\frac{(2(k-j,j)+|j-k|^2+m)^4}{|j|^7} + O\left(\frac{|j-k|^5}{|j|^4}\right),\end{aligned}$$

which leads to (2-9) where (we use that $|(k-j,j)| \le |j-k||j|$ and $|j-k| \le |j|^{1/2}$)

$$\mathfrak{g}(x,y,z,m) = \frac{-y^2}{2x} + \frac{(-2z+y^2+m)^2-m^2}{8x^3} + \frac{3}{48}\frac{8z^3-12z^2(y^2+m)}{x^5} + \frac{1}{4!}\frac{15}{16}\frac{16z^4}{x^7}. \qquad \square$$

We are now in position to prove the main result of this subsection.

Proof of Proposition 2.9. We want to control the small divisor

$$\Delta = \Lambda_{j_1} - \Lambda_{j_2} + \sigma_3\Lambda_{j_3} + \sigma_4\Lambda_{j_4}.$$

Let $\mathfrak{g}$ be the rational function introduced in Lemma 2.11. We write

$$\Delta = \sigma_3\Lambda_{j_3} + \sigma_4\Lambda_{j_4} + \frac{(j_1, j_1-j_2)}{|j_1|} + \mathfrak{g}(|j_1|, |j_1-j_2|, (j_1, j_1-j_2), m) + O\left(\frac{|j_1-j_2|^5}{|j_1|^4}\right).$$

Remember that by assumption $j_1 - j_2 + \sigma_3 j_3 + \sigma_4 j_4 = 0$ and in particular $|j_1-j_2| \le 2|j_3|$.

Fix $\gamma > 0$. Choosing

$$\kappa = \frac{\gamma}{|j_3|^{2d+6}|j_1|^\beta}$$

in Lemma 2.10 and assuming $2|j_3| \le |j_1|^{1/2}$ we have by Lemmas 2.10 and 2.11

$$|\Delta| \ge \frac{\gamma}{|j_3|^{2d+6}|j_1|^\beta} - C\frac{|j_3|^5}{|j_1|^4} \ge \frac{\gamma}{2|j_3|^{2d+6}|j_1|^\beta}$$

as soon as[3]

$$|j_1| \ge \left(\frac{C}{\gamma}\right)^{1/(4-\beta)} |j_3|^{(2d+11)/(4-\beta)} =: J(\gamma, |j_3|) \ge 5|j_3|^3$$

(where C is an universal constant) and $m \in \mathscr{D}(j_3, j_4, \sigma, \kappa, \sigma_3\,\mathfrak{g}(|j_1|, |j_1-j_2|, (j_1, j_1-j_2), \cdot))$ (the set $\mathscr{D}$ is defined in Lemma 2.10 and we set $\sigma = \sigma_3\sigma_4$). Then defining

$$\mathscr{C}(\gamma, j_3, j_4, \sigma_3, \sigma_4) := \left\{ m \in [1,2] : |\Delta| \ge \frac{\gamma}{2|j_3|^{2d+6}|j_1|^\beta} \text{ for all } (j_1, j_2) \right.$$
$$\left. \text{such that } |j_1| \ge \max(|j_2|, J(\gamma, |j_3|)),\ \ j_1 - j_2 + \sigma_3 j_3 + \sigma_4 j_4 = 0 \right\},$$

[3]Note that this estimate implies $|\partial_m \mathfrak{g}(|j_1|, [j_1-j_2|, (j_1-j_2, j_1), m)| \le 1/(2|j_1|^{3/2}) \le 1/(10|j_3|^3)$ and thus Lemma 2.10 applies.

we have

$$\mathscr{C}(\gamma, j_3, j_4, \sigma_3, \sigma_4) = \bigcap_{\mathfrak{g}} \mathscr{D}\left(j_3, j_4, \sigma, \frac{\gamma}{|j_3|^{2d+6}|j_1|^{\beta}}, \sigma_3 \mathfrak{g}(|j_1|, |j_1 - j_2|, (j_1, j_1 - j_2), \cdot)\right),$$

where the intersection is taken over all functions $\mathfrak{g}$ generated by $(j_1, j_2) \in (\mathbb{Z}^d)^2$ such that

$$|j_1| \geq \max(|j_2|, J(\gamma, |j_3|))$$

and $j_1 - j_2 + \sigma_3 j_3 + \sigma_4 j_4 = 0$. Thus by Lemma 2.10

$$\begin{aligned}&\operatorname{meas}([1, 2] \setminus \mathscr{C}(\gamma, j_3, j_4, \sigma_3, \sigma_4)) \\ &\qquad \leq \sum_{n \geq 1} \frac{10\gamma}{|j_3|^{2d+3} n^{\beta/2}} \#\{(|j_1|, |\sigma_3 j_3 + \sigma_4 j_4|, (j_1, \sigma_3 j_3 + \sigma_4 j_4)) : j_1 \in \mathbb{Z}^d,\ |j_1|^2 = n\}.\end{aligned}$$

The scalar product $(j_1, \sigma_3 j_3 + \sigma_4 j_4))$ takes only integer values smaller than $2|j_1||j_3|$. Then, since $\beta > 3$, we get

$$\operatorname{meas} \mathscr{C}(\gamma, j_3, j_4, \sigma_3, \sigma_4) \leq \frac{20\gamma}{|j_3|^{2d+2}} \sum_{n \geq 1} \frac{1}{n^{(\beta-1)/2}} \leq C_\beta \frac{\gamma}{|j_3|^{2d+2}}.$$

Then it remains to define

$$\mathscr{N}_\beta = \bigcup_{\gamma > 0} \bigcap_{\substack{(j_3, j_4) \in (\mathbb{Z}^d)^2 \\ |j_4| \leq |j_3| \\ \sigma_3, \sigma_4 \in \{-1, 1\}}} \mathscr{C}(\gamma, j_3, j_4, \sigma_3, \sigma_4)$$

to conclude the proof. □

Step 3: *Proof of Proposition 2.3.* We are now in a position to prove Proposition 2.3. Let $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in \{-1, 1\}$, $j_1, j_2, j_3, j_4 \in \mathbb{Z}^d$ satisfying $|j_1| \geq |j_2| \geq |j_3| \geq |j_4|$ and $\sigma_1 j_1 + \sigma_2 j_2 + \sigma_3 j_3 + \sigma_4 j_4 = 0$. If $\sigma_1 = \sigma_2$, then, since $|j_1| \geq |j_2| \geq |j_3| \geq |j_4|$, we conclude that the associated small divisor cannot be small except if $\sigma_1 = \sigma_2 = -\sigma_3 = -\sigma_4$. Then we have to control $|\Lambda_{j_1} + \Lambda_{j_2} - \Lambda_{j_3} - \Lambda_{j_4}|$ knowing that $|j_1| \geq |j_2| \geq |j_3| \geq |j_4|$. We first notice that if $|j_1|^2 \leq |j_3|^2 + 1$, then we can conclude using Lemma 2.4 that (2-6) is satisfied with $\alpha = \nu$ for $m \in \mathscr{M}_\nu$. On the other hand if $|j_1|^2 \geq |j_3|^2 + 1$ then

$$\Lambda_{j_1} + \Lambda_{j_2} - \Lambda_{j_3} - \Lambda_{j_4} \geq \Lambda_{j_1} - \Lambda_{j_3} \geq \frac{\Lambda_{j_1}^2 - \Lambda_{j_3}^2}{\Lambda_{j_1} + \Lambda_{j_3}} \geq \frac{1}{2\sqrt{|j_3|^2 + 2}},$$

which implies (2-6). Thus we can assume $\sigma_1 = -\sigma_2$ and we can apply Proposition 2.9, which implies the control (2-6) for $m \in \mathscr{N}_\beta$ with $\alpha = 2d + 3$ under the additional constraint $|j_1| \geq J(\gamma(m), |j_3|)$. Now if $|j_1| \leq J(\gamma(m), |j_3|)$, we can apply Lemma 2.4 to obtain that there exists $\nu > 0$ and full measure set $\mathscr{M}_\nu$ such that for $m \in \mathscr{M}_\nu \cap \mathscr{N}_\beta := \mathscr{C}_\beta$ there exists $\kappa(m) > 0$ such that

$$|\sigma_1 \Lambda_{j_1} + \sigma_2 \Lambda_{j_2} + \sigma_3 \Lambda_{j_3} + \sigma_4 \Lambda_{j_4}| \geq \frac{\kappa(m)}{|j_1|^\nu} \geq \frac{\kappa(m)}{J(\gamma(m), |j_3|)^\nu} = C \frac{\kappa(m)\gamma(m)^{4-\beta}}{|j_3|^\alpha},$$

with $\alpha = \nu(2d + 8)/(4 - \beta)$ which, of course, implies (2-6).

3. Functional setting

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

$$u(x) = \frac{1}{(2\pi)^{d/2}} \sum_{n\in\mathbb{Z}^d} \hat{u}(n) e^{\mathrm{i}n\cdot x}, \quad \hat{u}(n) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{T}^d} u(x) e^{-\mathrm{i}n\cdot x}\,dx. \tag{3-1}$$

We set $\langle j\rangle := \sqrt{1+|j|^2}$ for $j \in \mathbb{Z}^d$. We endow $H^s(\mathbb{T}^d;\mathbb{C})$ with the norm

$$\|u(\,\cdot\,)\|^2_{H^s} := \sum_{j\in\mathbb{Z}^d} \langle j\rangle^{2s} |\hat{u}(j)|^2. \tag{3-2}$$

For $U = (u_1,u_2) \in H^s(\mathbb{T}^d;\mathbb{C}^2)$ we set $\|U\|_{H^s} = \|u_1\|_{H^s} + \|u_2\|_{H^s}$. Moreover, for $r \in \mathbb{R}^+$, we denote by $B_r(H^s(\mathbb{T}^d;\mathbb{C}))$ (resp. $B_r(H^s(\mathbb{T}^d;\mathbb{C}^2))$) the ball of $H^s(\mathbb{T}^d;\mathbb{C}))$ (resp. $H^s(\mathbb{T}^d;\mathbb{C}^2))$) with radius r centered at the origin. We shall also write the norm in (3-2) as $\|u\|^2_{H^s} = (\langle D\rangle^s u, \langle D\rangle^s u)_{L^2}$, where $\langle D\rangle e^{\mathrm{i}j\cdot x} = \langle j\rangle e^{\mathrm{i}j\cdot x}$ for any $j\in\mathbb{Z}^d$, and $(\,\cdot\,,\,\cdot\,)_{L^2}$ denotes the standard complex L^2-scalar product

$$(u,v)_{L^2} := \int_{\mathbb{T}^d} u\cdot\bar{v}\,dx \quad \text{for all } u,v \in L^2(\mathbb{T}^d;\mathbb{C}). \tag{3-3}$$

Notation. We shall use the notation $A \lesssim B$ to denote $A \le CB$, where C is a positive constant depending on parameters fixed once for all, for instance d and s. We will emphasize by writing $\lesssim_q$ when the constant C depends on some other parameter q.

Basic paradifferential calculus. We follow the notation of [Feola and Iandoli 2022]. We introduce the symbols we shall use in this paper. We shall consider symbols $\mathbb{T}^d \times \mathbb{R}^d \ni (x,\xi) \to a(x,\xi)$ in the spaces $\mathcal{N}^m_s$, $m,s\in\mathbb{R}$, $s \ge 0$, defined by the norms

$$|a|_{\mathcal{N}^m_s} := \sup_{|\alpha|+|\beta|\le s} \sup_{\xi\in\mathbb{R}} \langle\xi\rangle^{-m+|\beta|} \|\partial_\xi^\beta \partial_x^\alpha a(x,\xi)\|_{L^\infty}. \tag{3-4}$$

The constant $m\in\mathbb{R}$ indicates the *order* of the symbols, while s denotes its differentiability. Let $0<\epsilon<\frac{1}{2}$ and consider a smooth function $\chi:\mathbb{R}\to[0,1]$,

$$\chi(\xi) = \begin{cases} 1 & \text{if } |\xi| \le \frac{5}{4}, \\ 0 & \text{if } |\xi| \ge \frac{8}{5} \end{cases} \quad \text{and define} \quad \chi_\epsilon(\xi) := \chi\left(\frac{|\xi|}{\epsilon}\right). \tag{3-5}$$

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

$$T_a h := \frac{1}{(2\pi)^d} \sum_{j\in\mathbb{Z}^d} e^{\mathrm{i}j\cdot x} \sum_{k\in\mathbb{Z}^d} \chi_\epsilon\left(\frac{|j-k|}{\langle j+k\rangle}\right) \hat{a}\left(j-k, \frac{j+k}{2}\right) \hat{h}(k), \tag{3-6}$$

where $\hat{a}(\eta,\xi)$ denotes the Fourier transform of $a(x,\xi)$ in the variable $x\in\mathbb{T}^d$. Moreover the definition of the operator T_a is independent of the choice of the cut-off function χ_ϵ up to smoothing terms; this will be proved later in Lemma 3.1.

Notation. Given a symbol $a(x,\xi)$, we shall also write

$$T_a[\,\cdot\,] := \mathrm{Op}^{\mathrm{BW}}(a(x,\xi))[\,\cdot\,] \tag{3-7}$$

to denote the associated paradifferential operator. In the notation B stands for Bony and W for Weyl.

We now collect some fundamental properties of paradifferential operators on tori. The results are similar to the ones given in [Feola and Iandoli 2022]. One could also look at [Berti et al. 2021c] for recent improvements.

Lemma 3.1. *The following hold*:

(i) *Let* $m_1, m_2 \in \mathbb{R}$, $s > \frac{1}{2}d$, $s \in \mathbb{N}$ *and* $a \in \mathcal{N}_s^{m_1}$, $b \in \mathcal{N}_s^{m_2}$. *One has*

$$|ab|_{\mathcal{N}_s^{m_1+m_2}} + |\{a,b\}|_{\mathcal{N}_{s-1}^{m_1+m_2-1}} \lesssim |a|_{\mathcal{N}_s^{m_1}} |b|_{\mathcal{N}_s^{m_2}}, \tag{3-8}$$

where

$$\{a,b\} := \sum_{j=1}^{d} \big((\partial_{\xi_j} a)(\partial_{x_j} b) - (\partial_{x_j} a)(\partial_{\xi_j} b)\big). \tag{3-9}$$

(ii) *Let* $s_0 > d$, $s_0 \in \mathbb{N}$, $m \in \mathbb{R}$ *and* $a \in \mathcal{N}_{s_0}^m$. *Then, for any* $s \in \mathbb{R}$, *one has*

$$\|T_a h\|_{H^{s-m}} \lesssim |a|_{\mathcal{N}_{s_0}^m} \|h\|_{H^s} \quad \textit{for all } h \in H^s(\mathbb{T}^d;\mathbb{C}). \tag{3-10}$$

(iii) *Let* $s_0 > d$, $s_0 \in \mathbb{N}$, $m \in \mathbb{R}$, $\rho \in \mathbb{N}$, *and* $a \in \mathcal{N}_{s_0+\rho}^m$. *For* $0 < \epsilon_2 \le \epsilon_1 < \frac{1}{2}$ *and any* $h \in H^s(\mathbb{T}^d;\mathbb{C})$, *we define*

$$R_a h := \frac{1}{(2\pi)^d} \sum_{j\in\mathbb{Z}^d} e^{\mathrm{i} j\cdot x} \sum_{k\in\mathbb{Z}^d} (\chi_{\epsilon_1} - \chi_{\epsilon_2})\left(\frac{|j-k|}{\langle j+k\rangle}\right) \hat{a}\left(j-k, \frac{j+k}{2}\right) \hat{h}(k), \tag{3-11}$$

where $\chi_{\epsilon_1}, \chi_{\epsilon_2}$ *are as in* (3-5). *Then one has*

$$\|R_a h\|_{H^{s+\rho-m}} \lesssim \|h\|_{H^s} |a|_{\mathcal{N}_{\rho+s_0}^m} \quad \textit{for all } h \in H^s(\mathbb{T}^d;\mathbb{C}). \tag{3-12}$$

Proof. (i) For any $|\alpha| + |\beta| \le s$ we have

$$\partial_x^\alpha \partial_\xi^\beta (a(x,\xi) b(x,\xi)) = \sum_{\substack{\alpha_1+\alpha_2=\alpha\\ \beta_1+\beta_2=\beta}} C_{\alpha,\beta} (\partial_x^{\alpha_1}\partial_\xi^{\beta_1} a)(x,\xi)(\partial_x^{\alpha_2}\partial_\xi^{\beta_2} b)(x,\xi)$$

for some combinatorial coefficients $C_{\alpha,\beta} > 0$. Then, recalling (3-4),

$$\|(\partial_x^{\alpha_1}\partial_\xi^{\beta_1} a)(x,\xi)(\partial_x^{\alpha_2}\partial_\xi^{\beta_2} b)(x,\xi)\|_{L^\infty} \lesssim_{\alpha,\beta} |a|_{\mathcal{N}_s^{m_1}} |b|_{\mathcal{N}_s^{m_2}} \langle\xi\rangle^{m_1+m_2-|\beta|}.$$

This implies (3-8) for the product ab. Inequality (3-8) for the symbol $\{a,b\}$ follows similarly using (3-9).

(ii) First of all notice that, since $a \in \mathcal{N}_{s_0}^m$, $s_0 \in \mathbb{N}$, we have (recall (3-4))

$$\|a(\,\cdot\,,\xi)\|_{H^{s_0}} \lesssim \langle\xi\rangle^m |a|_{\mathcal{N}_{s_0}^m} \quad \text{for all } \xi \in \mathbb{Z}^d,$$

which implies

$$|\hat{a}(j,\xi)| \lesssim \langle\xi\rangle^m |a|_{\mathcal{N}_{s_0}^m} \langle j\rangle^{-s_0} \quad \text{for all } j,\xi \in \mathbb{Z}^d. \tag{3-13}$$

Moreover, since $0<\epsilon<\frac{1}{2}$ we note that, for $\xi,\eta\in\mathbb{Z}^d$,

$$\chi_\epsilon\left(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\right)\neq 0 \quad\Longrightarrow\quad \begin{cases}(1-\tilde{\epsilon})|\xi|\le(1+\tilde{\epsilon})|\eta|,\\(1-\tilde{\epsilon})|\eta|\le(1+\tilde{\epsilon})|\xi|,\end{cases} \tag{3-14}$$

where $0<\tilde{\epsilon}<\frac{4}{5}$, and hence $\langle\xi+\eta\rangle\sim\langle\xi\rangle$. Therefore

$$\begin{aligned}\|T_a h\|^2_{H^{s-m}} &\lesssim \sum_{\xi\in\mathbb{Z}^d}\langle\xi\rangle^{2(s-m)}\left|\sum_{\eta\in\mathbb{Z}^d}\chi_\epsilon\left(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\right)\hat{a}\left(\xi-\eta,\frac{\xi+\eta}{2}\right)\hat{h}(\eta)\right|^2 && \text{(by (3-2))}\\ &\lesssim \sum_{\xi\in\mathbb{Z}^d}\langle\xi\rangle^{-2m}\left(\sum_{\eta\in\mathbb{Z}^d}\frac{\langle\xi\rangle^m}{\langle\xi-\eta\rangle^{s_0}}|\hat{h}(\eta)|\langle\eta\rangle^s\right)^2|a|^2_{\mathcal{N}^m_{s_0}} && \text{(by (3-13), (3-14))},\\ &\lesssim |a|^2_{\mathcal{N}^m_{s_0}}\sum_{\xi\in\mathbb{Z}^d}\left(\sum_{\eta\in\mathbb{Z}^d}\left|\hat{h}(\eta)\langle\eta\rangle^s\frac{1}{\langle\xi-\eta\rangle^{s_0}}\right|\right)^2\\ &\lesssim |a|^2_{\mathcal{N}^m_{s_0}}\|\hat{h}(\xi)\langle\xi\rangle^s\star\langle\xi\rangle^{-s_0}\|^2_{\ell^2(\mathbb{Z}^d)}\le|a|^2_{\mathcal{N}^m_{s_0}}\|\hat{h}(\xi)\langle\xi\rangle^s\|^2_{\ell^2(\mathbb{Z}^d)}\|\langle\xi\rangle^{-s_0}\|^2_{\ell^1(\mathbb{Z}^d)}\\ &\lesssim \|h\|^2_{H^s}|a|^2_{\mathcal{N}^m_{s_0}},\end{aligned} \tag{3-15}$$

where we denote by $\star$ the convolution between sequences, and in the penultimate inequality we used the Young inequality for sequences and in the last one that $\langle\xi\rangle^{-s_0}$ is in $\ell^1(\mathbb{Z}^d)$ since $s_0>d$.

(iii) Notice that the set of ξ,η such that $(\chi_{\epsilon_1}-\chi_{\epsilon_2})(|\xi-\eta|/\langle\xi+\eta\rangle)=0$ contains the set such that

$$|\xi-\eta|\ge\tfrac{8}{5}\epsilon_1\langle\xi+\eta\rangle \quad\text{or}\quad |\xi-\eta|\le\tfrac{5}{4}\epsilon_2\langle\xi+\eta\rangle.$$

Therefore $(\chi_{\epsilon_1}-\chi_{\epsilon_2})(|\xi-\eta|/\langle\xi+\eta\rangle)\neq 0$ implies

$$\tfrac{5}{4}\epsilon_2\langle\xi+\eta\rangle\le|\xi-\eta|\le\tfrac{8}{5}\epsilon_1\langle\xi+\eta\rangle. \tag{3-16}$$

For $\xi\in\mathbb{Z}^d$ we denote by $\mathscr{A}(\xi)$ the set of $\eta\in\mathbb{Z}^d$ such that (3-16) holds. Moreover (reasoning as in (3-13)), since $a\in\mathcal{N}^m_{s_0+\rho}$, we have

$$|\hat{a}(j,\xi)|\lesssim\langle\xi\rangle^m|a|_{\mathcal{N}^m_{s_0+\rho}}\langle j\rangle^{-s_0-\rho}\quad\text{for all } j,\xi\in\mathbb{Z}^d. \tag{3-17}$$

To estimate the remainder in (3-11) we reason as in (3-15). By (3-16) and setting $\rho=s-s_0$ we have

$$\begin{aligned}\|R_a h\|^2_{H^{s+\rho-m}} &\lesssim \sum_{\xi\in\mathbb{Z}^d}\langle\xi\rangle^{2(s+\rho-m)}\left|(\chi_{\epsilon_1}-\chi_{\epsilon_2})\left(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\right)\hat{a}\left(\xi-\eta,\frac{\xi+\eta}{2}\right)\hat{h}(\eta)\right|^2 && \text{(by (3-2))}\\ &\lesssim \sum_{\xi\in\mathbb{Z}^d}\langle\xi\rangle^{-2m}\left(\sum_{\eta\in\mathscr{A}(\xi)}\frac{\langle\xi-\eta\rangle^\rho\langle\xi+\eta\rangle^m}{\langle\xi-\eta\rangle^{\rho+s_0}}|\hat{h}(\eta)|\langle\eta\rangle^s\right)^2|a|^2_{\mathcal{N}^m_{s_0+\rho}} && \text{(by (3-17))}\\ &\lesssim \|\hat{h}(\xi)\langle\xi\rangle^s\star\langle\xi\rangle^{-s_0}\|^2_{\ell^2(\mathbb{Z}^d)}|a|^2_{\mathcal{N}^m_{\rho+s_0}}\\ &\lesssim \|\hat{h}(\xi)\langle\xi\rangle^s\|^2_{\ell^2(\mathbb{Z}^d)}\|\langle\xi\rangle^{-s_0}\|^2_{\ell^1(\mathbb{Z}^d)}|a|^2_{\mathcal{N}^m_{\rho+s_0}}\lesssim\|h\|^2_{H^s}|a|^2_{\mathcal{N}^m_{\rho+s_0}},\end{aligned} \tag{3-18}$$

where we denote by $\star$ the convolution between sequences, and in the penultimate inequality we used the Young inequality for sequences and in the last one we used that $\langle\xi\rangle^{-s_0}$ is in $\ell^1(\mathbb{Z}^d)$ since $s_0>d$. □

Proposition 3.2 (composition). *Fix $s_0 > d$, $s_0 \in \mathbb{N}$, and $m_1, m_2 \in \mathbb{R}$. For $a \in \mathcal{N}^{m_1}_{s_0+2}$ and $b \in \mathcal{N}^{m_2}_{s_0+2}$ we have* (*recall* (3-9))

$$T_a \circ T_b = T_{ab} + R_1(a,b), \quad T_a \circ T_b = T_{ab} + \frac{1}{2\mathrm{i}} T_{\{a,b\}} + R_2(a,b), \tag{3-19}$$

where $R_j(a,b)$ are remainders satisfying, for any $s \in \mathbb{R}$,

$$\|R_j(a,b)h\|_{H^{s-m_1-m_2+j}} \lesssim \|h\|_{H^s} |a|_{\mathcal{N}^{m_1}_{s_0+j}} |b|_{\mathcal{N}^{m_2}_{s_0+j}}. \tag{3-20}$$

Moreover, if $a, b \in H^{\rho+s_0}(\mathbb{T}^d; \mathbb{C})$ are functions (*independent of $\xi \in \mathbb{R}^n$*) *then, for all $s \in \mathbb{R}$,*

$$\|(T_a T_b - T_{ab})h\|_{H^{s+\rho}} \lesssim \|h\|_{H^s} \|a\|_{H^{\rho+s_0}} \|b\|_{H^{\rho+s_0}}. \tag{3-21}$$

Proof. We start by proving (3-21). For $\xi, \theta, \eta \in \mathbb{Z}^d$ we define

$$r_1(\xi,\theta,\eta) := \chi_\epsilon\Big(\frac{|\xi-\theta|}{\langle \xi+\theta\rangle}\Big)\chi_\epsilon\Big(\frac{|\theta-\eta|}{\langle \theta+\eta\rangle}\Big), \quad r_2(\xi,\eta) := \chi_\epsilon\Big(\frac{|\xi-\eta|}{\langle \xi+\eta\rangle}\Big). \tag{3-22}$$

Recalling (3-6) and that a, b are functions we have

$$\begin{aligned} R_0 h &:= (T_a T_b - T_{ab})h, \\ \widehat{(R_0 h)}(\xi) &= (2\pi)^{-3d/2} \sum_{\eta,\theta \in \mathbb{Z}^d} (r_1 - r_2)(\xi,\theta,\eta)\hat{a}(\xi-\theta)\hat{b}(\theta-\eta)\hat{h}(\eta). \end{aligned} \tag{3-23}$$

Let us define the sets

$$D := \{(\xi,\theta,\eta) \in \mathbb{Z}^{3d} : (r_1 - r_2)(\xi,\theta,\eta) = 0\}, \tag{3-24}$$

$$A := \Big\{(\xi,\theta,\eta) \in \mathbb{Z}^{3d} : \frac{|\xi-\theta|}{\langle \xi+\theta\rangle} \le \frac{5\epsilon}{4},\ \frac{|\xi-\eta|}{\langle \xi+\eta\rangle} \le \frac{5\epsilon}{4},\ \frac{|\theta-\eta|}{\langle \theta+\eta\rangle} \le \frac{5\epsilon}{4}\Big\}, \tag{3-25}$$

$$B := \Big\{(\xi,\theta,\eta) \in \mathbb{Z}^{3d} : \frac{|\xi-\theta|}{\langle \xi+\theta\rangle} \ge \frac{8\epsilon}{5},\ \frac{|\xi-\eta|}{\langle \xi+\eta\rangle} \ge \frac{8\epsilon}{5},\ \frac{|\theta-\eta|}{\langle \theta+\eta\rangle} \ge \frac{8\epsilon}{5}\Big\}. \tag{3-26}$$

We note that

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

Let $(\xi,\theta,\eta) \in D^c$ and assume in particular that $(\xi,\theta,\eta) \in \mathrm{Supp}(r_1) := \overline{\{(\xi,\theta,\eta) : r_1 \neq 0\}}$. Then, reasoning as in (3-14), we can note that

$$|\xi - \eta| \le \epsilon\langle \xi+\eta\rangle \quad \text{and} \quad \langle\xi\rangle \sim \langle\eta\rangle. \tag{3-27}$$

Notice also that $(\xi,\theta,\eta) \in \mathrm{Supp}(r_2)$ implies (3-27) as well. The rough idea of the proof is based on the fact that, if $(\xi,\theta,\eta) \in D^c$, then there are *at least three* equivalent frequencies among $\xi, \xi-\theta, \theta-\eta, \eta$; therefore (3-23) restricted to $(\xi,\theta,\eta) \in D^c$ is a regularizing operator. We need to estimate

$$\|R_0 h\|^2_{H^{s+\rho}} \lesssim \sum_{\xi \in \mathbb{Z}^d} \Big(\sum^{*}_{\eta,\theta} |\hat{a}(\xi-\theta)||\hat{b}(\theta-\eta)||\hat{h}(\eta)|\langle\xi\rangle^{s+\rho}\Big)^2 = I + II + III,$$

where $\sum^*_{\eta,\theta}$ denotes the sum over indexes satisfying (3-27), the term I denotes the sum on indexes satisfying also $|\xi-\theta| > c\epsilon|\xi|$, II denotes the sum on indexes satisfying also $|\eta-\theta| > c\epsilon|\eta|$ for some $0 < c \ll 1$

and III is defined by the difference. We estimate the term I. Using (3-27) and $|\xi-\theta|>c\epsilon|\xi|$ we get

$$
\begin{aligned}
I &\lesssim \sum_{\xi\in\mathbb{Z}^d}\Big(\sum_{\eta,\theta}^{*}|\hat{a}(\xi-\theta)|\,|\hat{b}(\theta-\eta)|\,|\hat{h}(\eta)|\langle\eta\rangle^{s}\langle\xi-\theta\rangle^{\rho}\Big)^2\\
&\lesssim \|\,|\hat{h}(\xi)|\langle\xi\rangle^{s}\star|\hat{a}(\xi)|\langle\xi\rangle^{\rho}\star|\hat{b}(\xi)|\,\|^2_{\ell^2(\mathbb{Z}^d)}\\
&\lesssim \|\,|\hat{h}(\xi)|\langle\xi\rangle^{s}\|^2_{\ell^2(\mathbb{Z}^d)}\|\,|\hat{a}(\xi)|\langle\xi\rangle^{\rho}\|^2_{\ell^1(\mathbb{Z}^d)}\|\,|\hat{b}(\xi)|\,\|^2_{\ell^1(\mathbb{Z}^d)}\\
&\lesssim \|h\|^2_{H^s}\|a\|^2_{H^{s_0+\rho}}\|b\|^2_{H^{s_0}},
\end{aligned}
$$

where in the last inequality we used Cauchy–Schwarz and $s_0>d>\frac{1}{2}d$.

Reasoning similarly one obtains $II\lesssim\|h\|^2_{H^s}\|a\|^2_{H^{s_0}}\|b\|^2_{H^{s_0+\rho}}$. The sum III is restricted to indexes satisfying (3-27) and $|\xi-\theta|\le c\epsilon|\xi|$, $|\eta-\theta|\le c\epsilon|\eta|$. For $c\ll 1$ small enough these restrictions imply that $(\xi,\eta,\zeta)\in A$, which is a contradiction since $(\xi,\eta,\zeta)\in D^c\subseteq A^c$.

Let us check (3-20). We prove that

$$
T_a\circ T_b=T_{ab}+\frac{1}{2\mathrm{i}}T_{\{a,b\}}+R_2(a,b),\qquad \|R_2(a,b)h\|_{H^{s-m_1-m_2+2}}\lesssim\|h\|_{H^s}|a|_{\mathcal{N}^{m_1}_{s_0+2}}|b|_{\mathcal{N}^{m_2}_{s_0+2}}. \tag{3-28}
$$

First of all we note that

$$
\widehat{(T_aT_bh)}(\xi)=\frac{1}{(\sqrt{2\pi})^{3d}}\sum_{\eta,\theta\in\mathbb{Z}^d}r_1(\xi,\theta,\eta)\hat{a}\Big(\xi-\theta,\frac{\xi+\theta}{2}\Big)\hat{b}\Big(\theta-\eta,\frac{\theta+\eta}{2}\Big)\hat{h}(\eta), \tag{3-29}
$$

$$
\widehat{(T_{ab}h)}(\xi)=\frac{1}{(\sqrt{2\pi})^{3d}}\sum_{\eta,\theta\in\mathbb{Z}^d}r_2(\xi,\eta)\hat{a}\Big(\xi-\theta,\frac{\xi+\eta}{2}\Big)\hat{b}\Big(\theta-\eta,\frac{\xi+\eta}{2}\Big)\hat{h}(\eta), \tag{3-30}
$$

$$
\begin{aligned}
\frac{1}{2\mathrm{i}}\widehat{(T_{\{a,b\}}h)}(\xi)=&\frac{1}{2\mathrm{i}(\sqrt{2\pi})^{3d}}\sum_{\eta,\theta\in\mathbb{Z}^d}r_2(\xi,\eta)\widehat{(\partial_\xi a)}\Big(\xi-\theta,\frac{\xi+\eta}{2}\Big)\cdot\widehat{(\partial_x b)}\Big(\theta-\eta,\frac{\xi+\eta}{2}\Big)\hat{h}(\eta)\\
&-\frac{1}{2\mathrm{i}(\sqrt{2\pi})^{3d}}\sum_{\eta,\theta\in\mathbb{Z}^d}r_2(\xi,\eta)\widehat{(\partial_x a)}\Big(\xi-\theta,\frac{\xi+\eta}{2}\Big)\cdot\widehat{(\partial_\xi b)}\Big(\theta-\eta,\frac{\xi+\eta}{2}\Big)\hat{h}(\eta).
\end{aligned} \tag{3-31}
$$

In the formulas above we used the notation $\partial_x=(\partial_{x_1},\dots,\partial_{x_d})$, similarly for ∂_ξ. We remark that we can substitute the cut-off function r_2 in (3-30), (3-31) with r_1 up to smoothing remainders. This follows because one can treat the cut-off function $r_1(\xi,\theta,\eta)-r_2(\xi,\eta)$ as done in the proof of (3-21). Write $\xi+\theta=\xi+\eta+(\theta-\eta)$. By Taylor expanding the symbols at $\xi+\eta$, we have

$$
\begin{aligned}
\hat{a}\Big(\xi-\theta,\frac{\xi+\theta}{2}\Big)=&\,\hat{a}\Big(\xi-\theta,\frac{\xi+\eta}{2}\Big)+\widehat{(\partial_\xi a)}\Big(\xi-\theta,\frac{\xi+\eta}{2}\Big)\cdot\frac{\theta-\eta}{2}\\
&+\frac{1}{4}\sum_{j,k=1}^{d}\int_0^1(1-\sigma)\widehat{(\partial_{\xi_j\xi_k}a)}\Big(\xi-\theta,\frac{\xi+\eta}{2}+\sigma\frac{\theta-\eta}{2}\Big)(\theta_j-\eta_j)(\theta_k-\eta_k)\,d\sigma.
\end{aligned} \tag{3-32}
$$

Similarly one obtains

$$
\begin{aligned}
\hat{b}\Big(\theta-\eta,\frac{\theta+\eta}{2}\Big)=&\,\hat{b}\Big(\theta-\eta,\frac{\xi+\eta}{2}\Big)+\widehat{(\partial_\xi b)}\Big(\theta-\eta,\frac{\xi+\eta}{2}\Big)\cdot\frac{\theta-\xi}{2}\\
&+\frac{1}{4}\sum_{j,k=1}^{d}\int_0^1(1-\sigma)\widehat{(\partial_{\xi_j\xi_k}b)}\Big(\theta-\eta,\frac{\xi+\eta}{2}+\sigma\frac{\theta-\xi}{2}\Big)(\theta_j-\xi_j)(\theta_k-\xi_k)\,d\sigma.
\end{aligned} \tag{3-33}
$$

By (3-32), (3-33) we deduce that

$$
\begin{aligned}
T_a T_b h - T_{ab} h - \frac{1}{2\mathrm{i}} T_{\{a,b\}} h &= \sum_{p=1}^{6} R_p h, \\
\widehat{(R_p h)}(\xi) &:= \frac{1}{(\sqrt{2\pi})^{3d}} \sum_{\eta,\theta\in\mathbb{Z}^d} r_1(\xi,\theta,\eta) g_p(\xi,\theta,\eta)\hat{h}(\eta),
\end{aligned}
\tag{3-34}
$$

where the symbols g_i are defined as

$$
g_1 := -\frac{1}{4}\sum_{j,k=1}^{d}\int_0^1 (1-\sigma)(\widehat{\partial_{x_k x_j} a})\Big(\xi-\theta, \frac{\xi+\eta}{2}\Big)(\widehat{\partial_{\xi_k\xi_j} b})\Big(\theta-\eta, \frac{\xi+\eta}{2}+\sigma\frac{\theta-\xi}{2}\Big)\,d\sigma, \tag{3-35}
$$

$$
g_2 := -\frac{1}{4}\sum_{j,k=1}^{d}\int_0^1 (1-\sigma)(\widehat{\partial_{\xi_k\xi_j} a})\Big(\xi-\theta, \frac{\xi+\eta}{2}+\sigma\frac{\theta-\eta}{2}\Big)(\widehat{\partial_{x_k x_j} b})\Big(\theta-\eta, \frac{\xi+\eta}{2}\Big)\,d\sigma, \tag{3-36}
$$

$$
g_3 := \frac{1}{4}\sum_{j,k=1}^{d} (\widehat{\partial_{x_j}\partial_{\xi_k} a})\Big(\xi-\theta, \frac{\xi+\eta}{2}\Big)(\widehat{\partial_{x_k}\partial_{\xi_j} b})\Big(\theta-\eta, \frac{\xi+\eta}{2}\Big), \tag{3-37}
$$

$$
g_4 := -\frac{1}{8\mathrm{i}}\sum_{j,k,p=1}^{d}\int_0^1 (1-\sigma)(\widehat{\partial_{x_k x_j \xi_p} a})\Big(\xi-\theta, \frac{\xi+\eta}{2}\Big)(\widehat{\partial_{x_p\xi_k\xi_j} b})\Big(\theta-\eta, \frac{\xi+\eta}{2}+\sigma\frac{\theta-\xi}{2}\Big)\,d\sigma, \tag{3-38}
$$

$$
g_5 := -\frac{1}{8\mathrm{i}}\sum_{j,k,p=1}^{d}\int_0^1 (1-\sigma)(\widehat{\partial_{\xi_k\xi_j x_p} a})\Big(\xi-\theta, \frac{\xi+\eta}{2}+\sigma\frac{\theta-\eta}{2}\Big)(\widehat{\partial_{\xi_p x_k x_j} b})\Big(\theta-\eta, \frac{\xi+\eta}{2}\Big)\,d\sigma, \tag{3-39}
$$

$$
\begin{aligned}
g_6 := \frac{1}{16}\sum_{j,k,p,q=1}^{d}\iint_0^1 &(1-\sigma_1)(1-\sigma_2)(\widehat{\partial_{\xi_j\xi_k x_p x_q} a})\Big(\xi-\theta, \frac{\xi+\eta}{2}+\sigma_1\frac{\theta-\eta}{2}\Big) \\
&\times (\widehat{\partial_{\xi_p\xi_q x_j x_k} b})\Big(\theta-\eta, \frac{\xi+\eta}{2}+\sigma_2\frac{\theta-\xi}{2}\Big)\,d\sigma_1\,d\sigma_2.
\end{aligned}
\tag{3-40}
$$

We prove the estimate (3-20) (with $j=2$) on each term of the sum in (3-34). First of all we note that $r_1(\xi,\theta,\eta)\neq 0$ implies

$$
(\theta,\eta)\in\Big\{\frac{|\xi-\theta|}{\langle\xi+\theta\rangle}\le\frac{8}{5}\epsilon\Big\}\cap\Big\{\frac{|\theta-\eta|}{\langle\theta+\eta\rangle}\le\frac{8}{5}\epsilon\Big\} =: \mathscr{B}(\xi), \quad \xi\in\mathbb{Z}^d. \tag{3-41}
$$

Moreover we note that

$$
(\theta,\eta)\in\mathscr{B}(\xi) \quad\Longrightarrow\quad |\xi|\lesssim|\theta|,\ |\theta|\lesssim|\eta|,\ |\eta|\lesssim|\xi|. \tag{3-42}
$$

We now study the term $R_3 h$ in (3-34) depending on $g_3(\xi,\theta,\eta)$ in (3-37). We need to bound from above, for any $j,k=1,\dots,d$, the $H^{s-m_1-m_2+2}$-Sobolev norm (see (3-41)) of a term like

$$
\begin{aligned}
\widehat{F}_{j,k}(\xi) &:= \sum_{(\theta,\eta)\in\mathscr{B}(\xi)} (\widehat{\partial_{x_j}\partial_{\xi_k} a})\Big(\xi-\theta, \frac{\xi+\eta}{2}\Big)(\widehat{\partial_{x_k}\partial_{\xi_j} b})\Big(\theta-\eta, \frac{\xi+\eta}{2}\Big)\hat{h}(\eta) \\
&= \sum_{\eta\in\mathbb{Z}^d} \hat{c}_{j,k}\Big(\xi-\eta, \frac{\xi+\eta}{2}\Big)\hat{h}(\eta),
\end{aligned}
\tag{3-43}
$$

where we let

$$\hat{c}_{j,k}(p,\zeta) := \sum_{\ell\in\mathbb{Z}^d} (\widehat{\partial_{x_j}\partial_{\xi_k} a})(p-\ell,\zeta)(\widehat{\partial_{x_k}\partial_{\xi_j} b})(\ell,\zeta)\mathtt{1}_{\mathscr{C}(p,\zeta)}, \quad p,\zeta\in\mathbb{Z}^d,$$

$$\mathscr{C}(p,\zeta) := \left\{\ell\in\mathbb{Z}^d : \frac{|p-\ell|}{\langle 2\zeta+\ell\rangle} \le \frac{8}{5}\epsilon\right\} \cap \left\{\ell\in\mathbb{Z}^d : \frac{|\ell|}{\langle \ell-p+2\zeta\rangle} \le \frac{8}{5}\epsilon\right\}$$

and $\mathtt{1}_{\mathscr{C}(p,\zeta)}$ is the characteristic function of the set $\mathscr{C}(p,\zeta)$. Reasoning as in (3-42), we can deduce that for $\ell\in\mathscr{C}(p,\zeta)$ one has

$$|2\zeta| \lesssim \tfrac{1}{2}|2\zeta+p|. \tag{3-44}$$

Indeed $\ell\in\mathscr{C}(p,\zeta)$ implies $(\theta,\eta)\in\mathscr{B}(\xi)$ by setting

$$2\xi = 2\zeta+p, \quad 2\theta = 2\ell+2\zeta-p, \quad 2\eta = 2\zeta-p. \tag{3-45}$$

Hence (3-44) follows by (3-42) by observing that $2\zeta=\xi+\eta$. Using that $a\in\mathscr{N}^{m_1}_{s_0+2}$, $b\in\mathscr{N}^{m_2}_{s_0+2}$ and reasoning as in (3-13), we deduce

$$|\hat{c}_{j,k}(p,\zeta)| \lesssim \langle\zeta\rangle^{m_1+m_2-2}\langle p\rangle^{-s_0}|a|_{\mathscr{N}^{m_1}_{s_0+2}}|b|_{\mathscr{N}^{m_2}_{s_0+2}}. \tag{3-46}$$

By (3-43), (3-42), (3-2), we get

$$\begin{aligned}
\|F_{j,k}\|^2_{H^{s-m_1-m_2+2}} &\lesssim \sum_{\xi\in\mathbb{Z}^d}\langle\xi\rangle^{-2m_1-2m_2+2}\biggl(\sum_{\eta\in\mathbb{Z}^d}\biggl|\hat{c}_{j,k}\biggl(\xi-\eta,\frac{\xi+\eta}{2}\biggr)\biggr||\hat{h}(\eta)|\langle\eta\rangle^s\biggr)^2\\
&\lesssim |a|^2_{\mathscr{N}^{m_1}_{s_0+2}}|b|^2_{\mathscr{N}^{m_2}_{s_0+2}}\sum_{\xi\in\mathbb{Z}^d}\biggl(\sum_{\eta\in\mathbb{Z}^d}|\hat{h}(\eta)|\langle\eta\rangle^s\frac{1}{\langle\xi-\eta\rangle^{s_0}}\biggr)^2 \quad \text{(by (3-46), (3-44), (3-45))}\\
&\lesssim |a|^2_{\mathscr{N}^{m_1}_{s_0+2}}|b|^2_{\mathscr{N}^{m_2}_{s_0+2}}\|\,|\hat{h}(\xi)|\langle\xi\rangle^s \star \langle\xi\rangle^{-s_0}\|_{\ell^2(\mathbb{Z}^d)}\\
&\lesssim \|h\|^2_{H^s}|a|^2_{\mathscr{N}^{m_1}_{s_0+2}}|b|^2_{\mathscr{N}^{m_2}_{s_0+2}},
\end{aligned}$$

where in the last step we used the Young inequality for sequences, the Cauchy–Schwarz inequality and that $\langle\xi\rangle^{-s_0}$ is in $\ell^1(\mathbb{Z}^d)$ if $s_0>d$. Since the estimate above holds for any $j,k=1,\dots,d$, we may absorb the remainder R_3h in (3-34) in $R_2(a,b)h$ satisfying (3-28). One can deal with the other terms g_1, g_2, g_4, g_5, g_6 similarly. □

Lemma 3.3. *Fix $s_0>\frac{1}{2}d$ and let $f,g,h\in H^s(\mathbb{T};\mathbb{C})$ for $s\ge s_0$. Then*

$$fgh = T_{fg}h + T_{gh}f + T_{fh}g + \mathscr{R}(f,g,h), \tag{3-47}$$

where

$$\begin{aligned}
\widehat{\mathscr{R}(f,g,h)}(\xi) &= \frac{1}{(2\pi)^d}\sum_{\eta,\zeta\in\mathbb{Z}^d} a(\xi,\eta,\zeta)\hat{f}(\xi-\eta-\zeta)\hat{g}(\eta)\hat{h}(\zeta),\\
|a(\xi,\eta,\zeta)| &\lesssim_\rho \frac{\max_2(|\xi-\eta-\zeta|,|\eta|,|\zeta|)^\rho}{\max(|\xi-\eta-\zeta|,|\eta|,|\zeta|)^\rho} \quad \textit{for all } \rho\ge 0.
\end{aligned} \tag{3-48}$$

Remark 3.4. An estimate of the form (3-48) implies that the function $(f,g,h)\mapsto\mathscr{R}(f,g,h)$ defines a continuous trilinear form on $H^s\times H^s\times H^s$ with values in $H^{s+\rho}$ as soon as $s>\rho+\frac{1}{2}d$. This will be proved in Lemma 3.7.

Proof. We start by proving the following claim: the term

$$T_{fg}h - \sum_{\xi\in\mathbb{Z}^d} e^{\mathrm{i}\xi\cdot x} \sum_{\eta,\zeta\in\mathbb{Z}^d} \chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\eta|}{\langle\zeta\rangle}\right)\hat{f}(\xi-\eta-\zeta)\hat{g}(\eta)\hat{h}(\zeta)$$

is a remainder of the form (3-48). By (3-6) this is actually true with coefficients $a(\xi,\eta,\zeta)$ of the form

$$a(\xi,\eta,\zeta) := \chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right) - \chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\eta|}{\langle\zeta\rangle}\right).$$

In order to prove this, we consider the following partition of the unity:

$$\Theta_\epsilon(\xi,\eta,\zeta) := 1-\chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\zeta|}{\langle\eta\rangle}\right) - \chi_\epsilon\left(\frac{|\eta|+|\zeta|}{\langle\xi-\eta-\zeta\rangle}\right) - \chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\eta|}{\langle\zeta\rangle}\right). \tag{3-49}$$

Then we can write

$$\begin{aligned}a(\xi,\eta,\zeta) = \left(\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)-1\right)\chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\eta|}{\langle\zeta\rangle}\right) + \chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)\chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\zeta|}{\langle\eta\rangle}\right)\\ + \chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)\chi_\epsilon\left(\frac{|\eta|+|\zeta|}{\langle\xi-\eta-\zeta\rangle}\right) + \chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)\Theta_\epsilon(\xi,\eta,\zeta).\end{aligned} \tag{3-50}$$

Using (3-5) one can prove that each summand in the right-hand side of the equation above is nonzero only if $\max_2(|\xi-\eta-\zeta|, |\eta, ||\zeta|) \sim \max_1(|\xi-\eta-\zeta|, |\eta, ||\zeta|)$. This implies that each summand defines a smoothing remainder as in (3-48). A similar property holds also for $T_{gh}f$ and $T_{fh}g$. At this point we write

$$\begin{aligned}fgh = \sum_{\xi\in\mathbb{Z}^d} e^{\mathrm{i}\xi\cdot x} \sum_{\eta,\zeta\in\mathbb{Z}^d} \Bigg[\Theta_\epsilon(\xi,\eta,\zeta) + \chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\zeta|}{\langle\eta\rangle}\right)\\ + \chi_\epsilon\left(\frac{|\eta|+|\zeta|}{\langle\xi-\eta-\zeta\rangle}\right) + \chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\eta|}{\langle\zeta\rangle}\right)\Bigg]\hat{f}(\xi-\eta-\zeta)\hat{g}(\eta)\hat{h}(\zeta).\end{aligned}$$

One concludes by using the claim at the beginning of the proof. □

Matrices of symbols and operators. Let us consider the subspace $\mathscr{U}$ defined as

$$\mathscr{U} := \{(u^+, u^-) \in L^2(\mathbb{T}^d;\mathbb{C}) \times L^2(\mathbb{T}^d;\mathbb{C}) : u^+ = \bar{u}^-\}. \tag{3-51}$$

Throughout the paper we shall deal with matrices of linear operators acting on $H^s(\mathbb{T}^d;\mathbb{C}^2)$ preserving the subspace $\mathscr{U}$. Consider two operators R_1, R_2 acting on $C^\infty(\mathbb{T}^d;\mathbb{C})$. We define the operator $\mathfrak{F}$ acting on $C^\infty(\mathbb{T}^d;\mathbb{C}^2)$ as

$$\mathfrak{F} := \begin{bmatrix} R_1 & R_2 \\ \bar{R}_2 & \bar{R}_1 \end{bmatrix}, \tag{3-52}$$

where the linear operators $\bar{R}_i[\,\cdot\,]$, $i=1,2$, are defined by the relation $\bar{R}_i[v] := \overline{R_i[\bar{v}]}$. We say that an operator of the form (3-52) is *real-to-real*. It is easy to note that real-to-real operators preserve $\mathscr{U}$ in (3-51). Consider now a symbol $a(x,\xi)$ of order m and set $A := T_a$. Using (3-6) one can check that

$$\bar{A}[h] = \overline{A[\bar{h}]} \quad \Longrightarrow \quad \bar{A} = T_{\tilde{a}}, \quad \tilde{a}(x,\xi) = \overline{a(x,-\xi)}, \tag{3-53}$$

$$\text{(adjoint)} \quad (Ah, v)_{L^2} = (h, A^* v)_{L^2} \quad \Longrightarrow \quad A^* = T_{\bar{a}}. \tag{3-54}$$

By (3-54) we deduce that the operator A is self-adjoint with respect to the scalar product (3-3) if and only if the symbol $a(x,\xi)$ is real-valued. We need the following definition. Consider two symbols $a, b \in \mathscr{N}_s^m$ and the matrix

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

Define the operator (recall (3-7))

$$M := \mathrm{Op}^{\mathrm{BW}}(A(x,\xi)) := \begin{pmatrix} \mathrm{Op}^{\mathrm{BW}}(a(x,\xi)) & \mathrm{Op}^{\mathrm{BW}}(b(x,\xi)) \\ \mathrm{Op}^{\mathrm{BW}}(\overline{b(x,-\xi)}) & \mathrm{Op}^{\mathrm{BW}}(\overline{a(x,-\xi)}) \end{pmatrix}. \tag{3-55}$$

The matrix of paradifferential operators defined above has the following properties:

- *Real-to-real-ness* : by (3-53) we have that the operator M in (3-55) has the form (3-52); hence it is real-to-real.
- *Self-adjointness*: using (3-54) the operator M in (3-55) is self-adjoint with respect to the scalar product on (3-51)

$$(U,V)_{L^2} := \int_{\mathbb{T}^d} U \cdot \overline{V}\, dx, \quad U = \begin{bmatrix} u \\ \bar{u} \end{bmatrix},\ V = \begin{bmatrix} v \\ \bar{v} \end{bmatrix}. \tag{3-56}$$

if and only if

$$a(x,\xi) = \overline{a(x,\xi)}, \quad b(x,-\xi) = b(x,\xi). \tag{3-57}$$

Nonhomogeneous symbols. In this paper we deal with symbols satisfying (3-4) which depend nonlinearly on an extra function $u(t,x)$ (which in the application will be a solution either of (NLS) or (KG)). We are interested in providing estimates of the seminorms (3-4) in terms of the Sobolev norms of the function u.

We recall classical tame estimates for composition of functions; we refer to [Baldi 2013] (see also [Taylor 2000]). A function $f : \mathbb{T}^d \times B_R \to \mathbb{C}$, where $B_R := \{y \in \mathbb{R}^m : |y| < R\}$, $R > 0$, induces the composition operator (Nemytskii)

$$\tilde{f}(u) := f(x, u(x), Du(x), \dots, D^p u(x)), \tag{3-58}$$

where $D^k u(x)$ denote the derivatives ∂_x^α of order $|\alpha| = k$ (the number m of y-variables depends on p, d).

Lemma 3.5. *Fix $\gamma > 0$ and assume that $f \in C^\infty(\mathbb{T}^d \times B_R; \mathbb{R})$. Then, for any $u \in H^{\gamma+p}$ with $\|u\|_{W^{p,\infty}} < R$, one has*

$$\|\tilde{f}(u)\|_{H^\gamma} \le C\|f\|_{C^\gamma}(1 + \|u\|_{H^{\gamma+p}}), \tag{3-59}$$

$$\Big\| \tilde{f}(u+h) - \sum_{n=0}^{N} \frac{1}{n!}\partial_u^n \tilde{f}[h,\dots,h] \Big\|_{H^\gamma} \le C\|h\|_{W^{p,\infty}}^N (\|h\|_{H^\gamma} + \|h\|_{W^{p,\infty}}\|u\|_{H^{\gamma+p}}) \tag{3-60}$$

for any $h \in H^{\gamma+p}$ with $\|h\|_{W^{p,\infty}} < \frac{1}{2}R$ and where $C > 0$ is a constant depending on γ and the norm $\|u\|_{W^{p,\infty}}$.

Consider a function $F(y_0, y_1, \dots, y_d)$ in $C^\infty(\mathbb{C}^{d+1}; \mathbb{R})$ in the *real* sense; i.e., F is C^∞ as function of $\mathrm{Re}(y_i)$, $\mathrm{Im}(y_i)$. Assume that F has a zero of order at least $p+2 \in \mathbb{N}$ at the origin. Consider a symbol $f(\xi)$, independent of $x \in \mathbb{T}^d$, such that $|f|_{\mathcal{N}_s^m} \le C < +\infty$ for some constant C. Let us define the symbol

$$a(x,\xi) := (\partial_{z_j^\alpha z_k^\beta} F)(u, \nabla u) f(\xi), \quad z_j^\alpha := \partial_{x_j}^\alpha u^\sigma,\ z_k^\beta := \partial_{x_k}^\beta u^{\sigma'} \tag{3-61}$$

for some $j,k = 1,\dots,d$, $\alpha,\beta \in \{0,1\}$ and $\sigma,\sigma' \in \{\pm\}$, where we use the notation $u^+ = u$ and $u^- = \bar{u}$.

Lemma 3.6. *Fix $s_0 > \frac{1}{2}d$. For $u \in B_R(H^{s+s_0+1}(\mathbb{T}^d;\mathbb{C}))$ with $0 < R < 1$, we have*

$$|a|_{\mathcal{N}_s^m} \lesssim \|u\|^p_{H^{s+s_0+1}},$$

where a is the symbol in (3-61). Moreover, the map $h \to (\partial_u a)(u;x,\xi)h$ is a $\mathbb{C}$-linear map from H^{s+s_0+1} to $\mathbb{C}$ and satisfies

$$|(\partial_u a)h|_{\mathcal{N}_s^m} \lesssim \|h\|_{H^{s+s_0+1}} \|u\|^{p-1}_{H^{s+s_0+1}}.$$

The same holds for $\partial_{\bar{u}} a$. Moreover if the symbol a does not depend on ∇u, then the same results are true with $s_0 + 1 \rightsquigarrow s_0$.

Proof. It follows from Lemma 3.5. □

Trilinear operators. Throughout the paper we shall deal with trilinear operators on the Sobolev spaces. We shall adopt a combination of notation introduced in [Berti and Delort 2018; Ionescu and Pusateri 2019]. In particular we are interested in studying properties of operators of the form

$$\begin{gathered} Q = Q[u_1,u_2,u_3] : (C^\infty(\mathbb{T}^d;\mathbb{C}))^3 \to C^\infty(\mathbb{T}^d;\mathbb{C}), \\ \widehat{Q}(\xi) = \frac{1}{(2\pi)^d} \sum_{\eta,\zeta\in\mathbb{Z}^d} q(\xi,\eta,\zeta)\hat{u}_1(\xi-\eta-\zeta)\hat{u}_2(\eta)\hat{u}_3(\zeta) \quad \text{for all } \xi \in \mathbb{Z}^d, \end{gathered} \tag{3-62}$$

where $q(\xi,\eta,\zeta) \in \mathbb{C}$ for any $\xi,\eta,\zeta \in \mathbb{Z}^d$. We now prove that, under certain conditions on the coefficients, the operators of the form (3-62) extend as continuous maps on the Sobolev spaces.

Lemma 3.7. *Let $\mu \ge 0$ and $m \in \mathbb{R}$. Assume that for any $\xi,\eta,\zeta \in \mathbb{Z}^d$ one has*

$$|q(\xi,\eta,\zeta)| \lesssim \frac{\max_2\{\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\}^\mu}{\max_1\{\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\}^m}. \tag{3-63}$$

Then, for $s \ge s_0 > \frac{1}{2}d + \mu$, the map Q in (3-62) with coefficients satisfying (3-63) extends as a continuous map form $(H^s(\mathbb{T}^d;\mathbb{C}))^3$ to $H^{s+m}(\mathbb{T}^d;\mathbb{C})$. Moreover one has

$$\|Q(u_1,u_2,u_3)\|_{H^{s+m}} \lesssim \sum_{i=1}^3 \|u_i\|_{H^s} \prod_{i\ne k} \|u_k\|_{H^{s_0}}. \tag{3-64}$$

Proof. By (3-2) we have

$$\begin{aligned} &\|Q(u_1,u_2,u_3)\|^2_{H^{s+m}} \\ &\quad\le \sum_{\xi\in\mathbb{Z}^d} \langle\xi\rangle^{2(s+m)} \Big(\sum_{\eta,\zeta\in\mathbb{Z}^d} |q(\xi,\eta,\zeta)||\hat{u}_1(\xi-\eta-\zeta)||\hat{u}_2(\eta)||\hat{u}_3(\zeta)| \Big)^2 \\ &\quad\lesssim \sum_{\xi\in\mathbb{Z}^d} \Big(\sum_{\eta,\zeta\in\mathbb{Z}^d} \langle\xi\rangle^s \max_2\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}^\mu |\hat{u}_1(\xi-\eta-\zeta)||\hat{u}_2(\eta)||\hat{u}_3(\zeta)| \Big)^2 \quad \text{(by (3-63))} \\ &\quad:= I + II + III, \end{aligned} \tag{3-65}$$

where I, II, III are the terms in (3-65) which are supported respectively on indexes such that

$$\max_1\{\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\} = \langle\xi-\eta-\zeta\rangle,$$
$$\max_1\{\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\} = \langle\eta\rangle,$$
$$\max_1\{\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\} = \langle\zeta\rangle.$$

Consider for instance the term III. By using the Young inequality for sequences we deduce

$$III \lesssim \|(\langle p\rangle^{\mu}\hat{u}_1(p)) * (\langle\eta\rangle^{\mu}\hat{u}_2(\eta)) * (\langle\zeta\rangle^{s}\hat{u}_3(\zeta))\|_{\ell^2} \lesssim \|u_1\|_{H^{s_0}}\|u_2\|_{H^{s_0}}\|u_3\|_{H^s},$$

which is (3-64). The bounds of I and II are similar. □

In the following lemma we shall prove that a class of "paradifferential" trilinear operators, having some decay on the coefficients, satisfies the hypothesis of the previous lemma.

Lemma 3.8. *Let $\mu \geq 0$ and $m \in \mathbb{R}$, $m \geq 0$. Consider a trilinear map Q as in* (3-62) *with coefficients satisfying*

$$q(\xi,\eta,\zeta) = f(\xi,\eta,\zeta)\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right), \quad |f(\xi,\eta,\zeta)| \lesssim \frac{|\xi-\zeta|^{\mu}}{\langle\zeta\rangle^{m}} \tag{3-66}$$

for any $\xi,\eta,\zeta \in \mathbb{Z}^d$ and $0<\epsilon \ll 1$. Then the coefficients $q(\xi,\eta,\zeta)$ satisfy (3-63) *with $\mu \rightsquigarrow \mu+m$.*

Proof. First of all we write $q(\xi,\eta,\zeta) = q_1(\xi,\eta,\zeta)+q_2(\xi,\eta,\zeta)$ with

$$q_1(\xi,\eta,\zeta) = f(\xi,\eta,\zeta)\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)\chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\eta|}{\langle\zeta\rangle}\right), \tag{3-67}$$

$$q_2(\xi,\eta,\zeta) = f(\xi,\eta,\zeta)\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)\left[\chi_\epsilon\left(\frac{|\xi-\eta-\zeta|+|\zeta|}{\langle\eta\rangle}\right)+\chi_\epsilon\left(\frac{|\eta|+|\zeta|}{\langle\xi-\eta-\zeta\rangle}\right)+\Theta_\epsilon(\xi,\eta,\zeta)\right], \tag{3-68}$$

where $\Theta_\epsilon(\xi,\eta,\zeta)$ is defined in (3-49). Recalling (3-5) one can check that if $q_1(\xi,\eta,\zeta) \neq 0$ then $|\xi-\eta-\zeta|+|\eta| \ll |\zeta| \sim |\xi|$. Together with the bound on $f(\xi,\eta,\zeta)$ in (3-66) we deduce that the coefficients in (3-67) satisfy (3-63). The coefficients in (3-68) satisfy (3-63) because of the support of the cut off function in (3-5). □

Hamiltonian formalism in complex variables. Given a Hamiltonian function $H : H^1(\mathbb{T}^d;\mathbb{C}^2) \to \mathbb{R}$, its Hamiltonian vector field has the form

$$X_H(U) := -\mathrm{i}J\nabla H(U) = -\mathrm{i}\begin{pmatrix}\nabla_{\bar{u}}H(U)\\ -\nabla_u H(U)\end{pmatrix}, \quad J = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix},\ U = \begin{bmatrix}u\\ \bar{u}\end{bmatrix}. \tag{3-69}$$

Indeed one has

$$dH(U)[V] = -\Omega(X_H(U),V) \quad \text{for all } U = \begin{bmatrix}u\\ \bar{u}\end{bmatrix},\ V = \begin{bmatrix}v\\ \bar{v}\end{bmatrix}, \tag{3-70}$$

where Ω is the nondegenerate symplectic form

$$\Omega(U,V) = -\int_{\mathbb{T}^d} U\cdot \mathrm{i}JV\,dx = -\int_{\mathbb{T}^d}\mathrm{i}(u\bar{v}-\bar{u}v)\,dx. \tag{3-71}$$

The Poisson brackets between two Hamiltonians H, G are defined as

$$\{G, H\} := \Omega(X_G, X_H) \overset{(3\text{-}71)}{=} -\int_{\mathbb{T}^d} \mathrm{i} J \nabla G \cdot \nabla H \, dx = -\mathrm{i} \int_{\mathbb{T}^d} \nabla_u H \nabla_{\bar{u}} G - \nabla_{\bar{u}} H \nabla_u G \, dx. \tag{3-72}$$

The nonlinear commutator between two Hamiltonian vector fields is given by

$$[X_G, X_H](U) = dX_G(U)[X_H(U)] - dX_H(U)[X_G(U)] = -X_{\{G,H\}}(U). \tag{3-73}$$

Hamiltonian formalism in real variables. Given a Hamiltonian function $H_{\mathbb{R}} : H^1(\mathbb{T}^d; \mathbb{R}^2) \to \mathbb{R}$, its Hamiltonian vector field has the form

$$X_{H_{\mathbb{R}}}(\psi, \phi) := J \nabla H_{\mathbb{R}}(\psi, \phi) = \begin{pmatrix} \nabla_\phi H_{\mathbb{R}}(\psi, \phi) \\ -\nabla_\psi H_{\mathbb{R}}(\psi, \phi) \end{pmatrix}, \tag{3-74}$$

where J is in (3-69). Indeed one has

$$dH_{\mathbb{R}}(\psi, \phi)[h] = -\widetilde{\Omega}(X_{H_{\mathbb{R}}}(\psi, \phi), h) \quad \text{for all } \begin{bmatrix} \psi \\ \phi \end{bmatrix}, \ h = \begin{bmatrix} \hat{\psi} \\ \hat{\phi} \end{bmatrix}, \tag{3-75}$$

where $\widetilde{\Omega}$ is the nondegenerate symplectic form

$$\widetilde{\Omega}\left(\begin{bmatrix} \psi_1 \\ \phi_1 \end{bmatrix}, \begin{bmatrix} \psi_2 \\ \phi_2 \end{bmatrix} \right) := \int_{\mathbb{T}^d} \begin{bmatrix} \psi_1 \\ \phi_1 \end{bmatrix} \cdot J^{-1} \begin{bmatrix} \psi_2 \\ \phi_2 \end{bmatrix} dx = \int_{\mathbb{T}^d} -(\psi_1 \phi_2 - \phi_1 \psi_2) \, dx. \tag{3-76}$$

We introduce the complex symplectic variables

$$\begin{pmatrix} u \\ \bar{u} \end{pmatrix} = \mathscr{C} \begin{pmatrix} \psi \\ \phi \end{pmatrix} := \frac{1}{\sqrt{2}} \begin{pmatrix} \Lambda_{\mathrm{KG}}^{1/2} \psi + \mathrm{i} \Lambda_{\mathrm{KG}}^{-1/2} \phi \\ \Lambda_{\mathrm{KG}}^{1/2} \psi - \mathrm{i} \Lambda_{\mathrm{KG}}^{-1/2} \phi \end{pmatrix}, \quad \begin{pmatrix} \psi \\ \phi \end{pmatrix} = \mathscr{C}^{-1} \begin{pmatrix} u \\ \bar{u} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Lambda_{\mathrm{KG}}^{-1/2}(u + \bar{u}) \\ -\mathrm{i} \Lambda_{\mathrm{KG}}^{1/2}(u - \bar{u}) \end{pmatrix}, \tag{3-77}$$

where Λ_{KG} is in (1-3). The symplectic form in (3-76) transforms, for

$$U = \begin{bmatrix} u \\ \bar{u} \end{bmatrix}, \quad V = \begin{bmatrix} v \\ \bar{v} \end{bmatrix},$$

into $\Omega(U, V)$ where Ω is in (3-71). In these coordinates the vector field $X_{H_{\mathbb{R}}}$ in (3-74) assumes the form X_H as in (3-69) with $H := H_{\mathbb{R}} \circ \mathscr{C}^{-1}$.

We now study some algebraic properties enjoyed by the Hamiltonian functions previously defined. Let us consider a homogeneous Hamiltonian $H : H^1(\mathbb{T}^d; \mathbb{C}^2) \to \mathbb{R}$ of degree 4 of the form

$$H(U) = (2\pi)^{-d} \sum_{\xi, \eta, \zeta \in \mathbb{Z}^d} \mathtt{h}_4(\xi, \eta, \zeta) \hat{u}(\xi - \eta - \zeta) \hat{\bar{u}}(\eta) \hat{u}(\zeta) \hat{\bar{u}}(-\xi), \quad U = \begin{bmatrix} u \\ \bar{u} \end{bmatrix}, \tag{3-78}$$

for some coefficients $\mathtt{h}_4(\xi, \eta, \zeta) \in \mathbb{C}$ such that

$$\begin{aligned} &\mathtt{h}_4(\xi, \eta, \zeta) = \mathtt{h}_4(-\eta, -\xi, \zeta) = \mathtt{h}_4(\xi, \eta, \xi - \eta - \zeta), \\ &\mathtt{h}_4(\xi, \eta, \zeta) = \overline{\mathtt{h}_4(\zeta, \eta + \zeta - \xi, \xi)} \quad \text{for all } \xi, \eta, \zeta \in \mathbb{Z}^d. \end{aligned} \tag{3-79}$$

By (3-79) one can check that the Hamiltonian H is real-valued and symmetric in its entries. Recalling (3-69) we have that its Hamiltonian vector field can be written as

$$\begin{pmatrix} -\mathrm{i}\nabla_{\bar u} H(U) \\ \mathrm{i}\nabla_u H(U) \end{pmatrix} = \begin{pmatrix} X_H^+(U) \\ \overline{X_H^+(U)} \end{pmatrix}, \tag{3-80}$$

$$\widehat{X_H^+(U)}(\xi) = (2\pi)^{-d} \sum_{\eta,\zeta\in\mathbb{Z}^d} f(\xi,\eta,\zeta)\hat u(\xi-\eta-\zeta)\hat{\bar u}(\eta)\hat u(\zeta), \tag{3-81}$$

where the coefficients $f(\xi,\eta,\zeta)$ have the form

$$f(\xi,\eta,\zeta) = -2\mathrm{i}\mathtt{h}_4(\xi,\eta,\zeta), \quad \xi,\eta,\zeta\in\mathbb{Z}^d. \tag{3-82}$$

We need the following definition.

Definition 3.9 (resonant set). We define the following set of *resonant* indexes:

$$\mathscr{R} := \{(\xi,\eta,\zeta)\in\mathbb{Z}^{3d} : |\xi|=|\zeta|,\ |\eta|=|\xi-\eta-\zeta|\}\cup\{(\xi,\eta,\zeta)\in\mathbb{Z}^{3d} : |\xi|=|\xi-\eta-\zeta|,\ |\eta|=|\zeta|\}. \tag{3-83}$$

Consider the vector field in (3-81) with Hamiltonian H defined in (3-78). We define the field $X_H^{+,\mathrm{res}}(U)$ by

$$\widehat{X_H^{+,\mathrm{res}}}(\xi) = (2\pi)^{-d} \sum_{\eta,\zeta\in\mathbb{Z}^d} f^{(\mathrm{res})}(\xi,\eta,\zeta)\hat u(\xi-\eta-\zeta)\hat{\bar u}(\eta)\hat u(\zeta), \tag{3-84}$$

where

$$f^{(\mathrm{res})}(\xi,\eta,\zeta) := f(\xi,\eta,\zeta)\mathtt{1}_{\mathscr{R}}(\xi,\eta,\zeta), \tag{3-85}$$

where $\mathtt{1}_{\mathscr{R}}$ is the characteristic function of the set $\mathscr{R}$ and f is defined in (3-82).

In the next lemma we prove a fundamental cancellation.

Lemma 3.10. *For $n\geq 0$ one has* (*recall* (3-2))

$$\mathrm{Re}(\langle D\rangle^n X_H^{+,\mathrm{res}}(U), \langle D\rangle^n u)_{L^2} \equiv 0. \tag{3-86}$$

Proof. Using (3-83)–(3-85) one can check that

$$\widehat{X_H^{+,\mathrm{res}}}(\xi) = (2\pi)^{-d} \sum_{(\eta,\zeta)\in\mathscr{R}(\xi)} \mathscr{F}(\xi,\eta,\zeta)\hat u(\xi-\eta-\zeta)\hat{\bar u}(\eta)\hat u(\zeta),$$

with $\mathscr{R}(\xi) := \{(\eta,\zeta)\in\mathbb{Z}^{2d} : |\xi|=|\zeta|,\ |\eta|=|\xi-\eta-\zeta|\}$ for $\xi\in\mathbb{Z}^d$, and

$$\mathscr{F}(\xi,\eta,\zeta) := f(\xi,\eta,\zeta) + f(\xi,\eta,\xi-\eta-\zeta). \tag{3-87}$$

By an explicit computation we have

$$\begin{aligned}\mathrm{Re}(\langle D\rangle^s &X_H^{+,\mathrm{res}}(U), \langle D\rangle^s u)_{L^2} \\ &= (2\pi)^{-d} \sum_{\xi\in\mathbb{Z}^d,(\eta,\zeta)\in\mathscr{R}(\xi)} \langle\xi\rangle^{2s}[\mathscr{F}(\xi,\eta,\zeta)+\overline{\mathscr{F}(\zeta,\zeta+\eta-\xi,\xi)}]\hat u(\xi-\eta-\zeta)\hat{\bar u}(\eta)\hat u(\zeta)\hat{\bar u}(-\xi).\end{aligned}$$

By (3-87), (3-82) and using the symmetries (3-79) we have $\mathscr{F}(\xi,\eta,\zeta)+\overline{\mathscr{F}(\zeta,\zeta+\eta-\xi,\xi)}=0$. □

Remark 3.11. Throughout the paper we shall deal with general Hamiltonian functions of the form

$$H(W)=(2\pi)^{-d}\sum_{\substack{\sigma_1,\sigma_2,\sigma_3,\sigma_4\in\{\pm\}\\ \xi,\eta,\zeta\in\mathbb{Z}^d}}\mathtt{h}^{\sigma_1,\sigma_2,\sigma_3,\sigma_4}(\xi,\eta,\zeta)\widehat{u^{\sigma_1}}(\xi-\eta-\zeta)\widehat{u^{\sigma_2}}(\eta)\widehat{u^{\sigma_3}}(\zeta)\widehat{u^{\sigma_4}}(-\xi),$$

where we use the notation

$$\widehat{u^{\sigma}}(\cdot)=\hat{u}(\cdot)\quad\text{if }\sigma=+\qquad\text{and}\qquad\widehat{u^{\sigma}}(\cdot)=\hat{\bar{u}}(\cdot)\quad\text{if }\sigma=-.\tag{3-88}$$

However, by the definition of the resonant set (3-83), we can note that the resonant vector field has still the form (3-84) and it depends only on the monomials in the Hamiltonian $H(U)$ which are *gauge*-invariant, i.e., of the form (3-78).

4. Paradifferential formulation of the problems

In this section we rewrite the equations in a paradifferential form by means of the paralinearization formula (à la [Bony 1981]). In Section 4A we deal with the problem (NLS) and in Section 4B we deal with (KG).

4A. ***Paralinearization of the NLS.*** In the following we paralinearize (NLS), with respect to the variables $(u,\bar{u})$. We recall that (NLS) may be rewritten as (1-12) and we define $\widetilde{P}(u):=P(u,\nabla u)-\frac{1}{2}|u|^4=\frac{1}{2}|\nabla h(|u|^2)|^2$. We set

$$\tilde{p}(u):=(\partial_{\bar{u}}\widetilde{P})(u,\nabla u)-\sum_{j=1}^{d}\partial_{x_j}(\partial_{\bar{u}_{x_j}}\widetilde{P})(u,\nabla u).\tag{4-1}$$

Lemma 4.1. *Fix $s_0>\frac{1}{2}d$ and $0\le\rho<s-s_0$, $s\ge s_0$. Consider $u\in H^s(\mathbb{T}^d;\mathbb{C})$. Then we have*

$$\begin{aligned}\tilde{p}(u)&=T_{\partial_{u\bar{u}}\widetilde{P}}[u]+T_{\partial_{\bar{u}\bar{u}}\widetilde{P}}[\bar{u}] &\text{(4-2)}\\ &\quad+\sum_{j=1}^{d}(T_{\partial_{\bar{u}u_{x_j}}\widetilde{P}}[u_{x_j}]+T_{\partial_{\bar{u}\bar{u}_{x_j}}\widetilde{P}}[\bar{u}_{x_j}])-\sum_{j=1}^{d}\partial_{x_j}(T_{\partial_{u\bar{u}_{x_j}}\widetilde{P}}[u]+T_{\partial_{\bar{u}\bar{u}_{x_j}}\widetilde{P}}[\bar{u}]) &\text{(4-3)}\\ &\quad-\sum_{j=1}^{d}\partial_{x_j}(T_{\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P}}[u_{x_j}]+T_{\partial_{\bar{u}_{x_j}\bar{u}_{x_j}}\widetilde{P}}[\bar{u}_{x_j}])+R(u), &\text{(4-4)}\end{aligned}$$

where $R(u)$ is a remainder satisfying

$$\|R(u)\|_{H^{s+\rho}}\lesssim C\|u\|_{H^s}^7\tag{4-5}$$

for some constant $C>0$ depending on s, s_0.

Proof. By using the Bony paralinearization formula, see [Bony 1981; Métivier 2008; Taylor 2000], and passing to the Weyl quantization we obtain

$$\begin{aligned}\tilde{p}(u)&=T_{\partial_{u\bar{u}}\widetilde{P}}[u]+T_{\partial_{\bar{u}\bar{u}}\widetilde{P}}[\bar{u}] &\text{(4-6)}\\ &\quad+\sum_{j=1}^{d}(T_{\partial_{\bar{u}u_{x_j}}\widetilde{P}}[u_{x_j}]+T_{\partial_{\bar{u}\bar{u}_{x_j}}\widetilde{P}}[\bar{u}_{x_j}])-\sum_{j=1}^{d}\partial_{x_j}(T_{\partial_{u\bar{u}_{x_j}}\widetilde{P}}[u]+T_{\partial_{\bar{u}\bar{u}_{x_j}}\widetilde{P}}[\bar{u}]) &\text{(4-7)}\\ &\quad-\sum_{j=1}^{d}\partial_{x_j}\sum_{k=1}^{d}(T_{\partial_{\bar{u}_{x_j}u_{x_k}}\widetilde{P}}[u_{x_k}]+T_{\partial_{\bar{u}_{x_j}\bar{u}_{x_k}}\widetilde{P}}[\bar{u}_{x_k}])+R(u), &\text{(4-8)}\end{aligned}$$

where $R(u)$ satisfies the estimate (4-5) since $h(x) \sim x^2$ for $x \sim 0$. The first term in (4-8) is equal to the first in (4-4) because $\partial_{\bar{u}_{x_j} u_{x_k}} \widetilde{P} = \frac{1}{2} \partial_{\bar{u}_{x_j} u_{x_k}} |\nabla h(|u|^2)|^2 = 0$ if $j \neq k$. □

We shall use the following notation throughout the rest of the paper:

$$U := \begin{bmatrix} u \\ \bar{u} \end{bmatrix}, \quad E := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbb{1} := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \operatorname{diag}(b) := b\mathbb{1}, \quad b \in \mathbb{C}. \tag{4-9}$$

Define the *real* symbols

$$\begin{aligned} a_2(x) &:= [h'(|u|^2)]^2 |u|^2, \quad b_2(x) := [h'(|u|^2)]^2 u^2, \\ \vec{a}_1(x) \cdot \xi &:= [h'(|u|^2)]^2 \sum_{j=1}^{d} \operatorname{Im}(u \bar{u}_{x_j}) \xi_j, \quad \xi = (\xi_1, \dots, \xi_d). \end{aligned} \tag{4-10}$$

We define also the matrix of functions

$$A_2(x) := A_2(U; x) := \begin{bmatrix} a_2(U; x) & b_2(U; x) \\ \overline{b_2(U; x)} & a_2(U; x) \end{bmatrix} = \begin{bmatrix} a_2(x) & b_2(x) \\ \overline{b_2(x)} & a_2(x) \end{bmatrix}, \tag{4-11}$$

with $a_2(x)$ and $b_2(x)$ defined in (4-10).

Proposition 4.2 (paralinearization of NLS). *Equation* (NLS) *is equivalent to the system*

$$\dot{U} = -\mathrm{i} E \operatorname{Op}^{\mathrm{BW}}((\mathbb{1} + A_2(x))|\xi|^2) U - \mathrm{i} E V * U - \mathrm{i} \operatorname{Op}^{\mathrm{BW}}(\operatorname{diag}(\vec{a}_1(x) \cdot \xi)) U + X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(U) + R(U), \tag{4-12}$$

where V is the convolution potential in (1-5), *the matrix $A_2(x)$ is the one in* (4-11), *the symbol $\vec{a}_1(x) \cdot \xi$ is in* (4-10) *and the vector field $X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(U)$ is defined as*

$$X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(U) = -\mathrm{i} E \left[\operatorname{Op}^{\mathrm{BW}} \left(\begin{bmatrix} 2|u|^2 & u^2 \\ \bar{u}^2 & 2|u|^2 \end{bmatrix} \right) U + Q_3(U) \right]. \tag{4-13}$$

The seminorms of the symbols satisfy the estimates

$$\begin{aligned} |a_2|_{\mathcal{N}_p^0} + |b_2|_{\mathcal{N}_p^0} &\lesssim \|u\|^6_{H^{p+s_0}} \quad \text{for all } p + s_0 \le s, \ p \in \mathbb{N}, \\ |\vec{a}_1 \cdot \xi|_{\mathcal{N}_p^1} &\lesssim \|u\|^6_{H^{p+s_0+1}} \quad \text{for all } p + s_0 + 1 \le s, \ p \in \mathbb{N}, \end{aligned} \tag{4-14}$$

where we have chosen $s_0 > d$. The remainder $Q_3(U)$ has the form $(Q_3^+(U), \overline{Q_3^+(U)})^T$ and

$$\widehat{Q_3^+}(\xi) = (2\pi)^{-d} \sum_{\eta, \zeta \in \mathbb{Z}^d} \mathtt{q}(\xi, \eta, \zeta) \hat{u}(\xi - \eta - \zeta) \hat{\bar{u}}(\eta) \hat{u}(\zeta) \tag{4-15}$$

for some $\mathtt{q}(\xi, \eta, \zeta) \in \mathbb{C}$. The coefficients of Q_3^+ satisfy

$$|\mathtt{q}(\xi, \eta, \zeta)| \lesssim \frac{\max_2\{\langle \xi - \eta - \zeta \rangle, \langle \eta \rangle, \langle \zeta \rangle\}^\rho}{\max\{\langle \xi - \eta - \zeta \rangle, \langle \eta \rangle, \langle \zeta \rangle\}^\rho} \quad \text{for all } \rho \ge 0. \tag{4-16}$$

The remainder $R(U)$ has the form $(R^+(U), \overline{R^+(U)})^T$. Moreover, for any $s > 2d + 2$, we have the estimates

$$\|R(U)\|_{H^s} \lesssim \|U\|^7_{H^s}, \quad \|Q_3(U)\|_{H^{s+2}} \lesssim \|U\|^3_{H^s}. \tag{4-17}$$

Proof. By Lemma 3.3 the cubic term $|u|^2u$ in (NLS) is equal to $2T_{|u|^2}u+T_{u^2}\bar{u}+\mathscr{R}(u,u,\bar{u})$. Setting $Q_3^+(U)=\mathscr{R}(u,u,\bar{u})$, we get (4-15) by (3-48). The second estimate in (4-17) is a consequence of Lemma 3.7 applied with $\rho=\mu=m=2$.

We now deal with the remaining quasilinear term $\tilde{p}(u)$ defined in (4-1). We start by noting that

$$\partial_{x_j}:=\mathrm{Op}^{\mathrm{BW}}(\mathrm{i}\xi_j),\qquad j=1,\dots d,\tag{4-18}$$

and that the quantization of a symbol $a(x)$ is given by $\mathrm{Op}^{\mathrm{BW}}(a(x))$. We also remark that the symbols appearing in (4-2), (4-3) and (4-4) can be estimated (in the norm $|\cdot|_{\mathcal{N}_s^m}$) by using Lemma 3.6. Consider now the first paradifferential term in (4-4). We have, for any $j=1,\dots,d$,

$$\partial_{x_j}T_{\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P}}\partial_{x_j}u=\mathrm{Op}^{\mathrm{BW}}(\mathrm{i}\xi_j)\circ\mathrm{Op}^{\mathrm{BW}}(\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P})\circ\mathrm{Op}^{\mathrm{BW}}(\mathrm{i}\xi_j)u.$$

By applying Proposition 3.2 and recalling the Poisson bracket in (3-9), we deduce

$$\begin{aligned}\mathrm{Op}^{\mathrm{BW}}(\mathrm{i}\xi_j)\circ\mathrm{Op}^{\mathrm{BW}}(\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P})\circ\mathrm{Op}^{\mathrm{BW}}(\mathrm{i}\xi_j)=\mathrm{Op}^{\mathrm{BW}}(-\xi_j^2\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P})&\qquad(4\text{-}19)\\ +\mathrm{Op}^{\mathrm{BW}}\Big(\frac{\mathrm{i}}{2}\xi_j\partial_{x_j}(\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P})-\frac{\mathrm{i}\xi_j}{2}\partial_{x_j}(\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P})\Big)&\qquad(4\text{-}20)\\ +\widetilde{R}_j^{(1)}(u)+\widetilde{R}_j^{(2)}(u),&\qquad(4\text{-}21)\end{aligned}$$

where

$$\widetilde{R}_j^{(1)}(u):=\mathrm{Op}^{\mathrm{BW}}\big(-\tfrac{1}{4}\partial_{x_jx_j}(\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P})\big)$$

and $\widetilde{R}_j^{(2)}(u)$ is some bounded operator. More precisely, using (3-20), (3-10) and the estimates given by Lemma 3.6, we have, for all $h\in H^s(\mathbb{T}^d;\mathbb{C})$,

$$\|\widetilde{R}_j^{(2)}(u)h\|_{H^s}\le C\|h\|_{H^s}\|u\|_{H^s}^6,\qquad \|\widetilde{R}_j^{(1)}(u)h\|_{H^s}\le C\|h\|_{H^s}\|u\|_{H^{2s_0+3}}^6\tag{4-22}$$

for some constant $C>0$ and $s_0\ge d+1$, $s_0\in\mathbb{N}$. We set

$$\widetilde{R}(u):=\sum_{j=1}^d(\widetilde{R}_j^{(1)}(u)+\widetilde{R}_j^{(2)}(u)).$$

Then

$$\begin{aligned}-\sum_{j=1}^d\partial_{x_j}T_{\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P}}\partial_{x_j}u &=\mathrm{Op}^{\mathrm{BW}}\Big(\sum_{j=1}^d\xi_j^2\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P}\Big)+\widetilde{R}(u)-\frac{\mathrm{i}}{2}\mathrm{Op}^{\mathrm{BW}}\Big(\sum_{j=1}^d(-\xi_j\partial_{x_j}(\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P})+\xi_j\partial_{x_j}(\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P}))\Big)\\ &=\mathrm{Op}^{\mathrm{BW}}(a_2(x)|\xi|^2)+\widetilde{R}(u)+\frac{\mathrm{i}}{2}\mathrm{Op}^{\mathrm{BW}}\Big(\sum_{j=1}^d\xi_j\partial_{x_j}((\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P})-(\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P}))\Big)\quad\text{(by (4-10))}\\ &=\mathrm{Op}^{\mathrm{BW}}(a_2(x)|\xi|^2)+\widetilde{R}(u),\end{aligned}$$

where we used the symmetry of the matrix $\partial_{\overline{\nabla u}\nabla u}\widetilde{P}$ (recall $\widetilde{P}$ is real) and that

$$\partial_{\bar{u}_{x_j}u_{x_j}}\widetilde{P}(u)=\tfrac{1}{2}\partial_{\bar{u}_{x_j}u_{x_j}}|\nabla h(|u|^2)|^2\overset{(4\text{-}10)}{=}a_2(x).$$

By performing similar explicit computations on the other summands in (4-2)-(4-4) we get (4-12), (4-11) with symbols in (4-10). By the discussion above we deduce that the remainder $R(U)$ in (4-12) satisfies the bound (4-17). ◻

Remark 4.3. • The cubic term $X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(U)$ in (4-13) is the Hamiltonian vector field of the Hamiltonian function

$$\mathscr{H}^{(4)}_{\mathrm{NLS}}(U) := \frac{1}{2}\int_{\mathbb{T}^d} |u|^4\,dx, \qquad X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(U) = -\mathrm{i}|u|^2\begin{bmatrix} u \\ \bar{u}\end{bmatrix}. \tag{4-23}$$

• The operators

$$\mathrm{Op}^{\mathrm{BW}}((\mathbb{1}+A_2(x))|\xi|^2), \quad \mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\vec{a}_1(x)\cdot\xi)), \quad \mathrm{Op}^{\mathrm{BW}}\left(\begin{bmatrix} 2|u|^2 & u^2 \\ \bar{u}^2 & 2|u|^2\end{bmatrix}\right)$$

are self-adjoint thanks to (3-57) and (4-10).

4B. ***Paralinearization of the KG.*** In this section we rewrite (KG) as a paradifferential system. This is the content of Proposition 4.7. Before stating this result we need some preliminaries. In particular in Lemma 4.4 below we analyze some properties of the cubic terms in (KG). Define the *real* symbols

$$\begin{aligned} a_2(x,\xi) &:= a_2(u;x,\xi) := \sum_{j,k=1}^{d}(\partial_{\psi_{x_j}\psi_{x_k}}F)(\psi,\nabla\psi)\xi_j\xi_k, \qquad \psi = \frac{\Lambda_{\mathrm{KG}}^{-1/2}}{\sqrt{2}}(u+\bar{u}), \\ a_0(x,\xi) &:= a_0(u;x,\xi) := \tfrac{1}{2}(\partial_{y_1y_1}G)(\psi,\Lambda_{\mathrm{KG}}^{1/2}\psi) + (\partial_{y_1y_0}G)(\psi,\Lambda_{\mathrm{KG}}^{1/2}\psi)\Lambda_{\mathrm{KG}}^{-1/2}(\xi). \end{aligned} \tag{4-24}$$

We define also the matrices of symbols

$$\mathscr{A}_1(x,\xi) := \mathscr{A}_1(u;x,\xi) := \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1\end{bmatrix}\Lambda_{\mathrm{KG}}^{-2}(\xi)a_2(u;x,\xi), \tag{4-25}$$

$$\mathscr{A}_0(x,\xi) := \mathscr{A}_0(u;x,\xi) := \begin{bmatrix} 1 & 1 \\ 1 & 1\end{bmatrix}a_0(u;x,\xi), \tag{4-26}$$

and the Hamiltonian function

$$\mathscr{H}^{(4)}_{\mathrm{KG}}(U) := \int_{\mathbb{T}^d} G(\psi,\Lambda_{\mathrm{KG}}^{1/2}\psi)\,dx, \tag{4-27}$$

with G the function appearing in (1-14). First of all we study some properties of the vector field of the Hamiltonian $\mathscr{H}^{(4)}_{\mathrm{KG}}$.

Lemma 4.4. *We have*

$$X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(U) = -\mathrm{i}J\nabla\mathscr{H}^{(4)}_{\mathrm{KG}}(U) = -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathscr{A}_0(x,\xi))U + Q_3(u), \tag{4-28}$$

with $\mathscr{A}_0$ *in* (4-26). *The remainder* $Q_3(u)$ *has the form* $(Q_3^+(u),\overline{Q_3^+(u)})^T$ *and* (*recall* (3-88))

$$\widehat{Q_3^+}(\xi) = \frac{1}{(2\pi)^d}\sum_{\substack{\sigma_1,\sigma_2,\sigma_3\in\{\pm\}\\ \eta,\zeta\in\mathbb{Z}^d}} \mathtt{q}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\widehat{u^{\sigma_1}}(\xi-\eta-\zeta)\widehat{u^{\sigma_2}}(\eta)\widehat{u^{\sigma_3}}(\zeta) \tag{4-29}$$

for some $\mathtt{q}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\in\mathbb{C}$. *The coefficients of* Q_3^+ *satisfy*

$$|\mathtt{q}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)| \lesssim \frac{\max_2\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}}{\max\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}} \tag{4-30}$$

for any $\sigma_1, \sigma_2, \sigma_3 \in \{\pm\}$. *Finally, for* $s > 2d+1$, *we have*

$$|a_0|_{\mathcal{N}_p^0} \lesssim \|u\|^2_{H^{p+s_0}}, \qquad p+s_0 \le s, \quad s_0 > d, \tag{4-31}$$

$$\|X_{\mathcal{H}^{(4)}_{\mathrm{KG}}}(U)\|_{H^s} \lesssim \|u\|^3_{H^s}, \qquad \|Q_3(u)\|_{H^{s+1}} \lesssim \|u\|^3_{H^s}, \tag{4-32}$$

$$\|d_U X_{\mathcal{H}^{(4)}_{\mathrm{KG}}}(U)[h]\|_{H^s} \lesssim \|u\|^2_{H^s}\|h\|_{H^s} \quad \textit{for all } h \in H^s(\mathbb{T}^d;\mathbb{C}^2). \tag{4-33}$$

Proof. By an explicit computation and using (1-2) we get

$$X_{\mathcal{H}^{(4)}_{\mathrm{KG}}}(U) = (X^+_{\mathcal{H}^{(4)}_{\mathrm{KG}}}(U), \overline{X^+_{\mathcal{H}^{(4)}_{\mathrm{KG}}}(U)})^T, \quad X^+_{\mathcal{H}^{(4)}_{\mathrm{KG}}}(U) = -\mathrm{i}\frac{\Lambda_{\mathrm{KG}}^{-1/2}}{\sqrt{2}} g(\psi).$$

The function g is a homogeneous polynomial of degree 3. Hence, by using Lemma 3.3, we obtain

$$\mathrm{i}X^+_{\mathcal{H}^{(4)}_{\mathrm{KG}}}(U) = A_0 + A_{-1/2} + A_{-1} + Q^{-\rho}(u), \tag{4-34}$$

where

$$A_0 := \tfrac{1}{2}\mathrm{Op}^{\mathrm{BW}}(\partial_{y_1y_1}G(\psi, \Lambda^{1/2}_{\mathrm{KG}}\psi))[u+\bar u], \tag{4-35}$$

$$A_{-1/2} := \tfrac{1}{2}\mathrm{Op}^{\mathrm{BW}}(\partial_{y_1y_0}G(\psi, \Lambda^{1/2}_{\mathrm{KG}}\psi))[\Lambda^{-1/2}_{\mathrm{KG}}(u+\bar u)] + \tfrac{1}{2}\Lambda^{-1/2}_{\mathrm{KG}}\mathrm{Op}^{\mathrm{BW}}(\partial_{y_1y_0}G(\psi, \Lambda^{1/2}_{\mathrm{KG}}\psi))[u+\bar u], \tag{4-36}$$

$$A_{-1} := \tfrac{1}{2}\Lambda^{-1/2}_{\mathrm{KG}}\mathrm{Op}^{\mathrm{BW}}(\partial_{y_0y_0}G(\psi, \Lambda^{1/2}_{\mathrm{KG}}\psi))[\Lambda^{-1/2}_{\mathrm{KG}}(u+\bar u)], \tag{4-37}$$

and $Q^{-\rho}$ is a cubic smoothing remainder of the form (3-48) whose coefficients satisfy the bound (4-30). The symbols of the paradifferential operators have the form (using that G is a polynomial)

$$(\partial_{kj}G)\left(\frac{\Lambda^{-1/2}_{\mathrm{KG}}(u+\bar u)}{\sqrt{2}}, \frac{u+\bar u}{\sqrt{2}}\right) = (2\pi)^{-d}\sum_{\sigma_1,\sigma_2\in\{\pm\}}\sum_{\xi\in\mathbb{Z}^d} e^{\mathrm{i}\xi\cdot x}\sum_{\eta\in\mathbb{Z}^d} \mathtt{g}^{\sigma_1,\sigma_2}_{k,j}(\xi,\eta)\widehat{u^{\sigma_1}}(\xi-\eta)\widehat{u^{\sigma_2}}(\eta), \tag{4-38}$$

where $k, j \in \{y_0, y_1\}$ and where the coefficients $\mathtt{g}^{\sigma_1,\sigma_2}_{k,j}(\xi,\eta) \in \mathbb{C}$ satisfy $|\mathtt{g}^{\sigma_1,\sigma_2}_{k,j}(\xi,\eta)| \lesssim 1$.

We claim that the term in (4-37) is a cubic remainder of the form (4-29) with coefficients satisfying (4-30). By (3-6) we have

$$\begin{aligned}\widehat{A}_{-1}(\xi) &= \frac{1}{2(2\pi)^d}\sum_{\zeta\in\mathbb{Z}^d,\sigma\in\{\pm\}} \widehat{\partial_{y_0y_0}G}(\xi-\zeta)\Lambda^{-1/2}_{\mathrm{KG}}(\xi)\Lambda^{-1/2}_{\mathrm{KG}}(\zeta)\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)\widehat{u^\sigma}(\xi) \\ &= \frac{1}{2(2\pi)^d}\sum_{\substack{\sigma_1,\sigma_2,\sigma\in\{\pm\}\\ \eta,\zeta\in\mathbb{Z}^d}} \mathtt{g}^{\sigma_1,\sigma_2}_{y_0,y_0}(\xi-\zeta,\eta)\Lambda^{-1/2}_{\mathrm{KG}}(\xi)\Lambda^{-1/2}_{\mathrm{KG}}(\zeta) \\ &\qquad\qquad \times \chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)\widehat{u^{\sigma_1}}(\xi-\eta-\zeta)\widehat{u^{\sigma_2}}(\eta)\widehat{u^\sigma}(\zeta) \quad \text{(by (4-38))},\end{aligned}$$

which implies that A_{-1} has the form (4-29) with coefficients

$$\mathtt{a}^{\sigma_1,\sigma_2,\sigma_3}_{-1}(\xi,\eta,\zeta) = \tfrac{1}{2}\mathtt{g}^{\sigma_1,\sigma_2}_{y_0,y_0}(\xi-\zeta,\eta)\Lambda^{-1/2}_{\mathrm{KG}}(\xi)\Lambda^{-1/2}_{\mathrm{KG}}(\zeta)\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right). \tag{4-39}$$

By Lemma 3.8 we have that the coefficients in (4-39) satisfy (4-30). This proves the claim for the operator A_{-1}. We now study the term in (4-36). We remark that, by Proposition 3.2 (see the composition formula (3-19)), we have $A_{-1/2} = \mathrm{Op}^{\mathrm{BW}}(\Lambda^{-1/2}_{\mathrm{KG}}(\xi)\partial_{y_0y_1}G)$ up to a smoothing operator of order $-\frac{3}{2}$.

Actually to prove that such a remainder has the form (4-29) with coefficients (4-30) it is more convenient to compute the composition operator explicitly. In particular, recalling (3-6), we get

$$A_{-1/2} = \mathrm{Op}^{\mathrm{BW}}(\Lambda_{\mathrm{KG}}^{-1/2}(\xi)\partial_{y_0 y_1} G) + R_{-1}, \tag{4-40}$$

where

$$\widehat{R}_{-1}(\xi) = (2\pi)^{-d} \sum_{\substack{\sigma_1,\sigma_2,\sigma\in\{\pm\}\\ \eta,\zeta\in\mathbb{Z}^d}} \mathtt{r}^{\sigma_1,\sigma_2,\sigma}(\xi-\eta-\zeta,\eta,\zeta)\widehat{u^{\sigma_1}}(\xi-\eta-\zeta)\widehat{u^{\sigma_2}}(\eta)\widehat{u^{\sigma}}(\zeta),$$

$$\mathtt{r}^{\sigma_1,\sigma_2,\sigma}(\xi-\eta-\zeta,\eta,\zeta) = \frac{1}{2}\mathtt{g}^{\sigma_1,\sigma_2}_{y_0,y_1}(\xi-\zeta,\eta)\chi_\epsilon\biggl(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\biggr)\biggl[\Lambda_{\mathrm{KG}}^{-1/2}(\xi)+\Lambda_{\mathrm{KG}}^{-1/2}(\zeta)-2\Lambda_{\mathrm{KG}}^{-1/2}\biggl(\frac{\xi+\zeta}{2}\biggr)\biggr].$$

We note that

$$\Lambda_{\mathrm{KG}}^{-1/2}(\xi) = \Lambda_{\mathrm{KG}}^{-1/2}\biggl(\frac{\xi+\zeta}{2}\biggr) - \frac{1}{2}\int_0^1 \Lambda_{\mathrm{KG}}^{-3/2}\biggl(\frac{\xi+\zeta}{2}+\tau\frac{\xi-\zeta}{2}\biggr)\,d\tau.$$

Then we deduce

$$\biggl|\Lambda_{\mathrm{KG}}^{-1/2}(\xi)+\Lambda_{\mathrm{KG}}^{-1/2}(\zeta)-2\Lambda_{\mathrm{KG}}^{-1/2}\biggl(\frac{\xi+\zeta}{2}\biggr)\biggr| \lesssim |\xi|^{-3/2}+|\zeta|^{-3/2}.$$

Again by Lemma 3.8 one can conclude that $\mathtt{r}^{\sigma_1,\sigma_2,\sigma}(\xi-\eta-\zeta,\eta,\zeta)$ satisfies (4-30). By (4-40), (4-35), (4-37) and recalling the definition of $a_0(x,\xi)$ in (4-24), we obtain (4-28). The bound (4-32) for Q_3 follows by (4-30) and Lemma 3.7. Moreover the bound (4-31) follows by Lemma 3.6 recalling that $G(\psi,\Lambda_{\mathrm{KG}}^{1/2}\psi)\sim O(u^4)$. Then the bound (4-32) for $X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}$ follows by Lemma 3.1. Let us prove (4-33). By differentiating (4-28) we get

$$d_U X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(U)[h] = -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathscr{A}_0(x,\xi))h - \mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(d_U\mathscr{A}_0(x,\xi)h)U + d_U Q_3(u)[h]. \tag{4-41}$$

The first summand in (4-41) satisfies (4-33) by Lemma 3.1 and (4-31). Moreover using (4-38) and (4-24) one can check that

$$|d_U\mathscr{A}_0(x,\xi)h|_{\mathscr{N}_p^0} \lesssim \|u\|_{H^{p+s_0}}\|h\|_{H^{p+s_0}}, \quad p+s_0\le s.$$

Then the second summand in (4-41) verifies the bound (4-33) again by Lemma 3.1. The estimate on the third summand in (4-41) follows by (4-29), (4-30) and Lemma 3.7. □

Remark 4.5. We remark that the symbol $a_0(x,\xi)$ in (4-24) is homogeneous of degree 2 in the variables $u,\bar{u}$. In particular, by (4-38), we have

$$\begin{aligned} a_0(x,\xi) &= (2\pi)^{-d/2}\sum_{p\in\mathbb{Z}^d} e^{\mathrm{i}p\cdot x}\hat{a}_0(p,\xi),\\ \hat{a}_0(p,\xi) &= (2\pi)^{-d}\sum_{\substack{\sigma_1,\sigma_2\in\{\pm\}\\ \eta\in\mathbb{Z}^d}} a_0^{\sigma_1,\sigma_2}(p,\eta,\xi)\widehat{u^{\sigma_1}}(p-\eta)\widehat{u^{\sigma_2}}(\eta), \\ a_0^{\sigma_1,\sigma_2}(p,\eta,\xi) &:= \tfrac{1}{2}\mathtt{g}^{\sigma_1,\sigma_2}_{y_1,y_1}(p,\eta)+\mathtt{g}^{\sigma_1,\sigma_2}_{y_0,y_1}(p,\eta)\Lambda_{\mathrm{KG}}^{-1/2}(\xi). \end{aligned} \tag{4-42}$$

Moreover one has $|a_0^{\sigma_1,\sigma_2}(p,\eta,\xi)|\lesssim 1$. Since the symbol $a_0(x,\xi)$ is real-valued, one can check that

$$a_0^{\sigma_1,\sigma_2}(p,\eta,\xi) = \overline{a_0^{-\sigma_1,-\sigma_2}(-p,-\eta,\xi)} \quad \text{for all } \xi,p,\eta\in\mathbb{Z}^d,\ \sigma_1,\sigma_2\in\{\pm\}. \tag{4-43}$$

Remark 4.6. Consider the special case when the function G in (1-2) is independent of y_1. Following the proof of Lemma 4.4 one can obtain the formula (4-28) with symbol $a_0(x,\xi)$ of order -1 given by (see (4-37))

$$a_0(x,\xi) := \tfrac{1}{2}\partial_{y_0 y_0} G(\psi)\Lambda_{\mathrm{KG}}^{-1}(\xi).$$

The remainder Q_3 would satisfy (4-30) with better denominator $\max\{\langle \xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\}^2$.

The main result of this section is the following.

Proposition 4.7 (paralinearization of KG). *The system* (1-13) *is equivalent to*

$$\dot{U} = -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}\big((\mathbb{1}+\mathscr{A}_1(x,\xi))\Lambda_{\mathrm{KG}}(\xi)\big)U + X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(U) + R(u), \tag{4-44}$$

where

$$U := \begin{bmatrix} u \\ \bar{u} \end{bmatrix} := \mathscr{C}\begin{bmatrix} \psi \\ \phi \end{bmatrix}$$

(*see* (3-77)), $\mathscr{A}_1(x,\xi)$ *is in* (4-25), *and* $X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(U)$ *is the Hamiltonian vector field of* (4-27). *The operator* $R(u)$ *has the form* $(R^+(u), \overline{R^+(u)})^T$. *Moreover we have*

$$|\mathscr{A}_1|_{\mathcal{N}_p^0} + |a_2|_{\mathcal{N}_p^2} + \lesssim \|u\|^3_{H^{p+s_0+1}} \quad \textit{for all } p+s_0+1 \le s,\ \ p\in\mathbb{N}, \tag{4-45}$$

where we have chosen $s_0 > d$. *Finally there is* $\mu > 0$ *such that*, *for any* $s > 2d+\mu$, *the remainder* $R(u)$ *satisfies*

$$\|R(u)\|_{H^s} \lesssim \|u\|^4_{H^s}. \tag{4-46}$$

Proof. First of all we note that system (1-13) in the complex coordinates (3-77) reads

$$\partial_t u = -\mathrm{i}\Lambda_{\mathrm{KG}} u - \mathrm{i}\frac{\Lambda_{\mathrm{KG}}^{-1/2}}{\sqrt{2}}(f(\psi)+g(\psi)), \qquad \psi = \frac{\Lambda_{\mathrm{KG}}^{-1/2}(u+\bar{u})}{\sqrt{2}}, \tag{4-47}$$

with $f(\psi)$, $g(\psi)$ in (1-1), (1-2). The term $(-\mathrm{i}/\sqrt{2})\Lambda_{\mathrm{KG}}^{-1/2}g(\psi)$ is the first component of the vector field $X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(U)$, which was studied in Lemma 4.4. By using the Bony paralinearization formula (see [Bony 1981; Métivier 2008; Taylor 2000]), passing to the Weyl quantization and (1-1) we get

$$f(\psi) = -\sum_{j,k=1}^{d} \partial_{x_j}\circ \mathrm{Op}^{\mathrm{BW}}((\partial_{\psi_{x_j}\psi_{x_k}}F)(\psi,\nabla\psi))\circ\partial_{x_k}\psi \tag{4-48}$$

$$+\sum_{j=1}^{d}[\mathrm{Op}^{\mathrm{BW}}((\partial_{\psi\psi_{x_j}}F)(\psi,\nabla\psi)),\partial_{x_j}]\psi + \mathrm{Op}^{\mathrm{BW}}((\partial_{\psi\psi}F)(\psi,\nabla\psi))\psi + R^{-\rho}(\psi), \tag{4-49}$$

where $R^{-\rho}(\psi)$ satisfies $\|R^{-\rho}(\psi)\|_{H^{s+\rho}} \lesssim \|\psi\|^4_{H^s}$ for any $s \ge s_0 > d+\rho$. By Lemma 3.6, and recalling that $F(\psi,\nabla\psi)\sim O(\psi^5)$, we have

$$|\partial_{\psi_{x_k}\psi_{x_j}}F|_{\mathcal{N}_p^0} + |\partial_{\psi\psi_{x_j}}F|_{\mathcal{N}_p^0} + |\partial_{\psi\psi}F|_{\mathcal{N}_p^0} \lesssim \|\psi\|^3_{H^{p+s_0+1}}, \quad p+s_0+1\le s, \tag{4-50}$$

where $s_0 > d$. Recall that $\partial_{x_j} = \mathrm{Op}^{\mathrm{BW}}(\xi_j)$. Then, by Proposition 3.2, we have

$$[\mathrm{Op}^{\mathrm{BW}}(\partial_{\psi\psi_{x_j}}F),\partial_{x_j}]\psi = \mathrm{Op}^{\mathrm{BW}}(-\mathrm{i}\{\partial_{\psi\psi_{x_j}}F,\xi_j\})\psi + Q(\psi),$$

with (see (3-20)) $\|Q(\psi)\|_{H^{s+1}} \lesssim |\partial_{\psi\psi_{x_j}} F|_{\mathcal{N}^0_{s_0+2}} \|\psi\|_{H^s}$. Then by (3-8), (4-50) and (3-10) (see Lemma 3.1 and Proposition 3.2) we deduce that the terms in (4-49) can be absorbed in a remainder satisfying (4-46) with $s \gg 2d$ large enough. We now consider the right-hand side of (4-48). We have

$$-\partial_{x_j} \circ \mathrm{Op}^{\mathrm{BW}}((\partial_{\psi_{x_j}\psi_{x_k}} F)(\psi, \nabla\psi)) \circ \partial_{x_k} = \mathrm{Op}^{\mathrm{BW}}(\xi_j)\, \mathrm{Op}^{\mathrm{BW}}((\partial_{\psi_{x_j}\psi_{x_k}} F)(\psi, \nabla\psi))\, \mathrm{Op}^{\mathrm{BW}}(\xi_k).$$

By using again Lemma 3.1 and Proposition 3.2 we get

$$f(\psi) = \mathrm{Op}^{\mathrm{BW}}(a_2(x, \xi))\psi + \widetilde{R}(\psi), \tag{4-51}$$

where a_2 is in (4-24) and $\widetilde{R}(\psi)$ is a remainder satisfying (4-46). The symbol $a_2(x, \xi)$ satisfies (4-45) by (4-50). Moreover

$$\frac{1}{\sqrt{2}} \Lambda_{\mathrm{KG}}^{-1/2} f(\psi) = \frac{1}{\sqrt{2}} \Lambda_{\mathrm{KG}}^{-1/2} f\left(\frac{\Lambda_{\mathrm{KG}}^{-1/2}(u+\bar{u})}{\sqrt{2}}\right) \overset{(4\text{-}51)}{=} \frac{1}{2}\, \mathrm{Op}^{\mathrm{BW}}(a_2(x, \xi)\Lambda_{\mathrm{KG}}^{-1}(\xi))[u+\bar{u}] \tag{4-52}$$

up to remainders satisfying (4-46). Here we used Proposition 3.2 to study the composition operator $\Lambda_{\mathrm{KG}}^{-1/2}\, \mathrm{Op}^{\mathrm{BW}}(a_2(x, \xi))\Lambda_{\mathrm{KG}}^{-1/2}$. By the discussion above and formula (4-47) we deduce (4-44). □

Remark 4.8. In the semilinear case, i.e., when $f = 0$ and g does not depend on y_1 (see (1-1), (1-2)), equation (4-44) reads

$$\dot{U} = -\mathrm{i} E\, \mathrm{Op}^{\mathrm{BW}}(\mathbb{1}\Lambda_{\mathrm{KG}}(\xi))U + X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(U),$$

where the vector field $X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}$ has the particular structure described in Remark 4.6.

5. Approximately symplectic maps

5A. *Paradifferential Hamiltonian vector fields.* In this section we shall construct some approximately symplectic changes of coordinates which will be important for the diagonalization procedure of Section 6.

Define the frequency localization

$$S_\xi w := \sum_{k \in \mathbb{Z}^d} \hat{w}(k)\chi_\epsilon\left(\frac{|k|}{\langle\xi\rangle}\right)e^{\mathrm{i}k\cdot x}, \quad \xi \in \mathbb{Z}^d, \tag{5-1}$$

for some $0 < \epsilon < 1$, where χ_ϵ is defined in (3-5). Consider the matrix of symbols

$$B_{\mathrm{NLS}}(W; x, \xi) := B_{\mathrm{NLS}}(x, \xi) := \begin{pmatrix} 0 & b_{\mathrm{NLS}}(x, \xi) \\ \overline{b_{\mathrm{NLS}}(x, -\xi)} & 0 \end{pmatrix}, \quad b_{\mathrm{NLS}}(x, \xi) = \tilde{\chi}(\xi) w^2 \frac{1}{2|\xi|^2}, \tag{5-2}$$

where $\tilde{\chi}(\xi)$ is a $C^\infty(\mathbb{R}; \mathbb{R}^+)$ function equal to 0 if $|\xi| \le \frac{1}{4}$ and 1 if $|\xi| \ge \frac{1}{2}$. Define also the Hamiltonian function

$$\mathscr{B}_{\mathrm{NLS}}(W) := \frac{1}{2} \int_{\mathbb{T}^d} \mathrm{i} E\, \mathrm{Op}^{\mathrm{BW}}(B_{\mathrm{NLS}}(S_\xi W; x, \xi))W \cdot \overline{W}\, dx, \tag{5-3}$$

where $S_\xi W := (S_\xi w, S_\xi \bar{w})^T$. The presence of truncation on the high modes (S_ξ) will be decisive in obtaining Lemma 5.1 (see comments in the proof of this lemma).

Analogously we define the following. Consider the matrix of symbols

$$B_{\mathrm{KG}}(W;x,\xi) := B_{\mathrm{KG}}(x,\xi) := \begin{pmatrix} 0 & b_{\mathrm{KG}}(x,\xi) \\ \overline{b_{\mathrm{KG}}(x,-\xi)} & 0 \end{pmatrix}, \quad b_{\mathrm{KG}}(W;x,\xi) = \frac{a_0(x,\xi)}{2\Lambda_{\mathrm{KG}}(\xi)}, \tag{5-4}$$

with $a_0(x,\xi)$ in (4-24) and Λ_{KG} in (1-4), and define the Hamiltonian function

$$\mathscr{B}_{\mathrm{KG}}(W) := \frac{1}{2}\int_{\mathbb{T}^d} \mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(B_{\mathrm{KG}}(S_\xi W;x,\xi))W\cdot \overline{W}\,dx, \tag{5-5}$$

where $S_\xi W := (S_\xi w, S_\xi \bar{w})^T$, where S_ξ is in (5-1).

In this section we study some properties of the maps generated by the Hamiltonians $\mathscr{B}_{\mathrm{NLS}}(W)$ in (5-3) and $\mathscr{B}_{\mathrm{KG}}(W)$ in (5-5). In the next lemma we show that their Hamiltonian vector fields are given by $\mathrm{Op}^{\mathrm{BW}}(B_{\mathrm{NLS}}(W;x,\xi))W$ and $\mathrm{Op}^{\mathrm{BW}}(B_{\mathrm{KG}}(W;x,\xi))W$ respectively, modulo smoothing remainders. More precisely we have the following.

Lemma 5.1. *Consider the Hamiltonian function $\mathscr{B}(W)$ equal to $\mathscr{B}_{\mathrm{NLS}}$ in* (5-3) *or $\mathscr{B}_{\mathrm{KG}}$ in* (5-5)*. One has that the Hamiltonian vector field of $\mathscr{B}(W)$ has the form*

$$X_{\mathscr{B}}(W) = -\mathrm{i}J\nabla\mathscr{B}(W) = \mathrm{Op}^{\mathrm{BW}}(B(W;x,\xi))W + Q_{\mathscr{B}}(W), \tag{5-6}$$

where $Q_{\mathscr{B}}(W)$ is a smoothing remainder of the form $(Q^+_{\mathscr{B}}(W), \overline{Q^+_{\mathscr{B}}(W)})^T$ and the symbol $B(W;x,\xi)$ is respectively equal to $B_{\mathrm{NLS}}(W;x,\xi)$ in (5-2) *or $B_{\mathrm{KG}}(W;x,\xi)$ in* (5-4)*. In particular the cubic remainder $Q_{\mathscr{B}}(W)$ has the form*

$$(\widehat{Q^+_{\mathscr{B}}(W)})(\xi) = \frac{1}{(2\pi)^d}\sum_{\substack{\sigma_1,\sigma_2,\sigma_3\in\{\pm\}\\ \eta,\zeta\in\mathbb{Z}^d}} \mathtt{q}_{\mathscr{B}}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\widehat{w^{\sigma_1}}(\xi-\eta-\zeta)\widehat{w^{\sigma_2}}(\eta)\widehat{w^{\sigma_3}}(\zeta), \quad \xi\in\mathbb{Z}^d, \tag{5-7}$$

where $\mathtt{q}_{\mathscr{B}}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\in\mathbb{C}$ satisfy, for any $\xi,\eta,\zeta\in\mathbb{Z}^d$, a bound like (3-48)*. In the case that $\mathscr{B}=\mathscr{B}_{\mathrm{NLS}}$ we have $\sigma_1=+$, $\sigma_2=-$, $\sigma_3=+$. Moreover, for $s>\frac{1}{2}d+\rho$, we have*

$$\|d_W^k Q_{\mathscr{B}}(W)[h_1,\dots,h_k]\|_{H^{s+\rho}} \lesssim \|w\|_{H^s}^{3-k}\prod_{i=1}^k\|h_i\|_{H^s} \quad \text{for all } h_i\in H^s(\mathbb{T}^d;\mathbb{C}^2),\ i=1,2,3, \tag{5-8}$$

for $k=0,1,2,3$. Moreover, for any $s>2d+2$, one has

$$\|d_W^k X_{\mathscr{B}_{\mathrm{NLS}}}(W)[h_1,\dots,h_k]\|_{H^{s+2}} \lesssim \|w\|_{H^s}^{3-k}\prod_{i=1}^k\|h_i\|_{H^s} \quad \text{for all } h_i\in H^s(\mathbb{T}^d;\mathbb{C}^2),\ i=1,2,3, \tag{5-9}$$

$$\|d_W^k X_{\mathscr{B}_{\mathrm{KG}}}(W)[h_1,\dots,h_k]\|_{H^{s+1}} \lesssim \|w\|_{H^s}^{3-k}\prod_{i=1}^k\|h_i\|_{H^s} \quad \text{for all } h_i\in H^s(\mathbb{T}^d;\mathbb{C}^2),\ i=1,2,3, \tag{5-10}$$

with $k=0,1,2,3$.

Proof. We prove the statement in the case $\mathscr{B}=\mathscr{B}_{\mathrm{NLS}}$; the other case is similar. Using the formulas (5-2), (5-3) we obtain $\mathscr{B}_{\mathrm{NLS}}(W) = -G_1(W)-G_2(W)$ with

$$G_1(W) := -\frac{\mathrm{i}}{2}\int_{\mathbb{T}^d}\mathrm{Op}^{\mathrm{BW}}(b_{\mathrm{NLS}}(S_\xi w))\bar{w}\bar{w}\,dx, \quad G_2(W) := \frac{\mathrm{i}}{2}\int_{\mathbb{T}^d}\mathrm{Op}^{\mathrm{BW}}(\overline{b_{\mathrm{NLS}}(S_\xi w)})ww\,dx,$$

where we recall (5-1). By (5-2) we obtain that $\nabla_{\bar w} G_1(W) = -\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(b_{\mathrm{NLS}}(S_\xi w))\bar w$. We compute the gradient with respect $\bar w$ of the term $G_2(W)$. We have

$$\begin{aligned}
d_{\bar w} G_2(W)(\bar h) &= \frac{\mathrm{i}}{2}\int_{\mathbb{T}^d} \mathrm{Op}^{\mathrm{BW}}\Big(S_\xi(\bar w)S_\xi(\bar h)\frac{1}{|\xi|^2}\tilde\chi(\xi)\Big)ww\,dx\\
&= \frac{\mathrm{i}}{2}\frac{1}{(2\pi)^d}\sum_{\xi,\eta,\zeta\in\mathbb{Z}^d} \widehat{S_{(\xi+\zeta)/2}(\bar w)}(\xi-\eta-\zeta)\widehat{S_{(\xi+\zeta)/2}(\bar h)}(\eta)\hat w(\zeta)\\
&\qquad\qquad \times \frac{4}{|\zeta+\xi|^2}\tilde\chi\Big(\frac{\zeta+\xi}{2}\Big)\chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big)\hat w(-\xi) \quad \text{(by (3-6))}\\
&= 2\mathrm{i}\frac{1}{(2\pi)^d}\sum_{\xi,\eta,\zeta\in\mathbb{Z}^d}\tilde\chi\Big(\frac{\zeta+\xi}{2}\Big)\frac{1}{|\zeta+\xi|^2}\chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big)\chi_\epsilon\Big(\frac{2|\xi-\eta-\zeta|}{\langle\xi+\zeta\rangle}\Big)\chi_\epsilon\\
&\qquad\qquad \times\Big(\frac{2|\eta|}{\langle\xi+\zeta\rangle}\Big)\hat{\bar w}(\xi-\eta-\zeta)\hat{\bar h}(\eta)\hat w(\zeta)\hat w(-\xi) \quad \text{(by (5-1))}\\
&= 2\mathrm{i}\frac{1}{(2\pi)^d}\sum_{\eta\in\mathbb{Z}^d}\hat{\bar h}(-\eta)\sum_{\xi,\zeta\in\mathbb{Z}^d}\tilde\chi\Big(\frac{\zeta+\xi}{2}\Big)\frac{1}{|\zeta+\xi|^2}\chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big)\\
&\qquad\qquad \times\chi_\epsilon\Big(\frac{2|\xi+\eta-\zeta|}{\langle\xi+\zeta\rangle}\Big)\chi_\epsilon\Big(\frac{2|\eta|}{\langle\xi+\zeta\rangle}\Big)\hat{\bar w}(\xi+\eta-\zeta)\hat w(\zeta)\hat w(-\xi).
\end{aligned}$$

Recalling (3-69) and the computations above, after some changes of variables in the summations, we obtain

$$X_{\mathscr{B}_{\mathrm{NLS}}}(W) = \mathrm{Op}^{\mathrm{BW}}(B_{\mathrm{NLS}}(S_\xi W; x,\xi))W + R_1(W),$$

where the remainder $R_1(W)$ has the form $(R_1^+(W), \overline{R_1^+(W)})^T$, where (recall (3-5))

$$\begin{aligned}
\widehat{(R_1^+(W))}(\xi) &= \frac{1}{(2\pi)^d}\sum_{\eta,\zeta\in\mathbb{Z}^d} r_1(\xi,\eta,\zeta)\hat w(\xi-\eta-\zeta)\hat{\bar w}(\eta)\hat w(\zeta), \quad \xi\in\mathbb{Z}^d,\\
r_1(\xi,\eta,\zeta) &= -\frac{2}{|2\zeta-\xi+\eta|^2}\tilde\chi\Big(\frac{2\zeta-\xi+\eta}{2}\Big)\chi_\epsilon\Big(\frac{|\eta-\xi|}{\langle 2\zeta-\xi+\eta\rangle}\Big)\chi_\epsilon\Big(\frac{2|\xi|}{\langle\xi-\eta-2\zeta\rangle}\Big)\chi_\epsilon\Big(\frac{2|\eta|}{\langle\xi-\eta-2\zeta\rangle}\Big).
\end{aligned}$$

One can check, for $0<\epsilon<1$ small enough, $|\xi|+|\eta| \ll |\xi-\eta-\zeta| \sim |\zeta|$. Therefore the coefficient $r_1(\xi,\eta,\zeta)$ satisfies (3-48). Here we really need the truncation operator S_ξ: if you don't insert it in the definition of $\mathscr{B}_{\mathrm{NLS}}$ (see (5-3)) then R_1 is not a regularizing operator. Furthermore this truncation does not affect the leading term: Define the operator

$$R_2(W) = \begin{pmatrix} R_2^+(W)\\ \overline{R_2^+(W)}\end{pmatrix} := \mathrm{Op}^{\mathrm{BW}}\big(B_{\mathrm{NLS}}(S_\xi W; x,\xi) - B_{\mathrm{NLS}}(W;x,\xi)\big)W.$$

We are going to prove that R_2 is also a regularizing operator. By an explicit computation using (3-6), (5-1) and (5-2) one can check that

$$\begin{aligned}
\widehat{(R_2^+(W))}(\xi) &= \frac{1}{(2\pi)^d}\sum_{\eta,\zeta\in\mathbb{Z}^d} r_2(\xi,\eta,\zeta)\hat w(\xi-\eta-\zeta)\hat{\bar w}(\eta)\hat w(\zeta), \quad \xi\in\mathbb{Z}^d,\\
r_2(\xi,\eta,\zeta) &= -\frac{1}{|\xi+\zeta|^2}\tilde\chi\Big(\frac{\xi+\zeta}{2}\Big)\chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big)\Big(1-\chi_\epsilon\Big(\frac{|\xi-\eta-\zeta|}{\langle\xi+\zeta\rangle}\Big)\chi_\epsilon\Big(\frac{|\eta|}{\langle\xi+\zeta\rangle}\Big)\Big).
\end{aligned}$$

We write $1 \cdot r_2(\xi, \eta, \zeta)$ and we use the partition of the unity in (3-49). Hence using (3-5) one can check that each summand satisfies the bound in (3-48). Therefore the operator $Q_G := R_1 + R_2$ has the form (5-7) and (5-6) is proved. The estimates (5-8) follow by Lemma 3.7. We note that

$$d_W\big(\mathrm{Op}^{\mathrm{BW}}(B_{\mathrm{NLS}}(W; x, \xi))W\big)[h] = \mathrm{Op}^{\mathrm{BW}}(B_{\mathrm{NLS}}(W; x, \xi))h + \mathrm{Op}^{\mathrm{BW}}(d_W B_{\mathrm{NLS}}(W; x, \xi)[h])W.$$

Then the estimates (5-9) with $k = 0, 1$ follow by using (5-8), the explicit formula of $B(W; x, \xi)$ in (5-2) and Lemma 3.1. Reasoning similarly one can prove (5-9) with $k = 2, 3$. □

In the next proposition we define the changes of coordinates generated by the Hamiltonian vector fields $X_{\mathscr{B}_{\mathrm{NLS}}}$ and $X_{\mathscr{B}_{\mathrm{KG}}}$ and we study their properties as maps on Sobolev spaces.

Proposition 5.2. *For any $s \geq s_0 > 2d + 2$ there is $r_0 > 0$ such that for $0 \leq r \leq r_0$ and*

$$W = \begin{bmatrix} w \\ \bar{w} \end{bmatrix} \in B_r(H^s(\mathbb{T}^d; \mathbb{C}^2))$$

the following holds. Define

$$Z := \Phi_{\mathscr{B}_\star}(W) := W + X_{\mathscr{B}_\star}(W), \tag{5-11}$$

where $\star \in \{\mathrm{NLS}, \mathrm{KG}\}$ (recall (5-3), (5-5)). Then one has

$$\|Z\|_{H^s} \leq \|w\|_{H^s}(1 + C\|w\|^2_{H^s}) \tag{5-12}$$

for some $C > 0$ depending on s, and

$$W = Z - X_{\mathscr{B}_\star}(Z) + r(w), \tag{5-13}$$

where

$$\|r(w)\|_{H^s} \lesssim \|w\|^5_{H^s}. \tag{5-14}$$

Proof. By (5-11) we can write

$$W = Z - X_{\mathscr{B}_\star}(W) = Z - X_{\mathscr{B}_\star}(Z) + [X_{\mathscr{B}_\star}(W) - X_{\mathscr{B}_\star}(Z)].$$

By using estimates (5-9) or (5-10) one can deduce that $X_{\mathscr{B}_\star}(W) - X_{\mathscr{B}_\star}(Z)$ satisfies the bound (5-14). The bound (5-12) follows by Lemma 5.1. □

5B. *Conjugations.*

Recalling (1-25) and (4-23) we set

$$\mathscr{H}^{(\leq 4)}_{\mathrm{NLS}}(W) := \mathscr{H}^{(2)}_{\mathrm{NLS}}(W) + \mathscr{H}^{(4)}_{\mathrm{NLS}}(W), \quad \mathscr{H}^{(2)}_{\mathrm{NLS}}(Z) := \int_{\mathbb{T}^d} \Lambda_{\mathrm{NLS}} z \cdot \bar{z}\, dx. \tag{5-15}$$

Analogously, recalling (4-27) and (1-4), we set

$$\mathscr{H}^{(\leq 4)}_{\mathrm{KG}}(W) := \mathscr{H}^{(2)}_{\mathrm{KG}}(W) + \mathscr{H}^{(4)}_{\mathrm{KG}}(W), \quad \mathscr{H}^{(2)}_{\mathrm{KG}}(Z) := \int_{\mathbb{T}^d} \Lambda_{\mathrm{KG}} z \cdot \bar{z}\, dx. \tag{5-16}$$

In the following lemma we study how the Hamiltonian vector fields $X_{\mathscr{H}^{(\leq 4)}_{\mathrm{NLS}}}(W)$ in (5-15) and $X_{\mathscr{H}^{(\leq 4)}_{\mathrm{KG}}}(W)$ in (5-16) transform under the change of variables given by the previous lemma.

Lemma 5.3. *Let $s_0 > 2d+4$. Then for any $s \geq s_0$ there is $r_0 > 0$ such that for all $0 < r \leq r_0$ and*

$$Z = \begin{bmatrix} z \\ \bar{z} \end{bmatrix} \in B_r(H^s(\mathbb{T}^d; \mathbb{C}^2))$$

the following holds. Consider the Hamiltonian $\mathscr{B}_\star$ with $\star \in \{\mathrm{NLS}, \mathrm{KG}\}$ (recall (5-3), (5-5)) and the Hamiltonian $\mathscr{H}_\star^{(\leq 4)}$ (see (5-15), (5-16)). Then

$$d_W \Phi_{\mathscr{B}_\star}(W)[X_{\mathscr{H}_\star^{(\leq 4)}}(W)] = X_{\mathscr{H}_\star^{(\leq 4)}}(Z) + [X_{\mathscr{B}_\star}(Z), X_{\mathscr{H}_\star^{(2)}}(Z)] + R_5(Z), \tag{5-17}$$

where the remainder R_5 satisfies

$$\|R_5(Z)\|_{H^s} \lesssim \|z\|_{H^s}^5, \tag{5-18}$$

and $[\,\cdot\,,\,\cdot\,]$ is the nonlinear commutator defined in (3-73).

Proof. We prove the statement in the case $\mathscr{B}_\star = \mathscr{B}_{\mathrm{NLS}}$ and $\mathscr{H}_\star^{(\leq 4)} = \mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}$; the KG-case is similar. One can check that (5-17) follows by setting

$$\begin{aligned} R_5 := {} & d_W X_{\mathscr{B}_{\mathrm{NLS}}}(W)[X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(W) - X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(Z)] && (5\text{-}19) \\ & + (d_W X_{\mathscr{B}_{\mathrm{NLS}}}(W) - d_W X_{\mathscr{B}_{\mathrm{NLS}}}(Z))[X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(Z)] && (5\text{-}20) \\ & + X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(W) - X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(Z) + d_W X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(Z)[X_{\mathscr{B}_{\mathrm{NLS}}}(Z)] && (5\text{-}21) \\ & + [X_{\mathscr{B}_{\mathrm{NLS}}}(Z), X_{\mathscr{H}_{\mathrm{NLS}}^{(4)}}(Z)]. && (5\text{-}22) \end{aligned}$$

We are left to prove that R_5 satisfies (5-18). We start from the term in (5-19). First of all we note that

$$X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(W) - X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(Z) = -\mathrm{i} E \Lambda_{\mathrm{NLS}}(W - Z) + X_{\mathscr{H}_{\mathrm{NLS}}^{(4)}}(W) - X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(Z),$$

where we used that $X_{\mathscr{H}_{\mathrm{NLS}}^{(2)}}(W) = -\mathrm{i} E \Lambda_{\mathrm{NLS}} W$. By Proposition 5.2, (4-23) and (5-9) we deduce that

$$\|X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(W) - X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(Z)\|_{H^s} \lesssim \|w\|_{H^s}^3.$$

Hence using again the bounds (5-9) we obtain

$$\|d_W X_{\mathscr{B}_{\mathrm{NLS}}}(W)[X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq 4)}}(W) - X_{\mathscr{H}_{\mathrm{NLS}}^{(\leq)}}(Z)]\|_{H^s} \lesssim \|w\|_{H^s}^5.$$

Reasoning in the same way, using also (5-13), one can check that the terms in (5-20), (5-21), (5-22) satisfy the same quintic estimates. □

In the next lemma we study the structure of the cubic terms in the vector field in (5-17) in the NLS case.

Lemma 5.4. *Consider the Hamiltonian $\mathscr{B}_{\mathrm{NLS}}(W)$ in (5-3) and recall (4-23), (5-15). Then we have*

$$X_{\mathscr{H}_{\mathrm{NLS}}^{(4)}}(Z) + [X_{\mathscr{B}_{\mathrm{NLS}}}(Z), X_{\mathscr{H}_{\mathrm{NLS}}^{(2)}}(Z)] = -\mathrm{i} E \,\mathrm{Op}^{\mathrm{BW}} \begin{pmatrix} 2|z|^2 & 0 \\ 0 & 2|z|^2 \end{pmatrix} Z + Q_{\mathrm{H}_{\mathrm{NLS}}^{(4)}}(Z), \tag{5-23}$$

where the remainder $Q_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}$ has the form $Q_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z) = (Q^{+}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z), \overline{Q^{+}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z)})^T$ and

$$\widehat{(Q^{+}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z))}(\xi) = \frac{1}{(2\pi)^d} \sum_{\eta,\zeta\in\mathbb{Z}^d} \mathtt{q}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(\xi,\eta,\zeta)\hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}(\zeta), \quad \xi\in\mathbb{Z}^d, \tag{5-24}$$

with symbol satisfying

$$|\mathtt{q}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(\xi,\eta,\zeta)| \lesssim \frac{\max_2\{\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\}^4}{\max_1\{\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\}^2}. \tag{5-25}$$

Proof. We start by considering the commutator between $X_{\mathscr{B}_{\mathrm{NLS}}}$ and $X_{\mathscr{H}^{(2)}_{\mathrm{NLS}}}$. First of all notice that (see (5-6) and (5-2))

$$X_{\mathscr{B}_{\mathrm{NLS}}}(Z) = \begin{pmatrix} X^{+}_{\mathscr{B}_{\mathrm{NLS}}}(Z) \\ \overline{X^{+}_{\mathscr{B}_{\mathrm{NLS}}}(Z)} \end{pmatrix}, \quad X^{+}_{\mathscr{B}_{\mathrm{NLS}}}(Z) := \mathrm{Op}^{\mathrm{BW}}\left(\frac{z^2}{2|\xi|^2}\tilde{\chi}(\xi)\right)[\bar{z}] + Q^{+}_{\mathscr{B}_{\mathrm{NLS}}}(Z),$$

and hence (recall (3-6)), for $\xi\in\mathbb{Z}^d$,

$$\begin{aligned} &\widehat{(X^{+}_{\mathscr{B}_{\mathrm{NLS}}}(Z))}(\xi) \\ &\quad = \frac{1}{(2\pi)^d} \sum_{\eta,\zeta\in\mathbb{Z}^d} \hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}(\zeta)\left[\frac{2}{|\xi+\eta|^2}\tilde{\chi}\left(\frac{\xi+\eta}{2}\right)\chi_\epsilon\left(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\right) + q_{\mathscr{B}_{\mathrm{NLS}}}(\xi,\eta,\zeta)\right], \end{aligned} \tag{5-26}$$

where $q_{\mathscr{B}_{\mathrm{NLS}}}(\zeta,\eta,\zeta)$ satisfies the bound in (3-48). Hence, by using formulas (1-25), (5-26), (3-73), one obtains

$$X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(Z) + [X_{\mathscr{B}_{\mathrm{NLS}}}(Z), X_{\mathscr{H}^{(2)}_{\mathrm{NLS}}}(Z)] = \begin{pmatrix} \mathscr{C}^{+}(Z) \\ \overline{\mathscr{C}^{+}(Z)} \end{pmatrix},$$

$$\widehat{(\mathscr{C}^{+}(Z))}(\xi) = \frac{-1}{(2\pi)^d} \sum_{\eta,\zeta\in\mathbb{Z}^d} \mathrm{i}\mathtt{c}(\xi,\eta,\zeta)\hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}(\zeta),$$

where

$$\begin{aligned} \mathtt{c}(\xi,\eta,\zeta) = 1 + &\left[\frac{2}{|\xi+\eta|^2}\tilde{\chi}\left(\frac{\xi+\eta}{2}\right)\chi_\epsilon\left(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\right) + q_{\mathscr{B}_{\mathrm{NLS}}}(\xi,\eta,\zeta)\right] \\ &\times [\Lambda_{\mathrm{NLS}}(\xi-\eta-\zeta) - \Lambda_{\mathrm{NLS}}(\eta) + \Lambda_{\mathrm{NLS}}(\zeta) - \Lambda_{\mathrm{NLS}}(\xi)]. \end{aligned} \tag{5-27}$$

We need to prove that this can be written as the right-hand side of (5-23). First we note that the term in (5-27),

$$q_{\mathscr{B}_{\mathrm{NLS}}}(\xi,\eta,\zeta)[\Lambda_{\mathrm{NLS}}(\xi-\eta-\zeta) - \Lambda_{\mathrm{NLS}}(\eta) + \Lambda_{\mathrm{NLS}}(\zeta) - \Lambda_{\mathrm{NLS}}(\xi)], \tag{5-28}$$

can be absorbed in $\mathtt{R}_1$ since (5-28) satisfies the same bound as in (5-25). Moreover, using (1-25) and (1-5), we have that the coefficients

$$\frac{2}{|\xi+\eta|^2}\tilde{\chi}\left(\frac{\xi+\eta}{2}\right)\chi_\epsilon\left(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\right)[\widehat{V}(\xi-\eta-\zeta) - \widehat{V}(\eta) + \widehat{V}(\zeta) - \widehat{V}(\xi)]$$

satisfy the bound in (5-25) by using also Lemma 3.8. Therefore the corresponding operator contributes to $\mathtt{R}_1$. The same holds for the operator corresponding to the coefficients

$$\frac{2}{|\xi+\eta|^2}\tilde{\chi}\left(\frac{\xi+\eta}{2}\right)\chi_\epsilon\left(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\right)[|\xi-\eta-\zeta|^2 + |\zeta|^2].$$

We are left with the most relevant terms in (5-27) containing the highest frequencies η and ξ. We have

$$\frac{-2(|\xi|^2+|\eta|^2)}{|\xi+\eta|^2}\chi_\epsilon\Big(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\Big)\tilde\chi\Big(\frac{\xi+\eta}{2}\Big)=-\chi_\epsilon\Big(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\Big)-r_1(\xi,\eta,\zeta),$$

where

$$r_1(\xi,\eta,\zeta)=\Big(\tilde\chi\Big(\frac{\xi+\eta}{2}\Big)-1\Big)\chi_\epsilon\Big(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\Big)+\frac{|\xi-\eta|^2}{|\xi+\eta|^2}\tilde\chi\Big(\frac{\xi+\eta}{2}\Big)\chi_\epsilon\Big(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\Big).$$

Again we note that the coefficients $r_1(\xi,\eta,\zeta)$, using Lemma 3.8 and the definition of $\tilde\chi$ below (5-2), satisfy (5-25). Then it remains to study the operator $\mathscr{R}^+(Z)$ with

$$\widehat{(\mathscr{R}^+(Z))}(\xi):=-\frac{1}{(2\pi)^d}\sum_{\eta,\zeta\in\mathbb{Z}^d}\mathrm{i}\Big(1-\chi_\epsilon\Big(\frac{|\xi-\eta|}{\langle\xi+\eta\rangle}\Big)\Big)\hat z(\xi-\eta-\zeta)\hat{\bar z}(\eta)\hat z(\zeta).$$

By formula (4-13) and (3-6), $\mathscr{R}^+(Z)=-\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(2|z|^2)z+Q_3^+(U)$, where Q_3 satisfies (4-15), (4-16). □

In the next lemma we study the structure of the the cubic terms in the vector field in (5-17) in the KG case.

Lemma 5.5. *Consider the Hamiltonian $\mathscr{B}_{\mathrm{KG}}(W)$ in* (5-5) *and recall* (4-27), (5-16). *Then we have*

$$X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(Z)+[X_{\mathscr{B}_{\mathrm{KG}}}(Z),X_{\mathscr{H}^{(2)}_{\mathrm{KG}}}(Z)]=-\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(a_0(x,\xi)))Z+Q_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(Z),\tag{5-29}$$

where the symbol $a_0(x,\xi)=a_0(u,x,\xi)$ is as in (4-24) *and the remainder $Q_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(Z)$ has the form $(Q^+_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(Z),\overline{Q^+_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(Z)})^T$, with*

$$\widehat{Q^+_{\mathtt{H}^{(4)}_{\mathrm{KG}}}}(\xi)=(2\pi)^{-d}\sum_{\substack{\sigma_1,\sigma_2,\sigma_3\in\{\pm\}\\ \eta,\zeta\in\mathbb{Z}^d}}\mathtt{q}^{\sigma_1,\sigma_2,\sigma_3}_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(\xi,\eta,\zeta)\widehat{z^{\sigma_1}}(\xi-\eta-\zeta)\widehat{z^{\sigma_2}}(\eta)\widehat{z^{\sigma_3}}(\zeta)\tag{5-30}$$

for some $\mathtt{q}^{\sigma_1,\sigma_2,\sigma_3}_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(\xi,\eta,\zeta)\in\mathbb{C}$ satisfying

$$|\mathtt{q}^{\sigma_1,\sigma_2,\sigma_3}_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(\xi,\eta,\zeta)|\lesssim\frac{\max_2\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}^\mu}{\max\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}}\tag{5-31}$$

for some $\mu>1$.

Proof. Using (5-6) (with $\mathscr{B}=\mathscr{B}_{\mathrm{KG}}$) we can note that

$$[X_{\mathscr{B}_{\mathrm{KG}}}(Z),X_{\mathscr{H}^{(2)}_{\mathrm{KG}}}(Z)]=[\mathrm{Op}^{\mathrm{BW}}(B_{\mathrm{KG}}(Z;x,\xi)),X_{\mathscr{H}^{(2)}_{\mathrm{KG}}}(Z)]+R_2(Z),\tag{5-32}$$

where $R_2(Z)=(R_2^+(Z),\overline{R_2^+(Z)})^T$, with

$$\begin{aligned}\widehat{(R_2^+(Z))}(\xi)&=(2\pi)^{-d}\sum_{\substack{\sigma_1,\sigma_2,\sigma_3\in\{\pm\}\\ \eta,\zeta\in\mathbb{Z}^d}}\mathtt{r}_2^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\widehat{z^{\sigma_1}}(\xi-\eta-\zeta)\widehat{z^{\sigma_2}}(\eta)\widehat{z^{\sigma_3}}(\zeta),\quad \xi\in\mathbb{Z}^d,\\ \mathtt{r}_2^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)&:=\mathtt{q}^{\sigma_1,\sigma_2,\sigma_3}_{\mathscr{B}_{\mathrm{KG}}}(\xi,\eta,\zeta)[\sigma_1\Lambda_{\mathrm{KG}}(\xi-\eta-\zeta)+\sigma_2\Lambda_{\mathrm{KG}}(\eta)+\sigma_3\Lambda_{\mathrm{KG}}(\zeta)-\Lambda_{\mathrm{KG}}(\xi)],\end{aligned}\tag{5-33}$$

where the coefficients are defined in (5-7). The remainder R_2 has the form (5-30) and we have that the coefficients $\mathtt{r}_2^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)$ satisfy the bound (5-31). On the other hand, recalling (5-4), (3-73), we have

$$[\mathrm{Op}^{\mathrm{BW}}(B_{\mathrm{KG}}(Z;x,\xi)), X_{\mathscr{H}^{(2)}_{\mathrm{KG}}}(Z)] = R_3(Z)+R_4(Z), \quad R_j(Z) = \begin{pmatrix} R_j^+(Z) \\ \overline{R_j^+(Z)} \end{pmatrix}, \quad j=3,4, \tag{5-34}$$

where

$$R_3^+(Z) := \mathrm{Op}^{\mathrm{BW}}(b_{\mathrm{KG}}(Z;x,\xi))[\mathrm{i}\Lambda_{\mathrm{KG}}\bar z] + \mathrm{i}\Lambda_{\mathrm{KG}}\,\mathrm{Op}^{\mathrm{BW}}(b_{\mathrm{KG}}(Z;x,\xi))[\bar z], \tag{5-35}$$

$$R_4^+(Z) := \mathrm{Op}^{\mathrm{BW}}((d_Z b_{\mathrm{KG}})(z;x,\xi)[X_{\mathscr{H}^{(2)}_{\mathrm{KG}}}(Z)])[\bar z]. \tag{5-36}$$

By Remark 4.5 and (3-6) we get

$$\widehat{R_4^+}(\xi) = (2\pi)^{-d} \sum_{\sigma_1,\sigma_2\in\{\pm\}\eta,\zeta\in\mathbb{Z}^d} a_0^{\sigma_1,\sigma_2}\Big(\xi-\zeta,\eta,\frac{\xi+\zeta}{2}\Big)\frac{1}{2\Lambda_{\mathrm{KG}}((\xi+\zeta)/2)} \times \chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big)[-\mathrm{i}\sigma_1\Lambda_{\mathrm{KG}}(\xi-\eta-\zeta)-\mathrm{i}\sigma_2\Lambda_{\mathrm{KG}}(\eta)]\widehat{z^{\sigma_1}}(\xi-\eta-\zeta)\widehat{z^{\sigma_2}}(\eta)\hat{\bar z}(\zeta).$$

Using the explicit form of the coefficients of R_4^+ and Lemma 3.8 one can conclude that the operator R_4^+ has the form (5-30) with coefficients satisfying (5-31). To summarize, by (5-32) and (5-34), we have obtained (recall also (4-28), (4-26))

$$\text{LHS of (5-29)} = \mathrm{Op}^{\mathrm{BW}}\begin{pmatrix} -\mathrm{i}a_0(x,\xi) & 0 \\ 0 & \mathrm{i}a_0(x,\xi)\end{pmatrix} Z + F_3(Z)+Q_3(Z)+R_2(Z)+R_4(Z), \tag{5-37}$$

where R_4 is in (5-36), R_2 is in (5-33), $Q_3(Z)$ is in (4-28) and

$$F_3(Z) = \begin{pmatrix} F_3^+(Z) \\ \overline{F_3^+(Z)}\end{pmatrix}, \quad F_3^+(Z) = -\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(a_0(x,\xi))[\bar z] + R_3^+(Z), \tag{5-38}$$

where R_3^+ is in (5-35). By the discussion above and by Lemma 4.4 we have that the remainders R_2, R_4 and Q_3 have the form (5-30) with coefficients satisfying (5-31). To conclude the prove we need to show that F_3 has the same property. This will be a consequence of the choice of the symbol $b_{\mathrm{KG}}(W;x,\xi)$ in (5-4). Indeed, by (5-4), Remark 4.5, (5-38), (5-35), we have

$$\widehat{F_3^+}(\xi) = (2\pi)^{-d} \sum_{\substack{\sigma_1,\sigma_2\in\{\pm\}\\ \eta,\zeta\in\mathbb{Z}^d}} \mathtt{f}_3^{\sigma_1,\sigma_2,-}(\xi,\eta,\zeta)\widehat{z^{\sigma_1}}(\xi-\eta-\zeta)\widehat{z^{\sigma_2}}(\eta)\hat{\bar z}(\zeta),$$

where

$$\mathtt{f}_3^{\sigma_1,\sigma_2,-}(\xi,\eta,\zeta) := a_0^{\sigma_1,\sigma_2}\Big(\xi-\zeta,\eta,\frac{\xi+\zeta}{2}\Big)\,\mathrm{i}\Big[\frac{\Lambda_{\mathrm{KG}}(\xi)+\Lambda_{\mathrm{KG}}(\zeta)}{2\Lambda_{\mathrm{KG}}((\xi+\zeta)/2)}-1\Big]\chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\zeta+\zeta\rangle}\Big). \tag{5-39}$$

By Taylor expanding the symbol $\Lambda_{\mathrm{KG}}(\xi)$ in (1-4) (see also Remark 4.5) one deduces that

$$\Big|a_0^{\sigma_1,\sigma_2}\Big(\xi-\zeta,\eta,\frac{\xi+\zeta}{2}\Big)\,\mathrm{i}\Big[\frac{\Lambda_{\mathrm{KG}}(\xi)+\Lambda_{\mathrm{KG}}(\zeta)}{2\Lambda_{\mathrm{KG}}((\xi+\zeta)/2)}-1\Big]\Big| \lesssim \frac{|\xi-\zeta|}{(\langle\xi\rangle+\langle\zeta\rangle)^{3/2}}.$$

Therefore, using Lemma 3.8, we have that the coefficients $\mathtt{f}_3^{\sigma_1,\sigma_2,-}(\xi,\eta,\zeta)$ in (5-39) satisfy (5-31). This implies (5-29). $\square$

6. Diagonalization

6A. ***Diagonalization of the NLS.*** In this section we diagonalize the system (4-12). We first diagonalize the matrix $E(\mathbb{1}+A_2(x))$ in (4-12) by means of a change of coordinates as the ones made in the papers [Feola and Iandoli 2021; 2022]. After that we diagonalize the matrix of symbols of order 0 at homogeneity 3, by means of an approximately *symplectic* change of coordinates. Throughout the rest of the section we shall assume the following.

Hypothesis 6.1. We restrict the solution of (NLS) on the interval of times $[0, T)$, with T such that

$$\sup_{t\in[0,T)} \|u(t,x)\|_{H^s} \le \epsilon, \quad \|u_0(x)\|_{H^s} \le \tfrac{1}{4}\epsilon.$$

Note that such a time $T > 0$ exists thanks to the local existence theorem in [Feola and Iandoli 2022].

6A1. *Diagonalization at order 2.* We consider the matrix $E(\mathbb{1} + A_2(x))$ in (4-12). We define

$$\lambda_{\rm NLS}(x) := \lambda_{\rm NLS}(U; x) := \sqrt{1+2|u|^2[h'(|u|^2)]^2}, \quad a_2^{(1)}(x) := \lambda_{\rm NLS}(x) - 1, \tag{6-1}$$

and we note that $\pm\lambda_{\rm NLS}(x)$ are the eigenvalues of the matrix $E(\mathbb{1} + A_2(x))$. We denote by S matrix of the eigenvectors of $E(\mathbb{1} + A_2(x))$; more explicitly

$$\begin{aligned} S &= \begin{pmatrix} s_1 & s_2 \\ \bar{s}_2 & s_1 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} s_1 & -s_2 \\ -\bar{s}_2 & s_1 \end{pmatrix}, \\ s_1(x) &:= \frac{1+|u|^2[h'(|u|^2)]^2 + \lambda_{\rm NLS}(x)}{\sqrt{2\lambda_{\rm NLS}(x)(1+[h'(|u|^2)]^2|u|^2+\lambda_{\rm NLS}(x))}}, \\ s_2(x) &:= \frac{-u^2[h'(|u|^2)]^2}{\sqrt{2\lambda_{\rm NLS}(x)(1+[h'(|u|^2)]^2|u|^2+\lambda_{\rm NLS}(x))}}. \end{aligned} \tag{6-2}$$

Since $\pm\lambda_{\rm NLS}(x)$ are the eigenvalues and $S(x)$ is the matrix of eigenvectors of $E(\mathbb{1} + A_2(x))$ we have

$$S^{-1}E(\mathbb{1} + A_2(x))S = E\,\mathrm{diag}(\lambda_{\rm NLS}(x)), \quad s_1^2 - |s_2|^2 = 1, \tag{6-3}$$

where we have used the notation (4-9). In the lemma we estimate the seminorms of the symbols defined above.

Lemma 6.2. *Let* $\mathbb{N} \ni s_0 > d$. *The symbols* $a_2^{(1)}$ *defined in* (6-1), $s_1 - 1$ *and* s_2 *defined in* (6-2) *satisfy the following estimate*

$$|a_2^{(1)}|_{\mathcal{N}_p^0} + |s_1 - 1|_{\mathcal{N}_p^0} + |s_2|_{\mathcal{N}_p^0} \lesssim \|u\|^6_{H^{p+s_0}}, \quad p + s_0 \le s, \ p \in \mathbb{N}.$$

Proof. The proof follows by using the estimate (4-14) on the symbols in (4-10), the fact that $h'(s) \sim s$ when $s \sim 0$, $\|u\|_s \ll 1$, and the explicit expression (6-1), (6-2). □

We now study how the system (4-12) transforms under the maps

$$\Phi_{\rm NLS} := \Phi_{\rm NLS}(U) := \mathrm{Op}^{\rm BW}(S^{-1}(U; x)), \quad \Psi_{\rm NLS} := \Psi_{\rm NLS}(U) := \mathrm{Op}^{\rm BW}(S(U; x)). \tag{6-4}$$

Lemma 6.3. *Let*

$$U = \begin{bmatrix} u \\ \bar{u} \end{bmatrix}$$

be a solution of (4-12) *and assume Hypothesis 6.1. Then for any* $s \geq 2s_0 + 2$, $\mathbb{N} \ni s_0 > d$, *we have the following*:

(i) *One has the upper bound*

$$\begin{aligned} \|\Phi_{\mathrm{NLS}}(U)W\|_{H^s} + \|\Psi_{\mathrm{NLS}}(U)W\|_{H^s} &\leq \|W\|_{H^s}(1 + C\|u\|^6_{H^{2s_0}}), \\ \|(\Phi_{\mathrm{NLS}}(U) - \mathbb{1})W\|_{H^s} + \|(\Psi_{\mathrm{NLS}}(U) - \mathbb{1})W\|_{H^s} &\lesssim \|W\|_{H^s}\|u\|^6_{H^{2s_0}} \quad \text{for all } W \in H^s(\mathbb{T}^d; \mathbb{C}), \end{aligned} \tag{6-5}$$

where the constant C depends on s.

(ii) *One has* $\Psi_{\mathrm{NLS}}(U) \circ \Phi_{\mathrm{NLS}}(U) = \mathbb{1} + R(u)$, *where R is a real-to-real remainder of the form* (3-52) *satisfying*

$$\|R(u)W\|_{H^{s+2}} \lesssim \|W\|_{H^s}\|u\|^6_{H^{2s_0+2}}. \tag{6-6}$$

The map $\mathbb{1} + R(u)$ *is invertible with inverse* $(\mathbb{1} + R(u))^{-1} := (1 + \widetilde{R}(u))$, *with* $\widetilde{R}(u)$ *of the form* (3-52) *and*

$$\|\widetilde{R}(u)W\|_{H^{s+2}} \lesssim \|W\|_{H^s}\|u\|^6_{H^{2s_0+2}}; \tag{6-7}$$

as a consequence the map Φ_{NLS} *is invertible and* $\Phi^{-1}_{\mathrm{NLS}} = (1 + \widetilde{R})\Psi_{\mathrm{NLS}}$ *with estimates*

$$\|\Phi^{-1}_{\mathrm{NLS}}(U)W\|_{H^s} \leq \|W\|_{H^s}(1 + C\|u\|^6_{H^{2s_0+2}}), \tag{6-8}$$

where the constant C depends on s.

(iii) *For any* $t \in [0, T)$, *one has* $\partial_t \Phi_{\mathrm{NLS}}(U)[\,\cdot\,] = \mathrm{Op}^{\mathrm{BW}}(\partial_t S^{-1}(U; x))$ *and*

$$|\partial_t S^{-1}(U; x)|_{\mathcal{N}^0_{s_0}} \lesssim \|u\|^6_{H^{2s_0+2}}, \quad \|\partial_t \Phi_{\mathrm{NLS}}(U)V\|_{H^s} \lesssim \|W\|_{H^s}\|u\|^6_{H^{2s_0+2}}. \tag{6-9}$$

Proof. (i) The bounds (6-5) follow by (3-10) and Lemma 6.2.

(ii) We apply Proposition 3.2 to the maps in (6-4); in particular the first part of the item follows by using the expansion (3-21) and recalling that symbols $s_1(x)$ and $s_2(x)$ do not depend on ξ. Inequality (6-7) is obtained by Neumann series by using that (see Hypothesis 6.1) $\|u\|_{H^s} \ll 1$.

(iii) We note that $\partial_t s_1(x, \xi) = (\partial_u s_1)(u; x, \xi)[\dot{u}] + (\partial_{\bar{u}} s_1)(u; x, \xi)[\dot{\bar{u}}]$. Since u solves (4-12) and satisfies Hypothesis 6.1, then using Lemma 3.1 and (4-17) we deduce that $\|\dot{u}\|_{H^s} \lesssim \|u\|_{H^{s+2}}$. Hence the estimates (6-9) follow by direct inspection by using the explicit structure of the symbols s_1, s_2 in (6-2), Lemma 3.6 and (3-10). ☐

We are now in position to state the following proposition.

Proposition 6.4 (diagonalization at order 2). *Consider the system* (4-12) *and set*

$$W = \Phi_{\mathrm{NLS}}(U)U, \tag{6-10}$$

with Φ_{NLS} *defined in* (6-4)*. Then* W *solves the equation*

$$\dot W = -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(1+a_2^{(1)}(U;x))|\xi|^2)W - \mathrm{i}EV*W \\ - \mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\vec a_1^{(1)}(U;x)\cdot\xi))W + X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(W) + R^{(1)}(U), \tag{6-11}$$

where the vector field $X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}$ *is defined in* (4-13)*. The symbols* $a_2^{(1)}$ *and* $\vec a_1^{(1)}\cdot\xi$ *are **real**-valued and satisfy the estimates*

$$\begin{aligned} |a_2^{(1)}|_{\mathcal{N}_p^0} &\lesssim \|u\|^6_{H^{p+s_0}} && \text{for all } p+s_0\le s,\ p\in\mathbb{N},\\ |\vec a_1^{(1)}\cdot\xi|_{\mathcal{N}_p^1} &\lesssim \|u\|^6_{H^{p+s_0+1}} && \text{for all } p+s_0+1\le s,\ p\in\mathbb{N}, \end{aligned} \tag{6-12}$$

where we have chosen $s_0>d$*. The remainder* $R^{(1)}$ *has the form* $(R^{(1,+)},\overline{R^{(1,+)}})^T$*. Moreover, for any* $s>2d+2$*, it satisfies the estimate*

$$\|R^{(1)}(U)\|_{H^s}\lesssim \|U\|^7_{H^s}. \tag{6-13}$$

Proof. The function W defined in (6-10) satisfies

$$\begin{aligned} \dot W &= [\partial_t\Phi_{\mathrm{NLS}}(U)]U+\Phi_{\mathrm{NLS}}(U)\dot U\\ &= -\Phi_{\mathrm{NLS}}(U)\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}((\mathbb{1}+A_2(U))|\xi|^2)\Psi_{\mathrm{NLS}}(U)W-\Phi_{\mathrm{NLS}}(U)\mathrm{i}EV*\Psi_{\mathrm{NLS}}(U)W && \text{(6-14)}\\ &\quad -\mathrm{i}\Phi_{\mathrm{NLS}}(U)\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\vec a_1(U)\cdot\xi))\Psi_{\mathrm{NLS}}(U)W && \text{(6-15)}\\ &\quad +\Phi_{\mathrm{NLS}}(U)X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(U) && \text{(6-16)}\\ &\quad +\Phi_{\mathrm{NLS}}(U)R(U)+\mathrm{Op}^{\mathrm{BW}}(\partial_tS^{-1}(U))U && \text{(6-17)}\\ &\quad -\Phi_{\mathrm{NLS}}(U)\mathrm{i}[E\,\mathrm{Op}^{\mathrm{BW}}((\mathbb{1}+A_2(U))|\xi|^2)+\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\vec a_1\cdot\xi))+EV*]\widetilde R(U)\Psi_{\mathrm{NLS}}(U)W, && \text{(6-18)} \end{aligned}$$

where we have used items (ii) and (iii) of Lemma 6.3.

We are going to analyze each term in the right-hand side of the equation above. Because of estimates (6-7), (6-5) (applied to the map Φ_{NLS}), Lemma 6.2 (applied to the symbols a_2, b_2 and $\vec a_1\cdot\xi$) and finally item (ii) of Lemma 3.1, we may absorb term (6-18) in the remainder $R^{(1)}(U)$ verifying (6-13). The term in (6-17) may be absorbed in $R^{(1)}(U)$ as well because of (4-17) and (6-5) for the first term and because of (6-9) and item (ii) of Lemma 3.1 for the second one.

We study the first term in (6-14). We recall (6-4) and (6-2), we apply Proposition 3.2 and we get, by direct inspection, that the new term, modulo contribution that may be absorbed in $R^{(1)}(U)$, is given by

$$-\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\lambda_{\mathrm{NLS}}))W - 2\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}\big(\mathrm{diag}(\mathrm{Im}\{(s_2\bar b_2)\nabla s_1+(s_1b_2+s_2(1+a_2))\nabla\bar s_2\}\cdot\xi)\big)W,$$

where by $\mathrm{Im}\{\vec b\}$, with $\vec b=(b_1,\dots,b_d)$, we denote the vector $(\mathrm{Im}(b_1),\dots,\mathrm{Im}(b_d))$. The second term in (6-14) is equal to $-\mathrm{i}EV*W$ modulo contributions to $R^{(1)}(U)$ thanks to (1-5) and (6-5).

Reasoning analogously one can prove that the term in (6-15) equals $-\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\vec a_1(U)\cdot\xi))W$, modulo contributions to $R^{(1)}(U)$. We are left with studying (6-16). First of all we note that $X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(U)=-\mathrm{i}E|u|^2U$; then we write

$$X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(U)=X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(W)+X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(U)-X_{\mathscr{H}^{(4)}_{\mathrm{NLS}}}(W).$$

Lemma 6.2 and Lemma 3.1(ii) (recall also (6-2)), imply $\|\Phi_{\rm NLS}(U)U-U\|_{H^s}\lesssim\|U\|^7_{H^s}$; therefore it is a contribution to $R^{(1)}(U)$. We have obtained $\Phi_{\rm NLS}(U)X_{\mathscr{H}^{(4)}_{\rm NLS}}(U)=X_{\mathscr{H}^{(4)}_{\rm NLS}}(W)$ modulo $R^{(1)}(U)$.

Summarizing we obtained (6-11) with symbols $a_2^{(1)}$ defined in (6-1) and

$$\vec a_1^{(1)}=\vec a_1+2\,\mathrm{Im}\{(s_2\bar b_2)\nabla s_1+(s_1b_2+s_2(1+a_2))\nabla\bar s_2\}\in\mathbb{R}, \tag{6-19}$$

with $\vec a_1$ in (4-10). □

6A2. *Diagonalization of cubic terms at order 0.* The aim of this section is to diagonalize the cubic vector field $X_{\mathscr{H}^{(4)}_{\rm NLS}}$ in (6-11) (see also (4-13)) up to smoothing remainder. In order to do this we will consider a change of coordinates which is *symplectic* up to high degree of homogeneity. We reason as follows.

Let

$$Z:=\begin{bmatrix} z\\ \bar z\end{bmatrix}:=\Phi_{\mathscr{B}_{\rm NLS}}(W):=W+X_{\mathscr{B}_{\rm NLS}}(W), \tag{6-20}$$

where $X_{\mathscr{B}_{\rm NLS}}$ is the Hamiltonian vector field of (5-3). We note that $\Phi_{\mathscr{B}_{\rm NLS}}$ is not symplectic; nevertheless it is close to the flow of $\mathscr{B}_{\rm NLS}(W)$, which is symplectic. The properties of $X_{\mathscr{B}_{\rm NLS}}$ and the estimates of $\Phi_{\mathscr{B}_{\rm NLS}}$ have been discussed in Lemma 5.1 and in Proposition 5.2.

Remark 6.5. Recall (6-10) and (6-20). One can note that, owing to Hypothesis 6.1, for $s>2d+2$, we have

$$\big(1-\tfrac{1}{100}\big)\|U\|_{H^s}\le\|W\|_{H^s}\le\big(1+\tfrac{1}{100}\big)\|U\|_{H^s},\quad \big(1-\tfrac{1}{100}\big)\|W\|_{H^s}\le\|Z\|_{H^s}\le\big(1+\tfrac{1}{100}\big)\|W\|_{H^s}. \tag{6-21}$$

This is a consequence of the estimates (6-5), (6-8), (5-12), (5-9), (5-14) tanking ϵ small enough depending on s.

We prove the following.

Proposition 6.6 (diagonalization at order 0). *Let $U=(u,\bar u)$ be a solution of* (4-12) *and assume Hypothesis 6.1. Define $W:=\Phi_{\rm NLS}(U)U$, where $\Phi_{\rm NLS}(U)$ is the map in* (6-4) *given in Lemma 6.3. Then the function*

$$Z=\begin{bmatrix} z\\ \bar z\end{bmatrix}$$

defined in (6-20) *satisfies* (*recall* (1-25))

$$\begin{aligned}\partial_t Z=-\mathrm{i}E\Lambda_{\rm NLS}Z-\mathrm{i}E\,\mathrm{Op}^{\rm BW}(\mathrm{diag}(a_2^{(1)}(x)|\xi|^2))Z\\ -\mathrm{i}\,\mathrm{Op}^{\rm BW}(\mathrm{diag}(\vec a_1^{(1)}(x)\cdot\xi))Z+X_{\mathrm{H}^{(4)}_{\rm NLS}}(Z)+R_5^{(2)}(U),\end{aligned} \tag{6-22}$$

where $a_2^{(1)}(x)$, $\vec a_1^{(1)}(x)$ are the real-valued symbols appearing in Proposition 6.4, the cubic vector field $X_{\mathrm{H}^{(4)}_{\rm NLS}}(Z)$ has the form (*see* (5-23))

$$X_{\mathrm{H}^{(4)}_{\rm NLS}}(Z):=-\mathrm{i}E\,\mathrm{Op}^{\rm BW}\begin{pmatrix}2|z|^2 & 0\\ 0 & 2|z|^2\end{pmatrix}Z+Q_{\mathrm{H}^{(4)}_{\rm NLS}}(Z), \tag{6-23}$$

the remainder $Q_{\mathrm{H}^{(4)}_{\rm NLS}}$ is given by Lemma 5.4 and satisfies (5-24)–(5-25). *The remainder $R_5^{(2)}(U)$ has the form $(R_5^{(2,+)},\overline{R_5^{(2,+)}})^T$. Moreover, for any $s>2d+4$,*

$$\|R_5^{(2)}(U)\|_{H^s}\lesssim\|U\|^5_{H^s}. \tag{6-24}$$

The vector field $X_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z)$ *in* (6-23) *is Hamiltonian*; *i.e.*, (*see* (3-69), (3-72)) $X_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z) := -\mathrm{i}J\nabla \mathtt{H}^{(4)}_{\mathrm{NLS}}(Z)$, *with*

$$\mathtt{H}^{(4)}_{\mathrm{NLS}}(Z) := \mathscr{H}^{(4)}_{\mathrm{NLS}}(Z) - \{\mathscr{B}_{\mathrm{NLS}}(Z), \mathscr{H}^{(2)}_{\mathrm{NLS}}(Z)\}, \quad \mathscr{H}^{(2)}_{\mathrm{NLS}}(Z) = \int_{\mathbb{T}^d} \Lambda_{\mathrm{NLS}} z \cdot \bar{z}\, dx, \tag{6-25}$$

where $\mathscr{H}^{(4)}_{\mathrm{NLS}}$ *is in* (4-23), *and* $\mathscr{B}_{\mathrm{NLS}}$ *is in* (5-3), (5-2).

Proof. Recall (5-15). We have that (6-11) reads

$$\partial_t W = X_{\mathscr{H}^{(\le 4)}_{\mathrm{NLS}}}(W) - \mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(A(U; x, \xi))W + R^{(1)}(U),$$

where we set

$$A(U; x, \xi) := E\,\mathrm{diag}(a_2^{(1)}(U; x)|\xi|^2) + \mathrm{diag}(\vec{a}_1^{(1)}(U; x)\cdot \xi). \tag{6-26}$$

Hence by (6-20) we get

$$\begin{aligned}\partial_t Z = (d_W \Phi_{\mathscr{B}_{\mathrm{NLS}}}(W))[-\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(A(U; x, \xi))W] + (d_W \Phi_{\mathscr{B}_{\mathrm{NLS}}})(W)[X_{\mathscr{H}^{(\le 4)}_{\mathrm{NLS}}}(W)]&\\ + (d_W \Phi_{\mathscr{B}_{\mathrm{NLS}}})(W)[R^{(1)}(U)].&\end{aligned} \tag{6-27}$$

We study each summand separately. First of all we have

$$\|d_W \Phi_{\mathscr{B}_{\mathrm{NLS}}}(W)[R^{(1)}(U)]\|_{H^s} \overset{(5\text{-}9),(6\text{-}13)}{\lesssim} \|u\|^7_{H^s}(1 + \|w\|^2_{H^s}) \overset{(6\text{-}21)}{\lesssim} \|u\|^7_{H^s}. \tag{6-28}$$

Let us now analyze the first summand in the right-hand side of (6-27). We write

$$\begin{aligned}(d_W \Phi_{\mathscr{B}_{\mathrm{NLS}}}(W))[\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(A(U; x, \xi))W] &= \mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(A(U; x, \xi))Z + P_1 + P_2,\\ P_1 &:= \mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(A(U; x, \xi))[W - Z],\\ P_2 &:= ((d_W \Phi_{\mathscr{B}_{\mathrm{NLS}}}(W)) - \mathbb{1})[\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(A(U; x, \xi))W].\end{aligned} \tag{6-29}$$

Fix $s_0 > d$, we have, for $s \ge 2s_0 + 4$,

$$\|P_2\|_{H^s} \overset{(5\text{-}9)}{\lesssim} \|w\|^2_{H^s}\|\mathrm{Op}^{\mathrm{BW}}(A(U; x, \xi))W\|_{H^{s-2}} \overset{(6\text{-}12),(3\text{-}10),(6\text{-}21)}{\lesssim} \|u\|^9_{H^s}. \tag{6-30}$$

By (6-20), (5-9) we get $\|W - Z\|_{H^s} \lesssim \|w\|^3_{H^{s-2}}$. Therefore, by (6-29), (6-26), (6-12), (3-10) and (6-21) we get

$$\|P_1\|_{H^s} \lesssim \|u\|^6_{H^{2s_0+1}}\|W - Z\|_{H^{s+2}} \lesssim \|u\|^6_{H^{2s_0+1}}\|w\|^3_{H^s} \lesssim \|u\|^9_{H^s}. \tag{6-31}$$

The estimates (6-28), (6-30), (6-31) imply that the terms P_1, P_2 and $d_W \Phi_{\mathscr{B}_{\mathrm{NLS}}}(W)[R^{(1)}(U)]$ can be absorbed in a remainder satisfying (6-24). Finally we consider the second summand in (6-27). By Lemma 5.3 we deduce

$$d_W \Phi_{\mathscr{B}_{\mathrm{NLS}}}(W)[X_{\mathscr{H}^{(\le 4)}_{\mathrm{NLS}}}(W)] = X_{\mathscr{H}^{(\le 4)}_{\mathrm{NLS}}}(Z) + [X_{\mathscr{B}_{\mathrm{NLS}}}(Z), X_{\mathscr{H}^{(2)}_{\mathrm{NLS}}}(Z)] + R_5(Z),$$

where R_5 is a remainder satisfying the quintic estimate (5-18). By Lemma 5.4 we also have

$$X_{\mathscr{H}^{(\le 4)}_{\mathrm{NLS}}}(Z) + [X_{\mathscr{B}_{\mathrm{NLS}}}(Z), X_{\mathscr{H}^{(2)}_{\mathrm{NLS}}}(Z)] = -\mathrm{i}E\Lambda_{\mathrm{NLS}}Z + X_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z),$$

with $X_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}$ as in (6-23). Moreover it is Hamiltonian with Hamiltonian as in (6-25) by (5-23) and (3-73). □

Remark 6.7. The Hamiltonian function in (6-25) may be rewritten, up to symmetrizations, as in (3-78) with coefficients $\mathtt{h}_4(\xi,\eta,\zeta)$ satisfying (3-79). The coefficients of its Hamiltonian vector field have the form (3-82) (see also (3-81)). Moreover, by (6-23), (3-6), (5-23), (5-24), we deduce that

$$-2\mathrm{i}\mathtt{h}_4(\xi,\eta,\zeta)=-2\mathrm{i}\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)+\mathtt{q}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(\xi,\eta,\zeta). \tag{6-32}$$

6B. *Diagonalization of the KG.* In this section we diagonalize the system (4-44) up to a smoothing remainder. This will be done into two steps. We first diagonalize the matrix $E(\mathbb{1}+\mathscr{A}_1(x,\xi))$ in (4-44) by means of a change of coordinates similar to the one made in the previous section for the (NLS) case. After that we diagonalize the matrix of symbols of order 0 at homogeneity 3, by means of an *approximately symplectic* change of coordinates. Consider the Cauchy problem associated to (KG). Throughout the rest of the section we shall assume the following.

Hypothesis 6.8. We restrict the solution of (KG) on the interval of times $[0,T)$, with T such that

$$\sup_{t\in[0,T)}(\|\psi(t,\cdot)\|_{H^{s+1/2}}+\|\partial_t\psi(t,\cdot)\|_{H^{s-1/2}})\le\epsilon,\quad \|\psi_0(\cdot)\|_{H^{s+1/2}}+\|\psi_1(\cdot)\|_{H^{s-1/2}}\le\tfrac{1}{32}\epsilon,$$

with $\psi(0,x)=\psi_0(x)$ and $(\partial_t\psi)(0,x)=\psi_1(x)$.

Note that such a T exists thanks to the local well-posedness proved in [Kato 1975].

Remark 6.9. Recall (3-77). Then one can note that

$$\tfrac{1}{4}(\|\psi(t,\cdot)\|_{H^{s+1/2}}+\|\partial_t\psi(t,\cdot)\|_{H^{s-1/2}})\le\|u\|_{H^s}\le 2(\|\psi(t,\cdot)\|_{H^{s+1/2}}+\|\partial_t\psi(t,\cdot)\|_{H^{s-1/2}}).$$

6B1. *Diagonalization at order 1.* Consider the matrix of symbols (see (4-24), (4-25))

$$E(\mathbb{1}+\mathscr{A}_1(x,\xi)),\quad \mathscr{A}_1(x,\xi):=\begin{bmatrix}1&1\\1&1\end{bmatrix}\tilde{a}_2(x,\xi),\quad \tilde{a}_2(x,\xi):=\tfrac{1}{2}\Lambda_{\mathrm{KG}}^{-2}(\xi)a_2(x,\xi). \tag{6-33}$$

Define

$$\lambda_{\mathrm{KG}}(x,\xi):=\sqrt{(1+\tilde{a}_2(x,\xi))^2-(\tilde{a}_2(x,\xi))^2},\quad \tilde{a}_2^+(x,\xi):=\lambda_{\mathrm{KG}}(x,\xi)-1. \tag{6-34}$$

Notice that the symbol $\lambda_{\mathrm{KG}}(x,\xi)$ is well-defined by taking $\|u\|_{H^s}\ll 1$ small enough. The matrix of eigenvectors associated to the eigenvalues of $E(\mathbb{1}+\mathscr{A}_1(x,\xi))$ is

$$\begin{gathered}S(x,\xi):=\begin{pmatrix}s_1(x,\xi)&s_2(x,\xi)\\s_2(x,\xi)&s_1(x,\xi)\end{pmatrix},\qquad S^{-1}(x,\xi):=\begin{pmatrix}s_1(x,\xi)&-s_2(x,\xi)\\-s_2(x,\xi)&s_1(x,\xi)\end{pmatrix},\\ s_1:=\frac{1+\tilde{a}_2+\lambda_{\mathrm{KG}}}{\sqrt{2\lambda_{\mathrm{KG}}(1+\tilde{a}_2+\lambda_{\mathrm{KG}})}},\qquad s_2:=\frac{-\tilde{a}_2}{\sqrt{2\lambda_{\mathrm{KG}}(1+\tilde{a}_2+\lambda_{\mathrm{KG}})}}.\end{gathered} \tag{6-35}$$

By a direct computation one can check that

$$S^{-1}(x,\xi)E(\mathbb{1}+\mathscr{A}_1(x,\xi))S(x,\xi)=E\,\mathrm{diag}(\lambda_{\mathrm{KG}}(x,\xi)),\quad s_1^2-|s_2|^2=1. \tag{6-36}$$

We shall study how the system (4-44) transforms under the maps

$$\Phi_{\mathrm{KG}}=\Phi_{\mathrm{KG}}(U)[\cdot]:=\mathrm{Op}^{\mathrm{BW}}(S^{-1}(x,\xi)),\quad \Psi_{\mathrm{KG}}=\Psi_{\mathrm{KG}}(U)[\cdot]:=\mathrm{Op}^{\mathrm{BW}}(S(x,\xi)). \tag{6-37}$$

Lemma 6.10. *Assume Hypothesis 6.8. We have the following*:

(i) *If $s_0 > d$, then*

$$|\tilde{a}_2^+|_{\mathcal{N}_p^0} + |\tilde{a}_2|_{\mathcal{N}_p^0} + |s_1 - 1|_{\mathcal{N}_p^0} + |s_2|_{\mathcal{N}_p^0} \lesssim \|u\|^3_{H^{p+s_0+1}}, \quad p + s_0 + 1 \le s. \tag{6-38}$$

(ii) *For any $s \in \mathbb{R}$ one has*

$$\|\Phi_{\mathrm{KG}}(U)V - V\|_{H^s} + \|\Psi_{\mathrm{KG}}(U)V - V\|_{H^s} \lesssim \|V\|_{H^s}\|u\|^3_{H^{2s_0+1}} \quad \text{for all } V \in H^s(\mathbb{T}^d; \mathbb{C}^2). \tag{6-39}$$

(iii) *One has $\Psi_{\mathrm{KG}}(U) \circ \Phi_{\mathrm{KG}}(U) = \mathbb{1} + Q(U)$, where Q is a real-to-real remainder satisfying*

$$\|Q(U)V\|_{H^{s+1}} \lesssim \|V\|_{H^s}\|u\|^3_{H^{2s_0+3}}. \tag{6-40}$$

(iv) *For any $t \in [0, T)$, one has $\partial_t \Phi_{\mathrm{KG}}(U)[\,\cdot\,] = \mathrm{Op}^{\mathrm{BW}}(\partial_t S^{-1}(x, \xi))$ and*

$$|\partial_t S^{-1}(x, \xi)|_{\mathcal{N}_{s_0}^0} \lesssim \|u\|^3_{H^{2s_0+3}}, \quad \|\partial_t \Phi_{\mathrm{KG}}(U)V\|_{H^s} \lesssim \|V\|_{H^s}\|u\|^3_{H^{2s_0+3}}. \tag{6-41}$$

Proof. (i) Inequality (6-38) follows by (4-45) using the explicit formulas (6-35), (6-34).

(ii) This follows by using (6-38) and Lemma 3.1(ii).

(iii) By formula (3-19) in Proposition 3.2 one gets

$$\Psi_{\mathrm{KG}}(U) \circ \Phi_{\mathrm{KG}}(U) = \mathbb{1} + \mathrm{Op}^{\mathrm{BW}}\begin{pmatrix} 0 & \mathrm{i}\{s_1, s_2\} \\ -\mathrm{i}\{s_1, s_2\} & 0 \end{pmatrix} + R(s_1, s_2)$$

for some remainder satisfying (3-20) with $a \rightsquigarrow s_1$ and $b \rightsquigarrow s_2$. Therefore (6-40) follows by using (3-8), (3-10) and (6-38).

(iv) This is similar to the proof of Lemma 6.3(iii). □

Proposition 6.11 (diagonalization at order 1). *Consider the system* (4-44) *and set*

$$W = \Phi_{\mathrm{KG}}(U)U, \tag{6-42}$$

with Φ_{KG} defined in (6-37). *Then W solves the equation* (*recall* (4-9))

$$\partial_t W = -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}\big(\mathrm{diag}(1 + \tilde{a}_2^+(x, \xi))\Lambda_{\mathrm{KG}}(\xi)\big)W + X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(W) + R^{(1)}(u), \tag{6-43}$$

where the vector field $X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}$ is defined in (4-28). *The symbol $\tilde{a}_2^+$ is defined in* (6-34). *The remainder $R^{(1)}$ has the form $(R^{(1,+)}, \overline{R^{(1,+)}})^T$. Moreover, for any $s > 2d + \mu$, for some $\mu > 0$, it satisfies the estimate*

$$\|R^{(1)}(u)\|_{H^s} \lesssim \|u\|^4_{H^s}. \tag{6-44}$$

Proof. By (6-42) and (4-44) we get

$$\begin{aligned}
\partial_t W &= \Phi_{\mathrm{KG}}(U)\dot{U} + (\partial_t \Phi_{\mathrm{KG}}(U))[U] \\
&= -\mathrm{i}\Phi_{\mathrm{KG}}(U)\,\mathrm{Op}^{\mathrm{BW}}\big(E(\mathbb{1} + \mathscr{A}_1(x, \xi))\Lambda_{\mathrm{KG}}(\xi)\big)\Psi_{\mathrm{KG}}(U)W + \Phi_{\mathrm{KG}}(U)X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(U) \\
&\quad + \Phi_{\mathrm{KG}}(U)R(u) + (\partial_t \Phi_{\mathrm{KG}}(U))[U] \\
&\quad\quad + \mathrm{i}\Phi_{\mathrm{KG}}(U)\,\mathrm{Op}^{\mathrm{BW}}\big(E(\mathbb{1} + \mathscr{A}_1(x, \xi))(\xi)\big)Q(U)U,
\end{aligned} \tag{6-45}$$

where we used items (ii), (iii) in Lemma 6.10. We study the first summand in the right-hand side of (6-45). By direct inspection, using Lemma 3.1 and Proposition 3.2 we get

$$\begin{aligned}-\mathrm{i}\Phi_{\mathrm{KG}}(U)\,\mathrm{Op}^{\mathrm{BW}}\big(E(\mathbb{1}+\mathscr{A}_1(x,\xi))\Lambda_{\mathrm{KG}}(\xi)\big)\Psi_{\mathrm{KG}}(U) &= -\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(S^{-1}E(\mathbb{1}+\mathscr{A}_1(x,\xi))S)+R(u)\\ &= -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\lambda_{\mathrm{KG}}(x,\xi)))+R(u) \quad \text{(by (6-36))},\end{aligned}$$

where $R(u)$ is a remainder satisfying (6-44). Thanks to the discussion above and (6-34) we obtain the highest-order term in (6-43). All the other summands in the right-hand side of (6-45) may be analyzed as done in the proof of Proposition 6.4 by using Lemma 6.10. □

6B2. *Diagonalization of cubic terms at order 0.* Above we showed that if the function U solves (4-44) then W in (6-42) solves (6-43). The cubic terms in the system (6-43) are the same as those in (4-44) and have the form (4-28). The aim of this section is to diagonalize the matrix of symbols of order zero $\mathscr{A}_0(x,\xi)$.

Let us define

$$Z := \begin{bmatrix} z \\ \bar{z} \end{bmatrix} := \Phi_{\mathscr{B}_{\mathrm{KG}}}(W) := W + X_{\mathscr{B}_{\mathrm{KG}}}(W), \tag{6-46}$$

where $X_{\mathscr{B}_{\mathrm{KG}}}$ is the Hamiltonian vector field of (5-5) and W is the function in (6-42). The properties of $X_{\mathscr{B}_{\mathrm{KG}}}$ and the estimates of $\Phi_{\mathscr{B}_{\mathrm{KG}}}$ have been discussed in Lemma 5.1 and in Proposition 5.2.

Remark 6.12. Recall (6-42) and (6-46). One can note that, owing to Hypothesis 6.8, for $s > 2d+3$, we have

$$\big(1-\tfrac{1}{100}\big)\|U\|_{H^s} \le \|W\|_{H^s} \le \big(1+\tfrac{1}{100}\big)\|U\|_{H^s}, \quad \big(1-\tfrac{1}{100}\big)\|W\|_{H^s} \le \|Z\|_{H^s} \le \big(1+\tfrac{1}{100}\big)\|W\|_{H^s}. \tag{6-47}$$

This is a consequence of the estimates (6-39), (6-40) (5-12), (5-10), (5-14) taking ϵ small enough.

Proposition 6.13 (diagonalization at order 0). *Let U be a solution of* (4-44) *and assume Hypothesis 6.8* (*see also Remark 6.9*). *Then the function Z defined in* (6-46), *with W given in* (6-42), *satisfies*

$$\partial_t Z = -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}\big(\mathrm{diag}(1+\tilde{a}_2^+(x,\xi))\Lambda_{\mathrm{KG}}(\xi)\big)Z + X_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z) + R_4^{(2)}(u), \tag{6-48}$$

where $\tilde{a}_2^+(x,\xi)$ is the real-valued symbol in (6-34), *the cubic vector field $X_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z)$ has the form*

$$X_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z) := -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(a_0(x,\xi)))Z + Q_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z), \tag{6-49}$$

the symbol $a_0(x,\xi)$ is as in (4-24), *and the remainder $Q_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z)$ is the cubic remainder given in Lemma 5.5. The remainder $R_4^{(2)}(u)$ has the form $(R_4^{(2,+)}(u), \overline{R_4^{(2,+)}(u)})^T$. Moreover, for any $s > 2d+\mu$, for some $\mu > 0$, we have the estimate*

$$\|R_4^{(2)}(u)\|_{H^s} \lesssim \|u\|_{H^s}^4. \tag{6-50}$$

Finally the vector field $X_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z)$ in (6-49) *is Hamiltonian*; *i.e.*, $X_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z) := -\mathrm{i}J\nabla \mathrm{H}^{(4)}_{\mathrm{KG}}(Z)$ *with*

$$\mathrm{H}^{(4)}_{\mathrm{KG}}(Z) := \mathscr{H}^{(4)}_{\mathrm{KG}}(Z) - \{\mathscr{B}_{\mathrm{KG}}(Z), \mathscr{H}^{(2)}_{\mathrm{KG}}(Z)\}, \quad \mathscr{H}^{(2)}_{\mathrm{KG}}(Z) = \int_{\mathbb{T}^d} \Lambda_{\mathrm{KG}} z \cdot \bar{z}\, dx, \tag{6-51}$$

where $\mathscr{H}^{(4)}_{\mathrm{KG}}$ is in (4-27), *and $\mathscr{B}_{\mathrm{KG}}$ is in* (5-5), (5-4).

Proof. We recall (5-16) and we rewrite (6-43) as

$$\partial_t W = X_{\mathscr{H}^{(\le 4)}_{\mathrm{KG}}}(W) - \mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\tilde{a}_2^{+}(x,\xi)\Lambda_{\mathrm{KG}}(\xi))W + R^{(1)}(u).$$

Then, using (6-46), we get

$$\begin{aligned} \partial_t Z &= d_W\Phi_{\mathscr{B}_{\mathrm{KG}}}(W)[\partial_t W] \\ &= d_W\Phi_{\mathscr{B}_{\mathrm{KG}}}(W)[X_{\mathscr{H}^{(\le 4)}_{\mathrm{KG}}}(W)] && \text{(6-52)} \\ &\quad + d_W\Phi_{\mathscr{B}_{\mathrm{KG}}}(W)[-\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\tilde{a}_2^{+}(x,\xi)\Lambda_{\mathrm{KG}}(\xi)))W] && \text{(6-53)} \\ &\quad\quad + d_W\Phi_{\mathscr{B}_{\mathrm{KG}}}(W)[R^{(1)}(u)]. && \text{(6-54)} \end{aligned}$$

By estimates (5-10) and (6-44) we have that the term in (6-54) can be absorbed in a remainder satisfying (6-50). Consider the term in (6-53). We write

$$\begin{aligned} \text{(6-53)} &= -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\tilde{a}_2^{+}(x,\xi)\Lambda_{\mathrm{KG}}(\xi)))Z + P_1 + P_2, \\ P_1 &:= -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\tilde{a}_2^{+}(x,\xi)\Lambda_{\mathrm{KG}}(\xi)))[W-Z], && \text{(6-55)} \\ P_2 &:= ((d_W\Phi_{\mathscr{B}_{\mathrm{KG}}}(W)) - \mathbb{1})[-\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\tilde{a}_2^{+}(x,\xi)\Lambda_{\mathrm{KG}}(\xi)))W]. \end{aligned}$$

We have, for $s \ge 2s_0 + 2$,

$$\|P_2\|_{H^s} \overset{\text{(5-10)}}{\lesssim} \|u\|^2_{H^s}\|\mathrm{Op}^{\mathrm{BW}}(\tilde{a}_2^{+}(x,\xi)\Lambda_{\mathrm{KG}}(\xi))w\|_{H^{s-1}} \overset{\text{(6-38),(3-10),(6-47)}}{\lesssim} \|u\|^6_{H^s},$$

which implies (6-50). By (5-14) in Proposition 5.2 and estimate (5-10) we deduce $\|W-Z\|_{H^{s+1}} \lesssim \|u\|^3_{H^s}$. Hence using again (6-38), (3-10), (6-47) we get P_1 satisfies (6-50). It remains to discuss the structure of the term in (6-52). By Lemma 5.3 we obtain

$$d_W\Phi_{\mathscr{B}_{\mathrm{KG}}}(W)[X_{\mathscr{H}^{(\le 4)}_{\mathrm{KG}}}(W)] = X_{\mathscr{H}^{(\le 4)}_{\mathrm{KG}}}(Z) + [X_{\mathscr{B}_{\mathrm{KG}}}(Z), X_{\mathscr{H}^{(2)}_{\mathrm{KG}}}(Z)], \tag{6-56}$$

modulo remainders that can be absorbed in $R_4^{(2)}$ satisfying (6-50). Then (6-56), (6-52)–(6-54) and the discussion above imply (6-48), where the cubic vector field has the form

$$X_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(Z) = X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}(Z) + [X_{\mathscr{B}_{\mathrm{KG}}}(Z), X_{\mathscr{H}^{(2)}_{\mathrm{KG}}}(Z)]. \tag{6-57}$$

Using (3-73), (3-72), we conclude that $X_{\mathtt{H}^{(4)}_{\mathrm{KG}}}$ is the Hamiltonian vector field of $\mathtt{H}^{(4)}_{\mathrm{KG}}$ in (6-51). Equation (6-49) follows by Lemma 5.5. □

Remark 6.14. In view of Remarks 4.6 and 4.8, following the same proof as Proposition 6.13, in the semilinear case we obtain that (6-48) reads

$$\partial_t Z = -\mathrm{i}E\,\mathrm{Op}^{\mathrm{BW}}(\mathrm{diag}(\Lambda_{\mathrm{KG}}(\xi)))Z + X_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(Z) + R_4^{(2)}(u),$$

where $X_{\mathtt{H}^{(4)}_{\mathrm{KG}}}$ has the form (6-49) with $a_0(x,\xi)$ a symbol of order -1 and $Q_{\mathtt{H}^{(4)}_{\mathrm{KG}}}$ a remainder of the form (5-30) with coefficients satisfying (5-31) with the better denominator $\max\{\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\}^2$.

7. Energy estimates

7A. *Estimates for the NLS.* In this section we prove a priori energy estimates on the Sobolev norms of the variable Z in (6-20). In Section 7A1 we introduce a convenient energy norm on $H^s(\mathbb{T}^d;\mathbb{C})$ which is equivalent to the classic H^s-norm. This is the content of Lemma 7.2. In Section 7A2, using the nonresonance conditions of Proposition 2.1, we provide bounds on the nonresonant terms appearing in the energy estimates. We deal with resonant interactions in Lemma 7.4.

7A1. *Energy norm.* Let us define the symbol

$$\mathscr{L} = \mathscr{L}(x,\xi) := |\xi|^2 + \Sigma, \quad \Sigma = \Sigma(x,\xi) := a_2^{(1)}(x)|\xi|^2 + \vec{a}_1^{(1)}(x)\cdot\xi, \tag{7-1}$$

where the symbols $a_2^{(1)}(x)$, $\vec{a}_1^{(1)}(x)$ are given in Proposition 6.4.

Lemma 7.1. *Assume Hypothesis 6.1 and let $\gamma > 0$. Then for $\epsilon > 0$ small enough we have the following*:

(i) *One has*

$$|\Sigma|_{\mathcal{N}_{s_0}^2} \le C\|u\|^6_{H^{2s_0+1}}, \quad |(1+\mathscr{L})^\gamma - (|\xi|^2+1)^\gamma|_{\mathcal{N}_{s_0}^{2\gamma}} \lesssim_\gamma C\|u\|^6_{H^{2s_0+1}} \tag{7-2}$$

for some $C>0$ depending on s_0.

(ii) *For any $s\in\mathbb{R}$ and any $h\in H^s(\mathbb{T}^d;\mathbb{C})$, one has*

$$\begin{gathered}\|T_{\mathscr{L}^\gamma}h\|_{H^{s-2\gamma}} \le \|h\|_{H^s}(1+C\|u\|^6_{H^{2s_0+1}}),\\ \|T_\Sigma h\|_{H^{s-2}} + \|T_{(1+\mathscr{L})^\gamma-(|\xi|^2+1)^\gamma}h\|_{H^{s-2\gamma}} \lesssim_\gamma \|h\|_{H^s}\|u\|^6_{H^{2s_0+1}}\end{gathered} \tag{7-3}$$

for some $C>0$ depending on s and γ.

(iii) *For any $t\in[0,T)$ one has $|\partial_t\Sigma|_{\mathcal{N}_{s_0}^2} \lesssim \|u\|^6_{H^{2s_0+3}}$. Moreover*

$$\|(T_{\partial_t(1+\mathscr{L})^\gamma})h\|_{H^{s-2\gamma}} \lesssim_\gamma \|h\|_{H^s}\|u\|^6_{H^{2s_0+3}} \quad \textit{for all } h\in H^s(\mathbb{T}^d;\mathbb{C}). \tag{7-4}$$

(iv) *The operators $T_{\mathscr{L}}$, $T_{(1+\mathscr{L})^\gamma}$ are self-adjoint with respect to the L^2-scalar product* (3-3).

Proof. (i)–(ii) Inequalities (7-2) follow by using (7-1), the bounds (6-12) on the symbols $a_2^{(1)}$ and $\vec{a}_1^{(1)}\cdot\xi$; (7-3) follows by Lemma 3.1.

(iii) The bound on $\partial_t\Sigma$ follows by reasoning as in Lemma 6.3(iii) using the explicit formula of $a_2^{(1)}$ in (6-1) and the formula for $a_1^{(1)}\cdot\xi$ in (6-19) (see also (6-2)). Then (3-10) implies (7-4).

(iv) This follows by (3-54) since the symbol $\mathscr{L}$ in (7-1) is real-valued. □

In the following we shall construct the *energy norm.* By using this norm we are able to achieve the *energy estimates* on the previously diagonalized system. For $s\in\mathbb{R}$ we define

$$z_n := T_{(1+\mathscr{L})^n}z, \quad Z_n = \begin{bmatrix} z_n \\ \bar{z}_n\end{bmatrix} := T_{(1+\mathscr{L})^n}\mathbb{1}Z, \quad Z = \begin{bmatrix} z \\ \bar{z}\end{bmatrix}, \quad n := \tfrac{1}{2}s. \tag{7-5}$$

Lemma 7.2 (equivalence of the energy norm). *Assume Hypothesis 6.1 with $s > 2d+4$. Then, for $\epsilon > 0$ small enough enough, one has*

$$\big(1-\tfrac{1}{100}\big)\|z\|_{H^s} \le \|z_n\|_{L^2} \le \big(1+\tfrac{1}{100}\big)\|z\|_{H^s}. \tag{7-6}$$

Proof. Let $s = 2n$. Then by (7-3) and (7-5) we have $\|z_n\|_{L^2} \le \|z\|_{H^s}(1 + C\|u\|^6_{H^{2s_0+1}}) \le 2\|z\|_{H^2}$, with $s_0 > d$. Moreover

$$\|z\|_{H^s} = \|T_{(1+|\xi|^2)^n} z\|_{L^2} \overset{(7\text{-}3)}{\le} \|z_n\|_{L^2} + C\|z\|_{H^s}\|u\|^6_{H^{2s_0+1}},$$

which implies $(1 - C\|u\|^6_{H^{2s_0+1}})\|z\|_{H^s} \le \|z_n\|_{L^2}$ for some constant C depending on s. The discussion above implies (7-6) by taking $\epsilon > 0$ in Hypothesis 6.1 small enough. □

Recalling (6-22), (1-25) and (7-1) we have

$$(\partial_t + \mathrm{i}\Lambda_{\mathrm{NLS}})z = -\mathrm{i}T_\Sigma z + X^+_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z) + R_5^{(2,+)}(U), \qquad Z = \begin{bmatrix} z \\ \bar{z} \end{bmatrix}, \tag{7-7}$$

where $X_{\mathtt{H}_4^{(4)}}$ is given in (6-23) (see also Remark 6.7) and $R_5^{(2,+)}$ is the remainder satisfying (6-24).

Lemma 7.3. *Fix $s > 2d + 4$ and recall* (7-7). *One has that the function z_n defined in* (7-5) *solves the problem*

$$\partial_t z_n = -\mathrm{i}T_{\mathscr{L}} z_n - \mathrm{i}V * z_n + T_{(1+|\xi|^2)^n} X^{+,\mathrm{res}}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z) + B_n^{(1)}(Z) + B_n^{(2)}(Z) + R_{5,n}(U), \tag{7-8}$$

where $X^{+,\mathrm{res}}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}$ is defined as in Definition 3.9,

$$\begin{aligned} \widehat{B_n^{(1)}(Z)}(\xi) &= \frac{1}{(2\pi)^d} \sum_{\eta,\zeta\in\mathbb{Z}^d} b^{(1)}(\xi,\eta,\zeta)\hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}_n(\zeta), \\ \widehat{B_n^{(2)}(Z)}(\xi) &= \frac{1}{(2\pi)^d} \sum_{\eta,\zeta\in\mathbb{Z}^d} b_n^{(2)}(\xi,\eta,\zeta)\hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}(\zeta), \end{aligned} \tag{7-9}$$

with

$$b^{(1)}(\xi,\eta,\zeta) := -2\mathrm{i}\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle \xi+\zeta\rangle}\right)\mathbb{1}_{\mathscr{R}^c}(\xi,\eta,\zeta), \tag{7-10}$$

$$|b_n^{(2)}(\xi,\eta,\zeta)| \lesssim \frac{\langle\xi\rangle^{2n}\max_2\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}^4}{\max_1\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}}\mathbb{1}_{\mathscr{R}^c}(\xi,\eta,\zeta), \tag{7-11}$$

and where the remainder $R_{5,n}$ satisfies

$$\|R_{5,n}(U)\|_{L^2} \lesssim \|u\|^5_{H^s}. \tag{7-12}$$

Proof. Recalling (3-84) we define

$$X^{+,\perp}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z) := X^+_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z) - X^{+,\mathrm{res}}_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z). \tag{7-13}$$

By differentiating (7-5) and using (7-1) and (7-7) we get

$$\begin{aligned} \partial_t z_n &= T_{(1+\mathscr{L})^n}\partial_t z + T_{\partial_t(1+\mathscr{L})^n} z \\ &= -\mathrm{i}T_{\mathscr{L}} z_n - \mathrm{i}T_{(1+\mathscr{L})^n}(V*z) + T_{(1+\mathscr{L})^n} X^+_{\mathtt{H}^{(4)}_{\mathrm{NLS}}}(Z) + T_{(1+\mathscr{L})^n} R_5^{(2,+)}(U) \\ &\qquad\qquad + T_{\partial_t(1+\mathscr{L})^n} z - \mathrm{i}[T_{(1+\mathscr{L})^n}, T_{\mathscr{L}}]z. \end{aligned} \tag{7-14}$$

By using Lemmas 3.1 and 7.1, Proposition 3.2, and (7-6), (6-21) one proves that the last summand gives a contribution to $R_{5,n}(U)$ satisfying (7-12). By using (7-4), (6-21), (6-24) we deduce that

$$\|T_{(1+\mathscr{L})^n} R_5^{(2,+)}(U)\|_{L^2} + \|T_{\partial_t(1+\mathscr{L})^n} z\|_{L^2} \lesssim \|u\|^5_{H^s}.$$

Secondly we write

$$\mathrm{i}T_{(1+\mathscr{L})^n}(V*z)=\mathrm{i}V*z_n+\mathrm{i}V*(T_{(1+|\xi|^2)^n-(1+\mathscr{L})^n}z)+\mathrm{i}T_{(1+\mathscr{L})^n-(1+|\xi|^2)^n}(V*z).$$

By (7-3), (6-21), and recalling (1-5) we conclude $\|T_{(1+\mathscr{L})^n}(V*z)-V*z_n\|_{L^2}\lesssim\|u\|_{H^s}^7$. We now study the third summand in (7-14). We have (see (7-13))

$$T_{(1+\mathscr{L})^n}X^{+}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z)=T_{(1+|\xi|^2)^n}X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z)+T_{(1+|\xi|^2)n}X^{+,\perp}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z)+T_{(1+\mathscr{L})^n-(1+|\xi|^2)^n}X^{+}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z).$$

By (7-3), (6-23), (3-10), Lemma 3.7 and using the estimate (5-25), one obtains

$$\|T_{(1+\mathscr{L})^n-(1+|\xi|^2)^n}X^{+}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z)\|_{L^2}\lesssim\|u\|_{H^s}^9.$$

Recalling (6-32) and (7-13) we write

$$\begin{aligned}
&T_{(1+|\xi|^2)^n}X^{+,\perp}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z)=\mathscr{C}_1+\mathscr{C}_2+\mathscr{C}_3,\quad \widehat{\mathscr{C}}_i(\xi)=\frac{1}{(2\pi)^d}\sum_{\eta,\zeta\in\mathbb{Z}^d}\mathtt{c}_i(\xi,\eta,\zeta)\hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}(\zeta),\\
&\qquad\mathtt{c}_1(\xi,\eta,\zeta):=-2\mathrm{i}\chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big)(1+|\zeta|^2)^n\mathbb{1}_{\mathscr{R}^c}(\xi,\eta,\zeta),\\
&\qquad\mathtt{c}_2(\xi,\eta,\zeta):=-2\mathrm{i}\chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big)[(1+|\xi|^2)^n-(1+|\zeta|^2)^n]\mathbb{1}_{\mathscr{R}^c}(\xi,\eta,\zeta),\\
&\qquad\mathtt{c}_3(\xi,\eta,\zeta):=\mathtt{q}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(\xi,\eta,\zeta)(1+|\xi|^2)^n\mathbb{1}_{\mathscr{R}^c}(\xi,\eta,\zeta).
\end{aligned}\tag{7-15}$$

We now consider the operator $\mathscr{C}_1$ with coefficients $\mathtt{c}_1(\xi,\eta,\zeta)$. First of all we remark that it can be written as $\mathscr{C}_1=M(z,\bar{z},z)$, where M is a trilinear operator of the form (3-62). Moreover, setting

$$z_n=T_{(1+|\xi|^2)^n}z+h_n,\quad h_n:=T_{(1+\mathscr{L})^n-(1+|\xi|^2)^n}z,$$

we can write $\mathscr{C}_1=B_n^{(1)}(Z)-M(z,\bar{z},h_n)$, where $B_n^{(1)}$ has the form (7-9) with coefficients as in (7-10). Using that $|\mathtt{c}_1(\xi,\eta,\zeta)|\lesssim 1$, Lemma 3.7 (with $m=0$) and (7-3) we deduce that $\|M(z,\bar{z},h_n)\|_{L^2}\lesssim\|u\|_{H^s}^9$. Therefore this is a contribution to $R_{5,n}(U)$ satisfying (7-12). The discussion above implies formula (7-8) by setting $B_n^{(2)}$ as the operator of the form (7-9) with coefficients $b_n^{(2)}(\xi,\eta,\zeta):=\mathtt{c}_2(\xi,\eta,\zeta)+\mathtt{c}_3(\xi,\eta,\zeta)$. The coefficient $\mathtt{c}_3(\xi,\eta,\zeta)$ satisfies (7-11) by (5-25). For the coefficient $\mathtt{c}_2(\xi,\eta,\zeta)$ one has to apply Lemma 3.8 with $\mu=m=1$ and $f(\xi,\eta,\zeta):=((1+|\xi|^2)^n-(1+|\zeta|^2)^n)\langle\xi\rangle^{-2n}$. □

In the following lemma we prove a key cancellation due to the fact that the *super actions* are prime integrals of the resonant Hamiltonian vector field $X^{+,\mathrm{res}}_{\mathrm{H}_4}(Z)$ in the spirit of [Faou et al. 2013]. We also prove an important algebraic property of the operator $B_n^{(1)}$ in (7-8).

Lemma 7.4. *For any $n\geq 0$ we have*

$$\mathrm{Re}(T_{\langle\xi\rangle^n}X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z),T_{\langle\xi\rangle^n}z)_{L^2}=0,\tag{7-16}$$

$$\mathrm{Re}(B_n^{(1)}(Z),z_n)_{L^2}=0,\tag{7-17}$$

where $X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}$ is defined in Lemma 7.3 and $B_n^{(1)}$ in (7-9), (7-10).

Proof. Equation (7-16) follows by Lemma 3.10. Let us check (7-17). By an explicit computation using (3-3), (7-9) we get

$$\begin{aligned}\mathrm{Re}(B_n^{(1)}(Z), z_n)_{L^2} &= \frac{1}{(2\pi)^d}\sum_{\xi,\eta,\zeta\in\mathbb{Z}^d} b^{(1)}(\xi,\eta,\zeta)\hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}_n(\zeta)\hat{\bar{z}}_n(-\xi)\\ &\qquad + \frac{1}{(2\pi)^d}\sum_{\xi,\eta,\zeta\in\mathbb{Z}^d}\overline{b^{(1)}(\xi,\eta,\zeta)}\hat{\bar{z}}(-\xi+\eta+\zeta)\hat{z}(-\eta)\hat{\bar{z}}_n(-\zeta)\hat{z}_n(\xi)\\ &= \frac{1}{(2\pi)^d}\sum_{\xi,\eta,\zeta\in\mathbb{Z}^d}[b^{(1)}(\xi,\eta,\zeta)+\overline{b^{(1)}(\zeta,\zeta+\eta-\xi,\xi)}]\hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}_n(\zeta)\hat{\bar{z}}_n(-\xi).\end{aligned}$$

By (7-10) we have

$$b^{(1)}(\xi,\eta,\zeta)+\overline{b^{(1)}(\zeta,\zeta+\eta-\xi,\xi)} = 2\mathrm{i}\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)[1_{\mathscr{R}^c}(\xi,\eta,\zeta)-1_{\mathscr{R}^c}(\zeta,\zeta+\eta-\xi,\xi)] = 0,$$

where we used the form of the resonant set $\mathscr{R}$ in (3-83). □

We conclude the section with the following proposition.

Proposition 7.5. *Let $u(t,x)$ be a solution of* (NLS) *satisfying Hypothesis 6.1 and consider the function z_n in* (7-5) (*see also* (6-20), (6-10)). *Then, setting $s = 2n > 2d+4$ we have*

$$\frac{1}{2^{1/4}}\|u(t)\|_{H^s} \le \|z_n(t)\|_{L^2} \le 2^{1/4}\|u(t)\|_{H^s} \tag{7-18}$$

and

$$\partial_t\|z_n(t)\|^2_{L^2} = \mathscr{B}(t)+\mathscr{B}_{>5}(t), \quad t\in[0,T), \tag{7-19}$$

where:

- *The term $\mathscr{B}(t)$ has the form*

$$\begin{aligned}\mathscr{B}(t) &= \frac{2}{(2\pi)^d}\sum_{\xi,\eta,\zeta\in\mathbb{Z}^d}\langle\xi\rangle^{2n}\mathtt{b}(\xi,\eta,\zeta)\hat{z}(\xi-\eta-\zeta)\hat{\bar{z}}(\eta)\hat{z}(\zeta)\hat{\bar{z}}(-\xi),\\ \mathtt{b}(\xi,\eta,\zeta) &= b_n^{(2)}(\xi,\eta,\zeta)+\overline{b_n^{(2)}(\zeta,\zeta+\eta-\xi,\xi)}, \quad \xi,\eta,\zeta\in\mathbb{Z}^d,\end{aligned} \tag{7-20}$$

where $b_n^{(2)}(\xi,\eta,\zeta)$ are the coefficients in (7-9), (7-11).

- *The term $\mathscr{B}_{>5}(t)$ satisfies*

$$|\mathscr{B}_{>5}(t)| \lesssim \|u\|^6_{H^s}, \quad t\in[0,T). \tag{7-21}$$

Proof. The norm $\|z_n\|_{L^2}$ is equivalent to $\|u\|_{H^s}$ by using Lemma 7.2 and Remark 6.5. Using (7-8) we get

$$\tfrac{1}{2}\partial_t\|z_n(t)\|^2_{L^2} = \mathrm{Re}(T_{\langle\xi\rangle^{2n}}X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z), z_n)_{L^2} \tag{7-22}$$

$$+\mathrm{Re}(-\mathrm{i}T_{\mathscr{L}}z_n, z_n)_{L^2}+\mathrm{Re}(B_n^{(1)}(Z), z_n)_{L^2}+\mathrm{Re}(-\mathrm{i}V*z_n, z_n)_{L^2} \tag{7-23}$$

$$+\mathrm{Re}(B_n^{(2)}(Z), z_n)_{L^2} \tag{7-24}$$

$$+\mathrm{Re}(R_{5,n}(Z), z_n)_{L^2}. \tag{7-25}$$

Recall that $T_{\mathscr{L}}$ is self-adjoint (see Lemma 7.1(iv)) and the convolution potential V has real Fourier coefficients. Then by using also Lemma 7.4 (see (7-17)) we deduce (7-23) $= 0$. Moreover by the

Cauchy–Schwarz inequality and estimates (7-12), (7-6) and (6-21) we obtain that the term in (7-25) is bounded from above by $\|u\|_{H^s}^6$. Consider the terms in (7-22) and (7-24). Recalling (7-5) and (7-1) we write

$$\begin{aligned}\operatorname{Re}(T_{\langle\xi\rangle^{2n}} X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z), z_n)_{L^2} &= \operatorname{Re}(T_{\langle\xi\rangle^{2n}} X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z), T_{\langle\xi\rangle^{2n}} z)_{L^2} + \operatorname{Re}(T_{\langle\xi\rangle^{2n}} X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z), T_{(1+\mathscr{L})^n-\langle\xi\rangle^{2n}} z)_{L^2}\\ &= \operatorname{Re}(T_{\langle\xi\rangle^{2n}} X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z), T_{(1+\mathscr{L})^n-\langle\xi\rangle^{2n}} z)_{L^2} \qquad \text{(by (7-16))}.\end{aligned}$$

Moreover we write

$$\operatorname{Re}(B_n^{(2)}(Z), z_n)_{L^2} = \operatorname{Re}(B_n^{(2)}(Z), T_{\langle\xi\rangle^{2n}} z)_{L^2} + \operatorname{Re}(B_n^{(2)}(Z), T_{(1+\mathscr{L})^n-\langle\xi\rangle^{2n}} z)_{L^2}.$$

Using the bound (7-3) in Lemma 7.1 to estimate the operator $T_{(1+\mathscr{L})^n-\langle\xi\rangle^{2n}}$ and Lemma 3.7 and (7-11) to estimate the operator $B_n^{(2)}(Z)$, we get

$$|\operatorname{Re}(T_{\langle\xi\rangle^{2n}} X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z), T_{(1+\mathscr{L})^n-\langle\xi\rangle^{2n}} z)_{L^2}| + |\operatorname{Re}(B_n^{(2)}(Z), T_{(1+\mathscr{L})^n-\langle\xi\rangle^{2n}} z)_{L^2}| \lesssim \|u\|_{H^s}^{10},$$

which means that these remainders can be absorbed in the term $\mathscr{B}_{>5}(t)$. Then we set

$$\mathscr{B}(t) := 2\operatorname{Re}(B_n^{(2)}(Z), T_{\langle\xi\rangle^{2n}} z)_{L^2}.$$

Formulas (7-20) follow by an explicit computation using (7-9), (7-11). □

7A2. *Estimates of nonresonant terms.* In this subsection we provide estimates on the term $\mathscr{B}(t)$ appearing in (7-19). We state the main result of this section.

Proposition 7.6. *Let $N > 0$. Then there is $s_0 = s_0(N_0)$, where $N_0 > 0$ is given by Proposition 2.1, such that, if Hypothesis 6.1 holds with $s \geq s_0$, one has*

$$\left|\int_0^t \mathscr{B}(\sigma)\, d\sigma\right| \lesssim \|u\|_{L^\infty H^s}^{10} TN + \|u\|_{L^\infty H^s}^{6} T + \|u\|_{L^\infty H^s}^{4} TN^{-1} + \|u\|_{L^\infty H^s}^{4}, \tag{7-26}$$

where $\mathscr{B}(t)$ is in (7-20).

We need some preliminary results. We consider the trilinear maps

$$\mathscr{B}_i = \mathscr{B}_i[z_1, z_2, z_3], \quad \widehat{\mathscr{B}_i}(\xi) = \frac{1}{(2\pi)^d} \sum_{\eta,\zeta\in\mathbb{Z}^d} \mathtt{b}_i(\xi,\eta,\zeta)\hat{z}_1(\xi-\eta-\zeta)\hat{z}_2(\eta)\hat{z}_3(\zeta), \ i = 1, 2, \tag{7-27}$$

$$\mathscr{T}_< = \mathscr{T}_<[z_1, z_2, z_3], \quad \widehat{\mathscr{T}_<}(\xi) = \frac{1}{(2\pi)^d} \sum_{\eta,\zeta\in\mathbb{Z}^d} \mathtt{t}_<(\xi,\eta,\zeta)\hat{z}_1(\xi-\eta-\zeta)\hat{z}_2(\eta)\hat{z}_3(\zeta), \tag{7-28}$$

where

$$\mathtt{b}_1(\xi,\eta,\zeta) = \mathtt{b}(\xi,\eta,\zeta)\mathtt{1}_{\{\max\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}\leq N\}}, \tag{7-29}$$

$$\mathtt{b}_2(\xi,\eta,\zeta) = \mathtt{b}(\xi,\eta,\zeta)\mathtt{1}_{\{\max\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}> N\}}, \tag{7-30}$$

$$\mathtt{t}_<(\xi,\eta,\zeta) = \frac{-1}{\mathrm{i}\omega_{\mathrm{NLS}}(\xi,\eta,\zeta)}\mathtt{b}_1(\xi,\eta,\zeta), \tag{7-31}$$

where $\mathtt{b}(\xi,\eta,\zeta)$ are the coefficients in (7-20), and ω_{NLS} is the phase in (2-1). We remark that if $(\xi,\eta,\zeta) \in \mathscr{R}$ (see Definition 3.9) then the coefficients $\mathtt{b}(\xi,\eta,\zeta)$ are equal to zero (see (7-20), (7-9), (7-11)). Therefore, since ω_{NLS} is nonresonant (see Proposition 2.1), the coefficients in (7-31) are well-defined. We now prove an abstract results on the trilinear maps introduced in (7-27)–(7-28).

Lemma 7.7. *One has that, for $s = 2n > \frac{1}{2}d + 4$,*

$$\|\mathscr{B}_2[z_1, z_2, z_3]\|_{L^2} \lesssim N^{-1} \sum_{i=1}^{3} \|z_i\|_{H^s} \prod_{i \neq k} \|z_k\|_{H^{d/2+4+\epsilon}} \quad \textit{for all } \epsilon > 0. \tag{7-32}$$

There is $s_0(N_0) > 0$ ($N_0 > 0$ given by Proposition 2.1) such that for $s \geq s_0(N_0)$ one has

$$\|\mathscr{T}_<[z_1, z_2, z_3]\|_{H^p} \lesssim N \sum_{i=1}^{3} \|z_i\|_{H^{s+p-2}} \prod_{i \neq k} \|z_k\|_{H^{s_0}}, \quad p \in \mathbb{N}, \tag{7-33}$$

$$\|\mathscr{T}_<[z_1, z_2, z_3]\|_{L^2} \lesssim \sum_{i=1}^{3} \|z_i\|_{H^s} \prod_{i \neq k} \|z_k\|_{H^{s_0}}. \tag{7-34}$$

Proof. Using (7-30), (7-20), (7-11) we get

$$\begin{aligned}\|\mathscr{B}_2[z_1, z_2, z_3]\|_{L^2}^2 &\lesssim \sum_{\xi \in \mathbb{Z}^d} \Big(\sum_{\eta, \zeta \in \mathbb{Z}^d} |\mathtt{b}_2(\xi, \eta, \zeta)| \, |\hat{z}_1(\xi - \eta - \zeta)| \, |\hat{z}_2(\eta)| \, |\hat{z}_3(\zeta)| \Big)^2 \\ &\lesssim N^{-2} \sum_{\xi \in \mathbb{Z}^d} \Big(\sum_{\eta, \zeta \in \mathbb{Z}^d} \langle \xi \rangle^{2n} \max_2 \{|\xi - \eta - \zeta|, |\eta|, |\zeta|\}^4 |\hat{z}_1(\xi - \eta - \zeta)| \, |\hat{z}_2(\eta)| \, |\hat{z}_3(\zeta)| \Big)^2.\end{aligned}$$

Then, by reasoning as in the proof of Lemma 3.7, one obtains (7-32). Let us prove the bound (7-33) for $p = 0$; the others are similar. Using (7-31), (2-2), (7-20), (7-11) we have

$$\begin{aligned}\|\mathscr{T}_<[z_1, z_2, z_3]\|_{L^2}^2 &\lesssim \sum_{\xi \in \mathbb{Z}^d} \Big(\sum_{\eta, \zeta \in \mathbb{Z}^d} |\mathtt{t}_<(\xi, \eta, \zeta)| \, |\hat{z}_1(\xi - \eta - \zeta)| \, |\hat{z}_2(\eta)| \, |\hat{z}_3(\zeta)| \Big)^2 \\ &\lesssim_\gamma N^2 \sum_{\xi \in \mathbb{Z}^d} \Big(\sum_{\eta, \zeta \in \mathbb{Z}^d} \frac{\langle \xi \rangle^{2n} \max_2 \{|\xi - \eta - \zeta|, |\eta|, |\zeta|\}^{N_0+4}}{\max_1 \{|\xi - \eta - \zeta|, |\eta|, |\zeta|\}^2} |\hat{z}_1(\xi - \eta - \zeta)| \, |\hat{z}_2(\eta)| \, |\hat{z}_3(\zeta)| \Big)^2.\end{aligned}$$

Again, reasoning as in the proof of Lemma 3.7, one obtains (7-33). Inequality (7-34) follows similarly. □

Proof of Proposition 7.6. By (7-27), (7-29), (7-30) and recalling the definition of $\mathscr{B}$ in (7-20), we can write

$$\int_0^t \mathscr{B}(\sigma)\, d\sigma = \int_0^t (\mathscr{B}_1[z, \bar{z}, z], T_{\langle \xi \rangle^{2n}} z)_{L^2}\, d\sigma + \int_0^t (\mathscr{B}_2[z, \bar{z}, z], T_{\langle \xi \rangle^{2n}} z)_{L^2}\, d\sigma. \tag{7-35}$$

By Lemma 7.7 we have

$$\Big| \int_0^t (\mathscr{B}_2[z, \bar{z}, z], T_{\langle \xi \rangle^{2n}} z)_{L^2}\, d\sigma \Big| \overset{(7\text{-}32)}{\lesssim} N^{-1} \int_0^t \|z\|_{H^s}^4\, d\sigma \overset{(6\text{-}21)}{\lesssim} N^{-1} \int_0^t \|u\|_{H^s}^4\, d\sigma. \tag{7-36}$$

Consider now the first summand in the right-hand side of (7-35). We claim that we have the identity

$$\begin{aligned}\int_0^t (\mathscr{B}_1[z, \bar{z}, z], T_{\langle \xi \rangle^{2n}} z)_{L^2}\, d\sigma &= \int_0^t (\mathscr{T}_<[z, \bar{z}, z], T_{\langle \xi \rangle^{2n}} (\partial_t + \mathrm{i}\Lambda_{\mathrm{NLS}}) z)_{L^2}\, d\sigma \\ &\quad + \int_0^t (\mathscr{T}_<[(\partial_t + \mathrm{i}\Lambda_{\mathrm{NLS}}) z, \bar{z}, z], T_{\langle \xi \rangle^{2n}} z)_{L^2}\, d\sigma \\ &\quad + \int_0^t (\mathscr{T}_<[z, \bar{z}, (\partial_t + \mathrm{i}\Lambda_{\mathrm{NLS}}) z], T_{\langle \xi \rangle^{2n}} z)_{L^2}\, d\sigma \\ &\quad + \int_0^t (\mathscr{T}_<[z, \overline{(\partial_t + \mathrm{i}\Lambda_{\mathrm{NLS}}) z}, z], T_{\langle \xi \rangle^{2n}} z)_{L^2}\, d\sigma + O(\|u\|_{H^s}^4).\end{aligned} \tag{7-37}$$

We use the claim, postponing its proof. Consider the first summand in the right-hand side of (7-37). Using the self-adjointness of $T_{\langle\xi\rangle^2}$ and (7-7) we write

$$\begin{aligned}(\mathscr{T}_{<}[z,\bar z,z],\,T_{\langle\xi\rangle^{2n}}(\partial_t+\mathrm{i}\Lambda_{\mathrm{NLS}})z)_{L^2}&\\ =(T_{\langle\xi\rangle^2}\mathscr{T}_{<}[z,\bar z,z],\,-T_{\langle\xi\rangle^{2n-2}}\mathrm{i}T_{\Sigma}z)_{L^2}&+\big(\mathscr{T}_{<}[z,\bar z,z],\,T_{\langle\xi\rangle^{2n}}(X^{+}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z)+R_5^{(2,+)}(U))\big)_{L^2}.\end{aligned}$$

We estimate the first summand in the right-hand side by means of the Cauchy–Schwarz inequality, (7-33) with $p=2$ and (7-3); analogously we estimate the second summand by means of the Cauchy–Schwarz inequality, (7-34), (6-23) and (6-24), obtaining

$$\left|\int_0^t(\mathscr{T}_{<}[z,\bar z,z],\,T_{\langle\xi\rangle^{2n}}(\partial_t+\mathrm{i}\Lambda_{\mathrm{NLS}})z)_{L^2}\,d\sigma\right|\le\int_0^t\|u(\sigma)\|^{10}_{H^s}N+\|u(\sigma)\|^{6}_{H^s}\,d\sigma.$$

The other terms in (7-37) are estimated in a similar way. We eventually obtain (7-26).

We now prove the claim (7-37). Recalling (7-7) we have

$$\partial_t\hat z(\xi)=-\mathrm{i}\Lambda_{\mathrm{NLS}}(\xi)\hat z(\xi)+\widehat{\mathscr{Q}}(\xi),\quad \xi\in\mathbb{Z}^d,\qquad \mathscr{Q}:=-\mathrm{i}T_{\Sigma}z+X^{+}_{\mathrm{H}^{(4)}_{\mathrm{NLS}}}(Z)+R_5^{(2,+)}(U).$$

We define $\hat g(\xi):=e^{\mathrm{i}t\Lambda_{\mathrm{NLS}}(\xi)}\hat z(\xi)$ for all $\xi\in\mathbb{Z}^d$. One can note that $\hat g(\xi)$ satisfies

$$\partial_t\hat g(\xi)=e^{\mathrm{i}t\Lambda_{\mathrm{NLS}}(\xi)}\widehat{\mathscr{Q}}(\xi)=e^{\mathrm{i}t\Lambda_{\mathrm{NLS}}(\xi)}(\partial_t+\mathrm{i}\Lambda_{\mathrm{NLS}})\hat z(\xi)\quad\text{for all }\xi\in\mathbb{Z}^d.\tag{7-38}$$

According to this notation and using (7-27) and (2-1) we have

$$\begin{aligned}\int_0^t(\mathscr{B}_1[z,\bar z,z],\,T_{\langle\xi\rangle^{2n}}z)_{L^2}\,d\sigma&\\ =\int_0^t\sum_{\xi,\eta,\zeta\in\mathbb{Z}^d}\frac{1}{(2\pi)^d}\mathtt{b}_1(\xi,\eta,\zeta)e^{-\mathrm{i}\sigma\omega_{\mathrm{NLS}}(\xi,\eta,\zeta)}&\hat g(\xi-\eta-\zeta)\hat{\bar g}(\eta)\hat g(\zeta)\hat{\bar g}(-\xi)\langle\xi\rangle^{2n}\,d\sigma.\end{aligned}$$

By integrating by parts in σ and using (7-38) one gets (7-37) with

$$O(\|u\|^4_{H^s})=(\mathscr{T}_{<}[z(t),\bar z(t),z(t)],\,T_{\langle\xi\rangle^{2n}}z(t))_{L^2}-(\mathscr{T}_{<}[z(0),\bar z(0),z(0)],\,T_{\langle\xi\rangle^{2n}}z(0))_{L^2}.$$

The remainder above is bounded from above by $\|u\|^4_{L^\infty H^s}$ using Cauchy–Schwarz and (7-34). □

7B. *Estimates for the KG.* In this section we provide a priori energy estimates on the variable Z solving (6-48). This implies similar estimates on the solution U of the system (4-44) thanks to the equivalence (6-47). In Section 7B1 we introduce an equivalent energy norm and we provide a first energy inequality. This is the content of Proposition 7.10. Then in Section 7B2 we give improved bounds on the nonresonant terms.

7B1. *First energy inequality.* We recall that the system (6-48) is diagonal up to smoothing terms plus some higher degree of homogeneity remainder. Hence, for simplicity, we pass to the scalar equation

$$\partial_t z+\mathrm{i}\Lambda_{\mathrm{KG}}z=-\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(\tilde a_2^{+}(x,\xi)\Lambda_{\mathrm{KG}}(\xi))z+X^{+}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z)+R_4^{(2,+)}(u),\tag{7-39}$$

where (recall (6-49)) $X^{+}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z) = -\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}(a_0(x,\xi))z + Q^{+}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z)$. For $n \in \mathbb{R}$ we define

$$z_n := \langle D\rangle^n z, \quad Z_n = \begin{bmatrix} z_n \\ \bar{z}_n \end{bmatrix} := \mathbb{1}\langle D\rangle^n Z, \quad Z = \begin{bmatrix} z \\ \bar{z} \end{bmatrix}. \tag{7-40}$$

Lemma 7.8. *Fix $n := n(d) \gg 1$ large enough and recall* (7-39)*. One has that the function z_n defined in* (7-40) *solves the problem*

$$\partial_t z_n = -\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}\big((1+\tilde{a}_2^+(x,\xi))\Lambda_{\mathrm{KG}}(\xi)\big)z_n + \langle D\rangle^n X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z) + B_n^{(1)}(Z) + B_n^{(2)}(Z) + R_{4,n}(U), \tag{7-41}$$

where the resonant vector field $X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}$ is defined as in Definition 3.9 (*see also Remark 3.11*)*, the cubic terms $B_n^{(i)}$, $i = 1, 2$, have the form*

$$\widehat{B_n^{(1)}(Z)}(\xi) = \frac{1}{(2\pi)^d} \sum_{\substack{\sigma_1,\sigma_2\in\{\pm\} \\ \eta,\zeta\in\mathbb{Z}^d}} b_1^{\sigma_1,\sigma_2}(\xi,\eta,\zeta)\widehat{z^{\sigma_1}}(\xi-\eta-\zeta)\widehat{z^{\sigma_2}}(\eta)\hat{z}_n(\zeta), \tag{7-42}$$

$$\widehat{B_n^{(2)}(Z)}(\xi) = \frac{1}{(2\pi)^d} \sum_{\substack{\sigma_1,\sigma_2,\sigma_3\in\{\pm\} \\ \eta,\zeta\in\mathbb{Z}^d}} b_{2,n}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\widehat{z^{\sigma_1}}(\xi-\eta-\zeta)\widehat{z^{\sigma_2}}(\eta)\widehat{z^{\sigma_3}}(\zeta), \tag{7-43}$$

with (*recall Remark 4.5*)

$$b_1^{\sigma_1,\sigma_2}(\xi,\eta,\zeta) := -\mathrm{i}a_0^{\sigma_1,\sigma_2}\Big(\xi-\zeta,\eta,\frac{\xi+\zeta}{2}\Big)\chi_\epsilon\Big(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\Big)1_{\mathscr{R}^c}(\xi,\eta,\zeta), \tag{7-44}$$

$$|b_{2,n}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)| \lesssim \frac{\langle\xi\rangle^n \max_2\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}^\mu}{\max_1\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}}1_{\mathscr{R}^c}(\xi,\eta,\zeta) \tag{7-45}$$

for some $\mu > 1$. The remainder satisfies

$$\|R_{4,n}(U)\|_{L^2} \lesssim \|u\|^4_{H^n}. \tag{7-46}$$

Proof. Recalling the definition of resonant vector fields in Definition 3.9 we set

$$X^{+,\perp}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z) := X^{+}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z) - X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z), \tag{7-47}$$

which represents the nonresonant terms in the cubic vector field of (7-39). By differentiating in t (7-40) and using (7-39) we get

$$\begin{aligned} \partial_t z_n = {} & -\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}\big((1+\tilde{a}_2^+(x,\xi))\Lambda_{\mathrm{KG}}(\xi)\big)z_n + \langle D\rangle^n X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z) \\ & - \mathrm{i}[\langle D\rangle^n, \mathrm{Op}^{\mathrm{BW}}((1+\tilde{a}_2^+(x,\xi))\Lambda_{\mathrm{KG}}(\xi))]z \qquad (7\text{-}48) \\ & + \langle D\rangle^n X^{+,\perp}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z) \qquad (7\text{-}49) \\ & + \langle D\rangle^n R_4^{(2,+)}(u). \qquad (7\text{-}50) \end{aligned}$$

We analyze each summand above separately. By estimate (6-50) we deduce $\|(7\text{-}50)\|_{L^2} \lesssim \|u\|^4_{H^n}$. Let us now consider the commutator term in (7-48). By Lemma 3.1, Proposition 3.2 and the estimate on the seminorm of the symbol $\tilde{a}_2^+(x,\xi)$ in (6-38), we obtain that $\|(7\text{-}48)\|_{L^2} \lesssim \|u\|^3_{H^n}\|z\|_{H^n} \lesssim \|u\|^4_{H^n}$; we have used also (6-47). The term in (7-49) is the most delicate. By (6-49) and (7-47) (recall also Remark 4.5 and (3-6))

$$\langle D\rangle^n X^{+,\perp}_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(Z) = B_n^{(1)}(Z) + \mathscr{C}_1 + \mathscr{C}_2, \tag{7-51}$$

with $B_n^{(1)}(Z)$ as in (7-42) and coefficients as in (7-44), the term $\mathscr{C}_1$ has the form

$$\begin{aligned} \widehat{\mathscr{C}}_1(\xi) &= \frac{1}{(2\pi)^d} \sum_{\substack{\sigma_1,\sigma_2\in\{\pm\}\\ \eta,\zeta\in\mathbb{Z}^d}} \mathtt{c}_1^{\sigma_1,\sigma_2}(\xi,\eta,\zeta)\widehat{z^{\sigma_1}}(\xi-\eta-\zeta)\widehat{z^{\sigma_2}}(\eta)\hat{z}(\zeta),\\ \mathtt{c}_1^{\sigma_1,\sigma_2}(\xi,\eta,\zeta) &= -\mathrm{i}a_0^{\sigma_1,\sigma_2}\left(\xi-\zeta,\eta,\frac{\xi+\zeta}{2}\right)\chi_\epsilon\left(\frac{|\xi-\zeta|}{\langle\xi+\zeta\rangle}\right)[\langle\xi\rangle^n-\langle\zeta\rangle^n]1_{\mathscr{R}^c}(\xi,\eta,\zeta), \end{aligned} \tag{7-52}$$

and the term $\mathscr{C}_2$ has the form (7-43) with coefficients (see (5-30))

$$\mathtt{c}_2^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta) := \mathtt{q}^{\sigma_1,\sigma_2,\sigma_3}_{\mathtt{H}^{(4)}_{\mathrm{KG}}}(\xi,\eta,\zeta)\langle\xi\rangle^n 1_{\mathscr{R}^c}(\xi,\eta,\zeta). \tag{7-53}$$

In order to conclude the proof we need to show that the coefficients in (7-52), (7-53) satisfy the bound (7-45). This is true for the coefficients in (7-53) thanks to the bound (5-31). Moreover notice that

$$|\langle\xi\rangle^n - \langle\zeta\rangle^n| \lesssim |\xi-\zeta| \max\{\langle\xi\rangle,\langle\zeta\rangle\}^{n-1}.$$

Then the coefficients in (7-52) satisfy (7-45) by using Remark 4.5 and Lemma 3.8. □

Remark 7.9. In view of Remarks 4.6, 4.8, 6.14 if (KG) is semilinear then the symbol $\tilde{a}_2^+$ in (7-41) is equal to zero and the coefficients $b_{2,n}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)$ in (7-43) satisfy the bound (7-45) with the better denominator $\max_1\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}^2$.

In view of Lemma 7.8 we deduce the following.

Proposition 7.10. *Let $\psi(t,x)$ be a solution of* (KG) *satisfying Hypothesis 6.8 and consider the function z_n in* (7-40) *(see also* (6-46), (6-42)*). Then, setting $s=n=n(d)\gg 1$ we have $\|z_n\|_{L^2}\sim\|\psi\|_{H^{s+1/2}}+\|\dot\psi\|_{H^{s-1/2}}$ and*

$$\partial_t\|z_n(t)\|_{L^2}^2 = \mathscr{B}(t) + \mathscr{B}_{>4}(t), \quad t\in[0,T), \tag{7-54}$$

where:

- *The term $\mathscr{B}(t)$ has the form*

$$\mathscr{B}(t) = \sum_{\substack{\sigma_1,\sigma_2,\sigma_3\in\{\pm\}\\ \xi,\eta,\zeta\in\mathbb{Z}^d}} \langle\xi\rangle^{2n}\mathtt{b}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\widehat{z^{\sigma_1}}(\xi-\eta-\zeta)\widehat{z^{\sigma_2}}(\eta)\widehat{z^{\sigma_3}}(\zeta)\hat{\bar{z}}(-\xi), \tag{7-55}$$

where $\mathtt{b}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\in\mathbb{C}$ satisfy, for $\xi,\eta,\zeta\in\mathbb{Z}^d$,

$$|\mathtt{b}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)| \lesssim \frac{\max_2\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}^\mu}{\max_1\{\langle\xi-\eta-\zeta\rangle,\langle\eta\rangle,\langle\zeta\rangle\}} 1_{\mathscr{R}^c}(\xi,\eta,\zeta) \tag{7-56}$$

for some $\mu>1$.

- *The term $\mathscr{B}_{>5}(t)$ satisfies*

$$|\mathscr{B}_{>4}(t)| \lesssim \|u\|_{H^s}^5, \quad t\in[0,T). \tag{7-57}$$

Proof. The equivalence between $\|z_n\|_{L^2}$ and $\|\psi\|_{H^{s+1/2}} + \|\dot{\psi}\|_{H^{s-1/2}}$ follows by Remarks 6.12 and 6.9. By using (7-41) we get

$$\frac{1}{2}\partial_t \|z_n(t)\|_{L^2}^2 = \mathrm{Re}\big(-\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}((1+\tilde{a}_2^+(x,\xi))\Lambda_{\mathrm{KG}}(\xi))z_n, z_n\big)_{L^2} \tag{7-58}$$
$$+\mathrm{Re}(\langle D\rangle^n X^{+,\mathrm{res}}_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z), z_n)_{L^2} \tag{7-59}$$
$$+\mathrm{Re}(B_n^{(1)}(Z), z_n)_{L^2} \tag{7-60}$$
$$+\mathrm{Re}(B_n^{(2)}(Z), z_n)_{L^2} \tag{7-61}$$
$$+\mathrm{Re}(R_{4,n}(Z), z_n)_{L^2}. \tag{7-62}$$

By (6-34), (6-33) and (4-24) we have that the symbol $(1+\tilde{a}_2^+(x,\xi))\Lambda_{\mathrm{KG}}(\xi)$ is real-valued. Hence the operator $\mathrm{i}\,\mathrm{Op}^{\mathrm{BW}}((1+\tilde{a}_2^+(x,\xi))\Lambda_{\mathrm{KG}}(\xi))$ is skew-self-adjoint. We deduce (7-58)$\equiv 0$. By Lemma 3.10 (see also Remark 3.11) we have (7-59) $\equiv 0$. We also have (7-60) $\equiv 0$; to see this one can reason as done in the proof of Lemma 7.4, by using Remark 4.5, in particular (4-43). By formula (7-43) and estimates (7-45) we have that the term in (7-61) has the form (7-55) with coefficients satisfying (7-56). By the Cauchy–Schwarz inequality and estimate (7-46) we get that the term in (7-62) satisfies the bound (7-57). □

Remark 7.11. In view of Remark 7.9, if (KG) is semilinear, then the coefficients $\mathtt{b}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)$ of the energy in (7-55) satisfy the bound (7-56) with the better denominator $\max_1\{\langle\xi-\eta-\zeta\rangle, \langle\eta\rangle, \langle\zeta\rangle\}^2$.

7B2. *Estimates of nonresonant terms.* In Proposition 7.10 we provided a precise structure of the term $\mathscr{B}(t)$ of degree 4 in (7-54). In this section we show that, actually, $\mathscr{B}(t)$ satisfies better bounds with respect to a general quartic multilinear map by using that it is *nonresonant*. We state the main result of this section.

Proposition 7.12. *Let $N>0$ and let β be as in Proposition 2.2. Then there is $s_0 = s_0(N_0)$, where $N_0>0$ is given by Proposition 2.2, such that, if Hypothesis 6.8 holds with $s \geq s_0$, one has*

$$\left|\int_0^t \mathscr{B}(\sigma)\,d\sigma\right| \lesssim \|u\|^6_{L^\infty H^s} T N^{\beta-1} + \|u\|^7_{L^\infty H^s} N^\beta T + \|u\|^4_{L^\infty H^s} T N^{-1} + N^{\beta-1}\|u\|^4_{L^\infty H^s}, \tag{7-63}$$

where $\mathscr{B}(t)$ is in (7-55).

We first introduce some notation. Let $\vec{\sigma} := (\sigma_1,\sigma_2,\sigma_3) \in \{\pm\}^3$ and consider the trilinear maps

$$\mathscr{B}_i^{\vec{\sigma}} = \mathscr{B}_i^{\vec{\sigma}}[z_1,z_2,z_3], \quad \widehat{\mathscr{B}_i^{\vec{\sigma}}}(\xi) = \frac{1}{(2\pi)^d}\sum_{\eta,\zeta\in\mathbb{Z}^d} \langle\xi\rangle^s \mathtt{b}_i^{\vec{\sigma}}(\xi,\eta,\zeta)\widehat{z^{\sigma_1}}_1(\xi-\eta-\zeta)\widehat{z^{\sigma_2}}_2(\eta)\widehat{z^{\sigma_3}}_3(\zeta), \tag{7-64}$$

$$\mathscr{T}_<^{\vec{\sigma}} = \mathscr{T}_<^{\vec{\sigma}}[z_1,z_2,z_3], \quad \widehat{\mathscr{T}_<^{\vec{\sigma}}}(\xi) = \frac{1}{(2\pi)^d}\sum_{\eta,\zeta\in\mathbb{Z}^d} \langle\xi\rangle^s \mathtt{t}_<^{\vec{\sigma}}(\xi,\eta,\zeta)\hat{z}_1^{\sigma_1}(\xi-\eta-\zeta)\hat{z}_2^{\sigma_2}(\eta)\widehat{z^{\sigma_3}}_3(\zeta), \tag{7-65}$$

where

$$\mathtt{b}_1^{\vec{\sigma}}(\xi,\eta,\zeta) = \mathtt{b}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\mathbb{1}_{\{\max\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}\leq N\}}, \tag{7-66}$$
$$\mathtt{b}_2^{\vec{\sigma}}(\xi,\eta,\zeta) = \mathtt{b}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)\mathbb{1}_{\{\max\{|\xi-\eta-\zeta|,|\eta|,|\zeta|\}> N\}}, \tag{7-67}$$
$$\mathtt{t}_<^{\vec{\sigma}}(\xi,\eta,\zeta) = \frac{-1}{\mathrm{i}\omega_{\mathrm{KG}}^{\vec{\sigma}}(\xi,\eta,\zeta)}\mathtt{b}_1^{\vec{\sigma}}(\xi,\eta,\zeta), \tag{7-68}$$

where $\mathtt{b}^{\sigma_1,\sigma_2,\sigma_3}(\xi,\eta,\zeta)$ are the coefficients in (7-55), and $\omega_{\mathrm{KG}}^{\vec{\sigma}}$ is the phase in (2-4). We remark that if $(\xi,\eta,\zeta)\in\mathscr{R}$ (see Definition 3.9) then the coefficients $\mathtt{b}(\xi,\eta,\zeta)$ are equal to zero (see (7-55), (7-43), (7-45)). Therefore, since $\omega_{\mathrm{KG}}^{\vec{\sigma}}$ is nonresonant (see Proposition 2.2), the coefficients in (7-68) are well-defined. We now state an abstract result on the trilinear maps introduced in (7-64)–(7-65).

Lemma 7.13. *Let $\mu>1$ as in* (7-56). *One has that, for $s>\frac{1}{2}d+\mu$,*

$$\|\mathscr{B}_2^{\vec{\sigma}}[z_1,z_2,z_3]\|_{L^2}\lesssim N^{-1}\sum_{i=1}^{3}\|z_i\|_{H^s}\prod_{i\neq k}\|z_k\|_{H^{d/2+\mu+\epsilon}} \tag{7-69}$$

for any $\vec{\sigma}\in\{\pm\}^3$ and any $\epsilon>0$. There is $s_0(N_0)>0$ ($N_0>0$ given by Proposition 2.2) such that for $s\geq s_0(N_0)$ one has

$$\|\mathscr{T}_{<}[z_1,z_2,z_3]\|_{H^p}\lesssim N^{\beta}\sum_{i=1}^{3}\|z_i\|_{H^{s+p-1}}\prod_{i\neq k}\|z_k\|_{H^{s_0}},\qquad p\in\mathbb{N}, \tag{7-70}$$

$$\|\mathscr{T}_{<}[z_1,z_2,z_3]\|_{L^2}\lesssim N^{\beta-1}\sum_{i=1}^{3}\|z_i\|_{H^s}\prod_{i\neq k}\|z_k\|_{H^{s_0}}, \tag{7-71}$$

where β is defined in Proposition 2.2.

Proof. The proof is similar to that of Lemma 7.7. One has to use Proposition 2.2 instead of Proposition 2.1 to estimate the small divisors. □

Remark 7.14. In view of Remark 7.11, if (KG) is semilinear we may improve (7-69) and (7-71) with

$$\|\mathscr{B}_2^{\vec{\sigma}}[z_1,z_2,z_3]\|_{L^2}\lesssim N^{-2}\sum_{i=1}^{3}\|z_i\|_{H^s}\prod_{i\neq k}\|z_k\|_{H^{d/2+\mu+\epsilon}},$$

$$\|\mathscr{T}_{<}[z_1,z_2,z_3]\|_{L^2}\lesssim N^{\beta-2}\sum_{i=1}^{3}\|z_i\|_{H^s}\prod_{i\neq k}\|z_k\|_{H^{s_0}}.$$

We are now in a position to prove Proposition 7.12.

Proof of Proposition 7.12. By (7-64), (7-66), (7-67), and recalling the definition of $\mathscr{B}$ in (7-55), we can write

$$\int_0^t\mathscr{B}(\tau)\,d\tau=\sum_{\vec{\sigma}\in\{\pm\}^3}\int_0^t(\mathscr{B}_1^{\vec{\sigma}}[z,z,z],\langle D\rangle^s z)_{L^2}\,d\tau+\sum_{\vec{\sigma}\in\{\pm\}^3}\int_0^t(\mathscr{B}_2^{\vec{\sigma}}[z,z,z],\langle D\rangle^s z)_{L^2}\,d\tau. \tag{7-72}$$

By Lemma 7.13 we have

$$\left|\int_0^t(\mathscr{B}_2^{\vec{\sigma}}[z,z,z],\langle D\rangle^s z)_{L^2}\,d\sigma\right|\overset{(7\text{-}69)}{\lesssim}N^{-1}\int_0^t\|z\|_{H^s}^4\,d\tau\overset{(6\text{-}47)}{\lesssim}N^{-1}\int_0^t\|u\|_{H^s}^4\,d\tau. \tag{7-73}$$

Consider now the first summand in the right-hand side of (7-72). Integrating by parts as done in the proof of Proposition 7.6 we have

$$\begin{aligned}\int_0^t (\mathscr{B}_1^{\vec{\sigma}}[z,\bar{z},z],\langle D\rangle^s z)_{L^2}\,d\tau &= \int_0^t (\mathscr{T}_<^{\vec{\sigma}}[z,z,z],\langle D\rangle^s(\partial_t+\mathrm{i}\Lambda_{\mathrm{KG}})z)_{L^2}\,d\tau\\ &\quad+\int_0^t (\mathscr{T}_<^{\vec{\sigma}}[(\partial_t+\mathrm{i}\Lambda_{\mathrm{KG}})z,z,z],\langle D\rangle^s z)_{L^2}\,d\tau\\ &\quad+\int_0^t (\mathscr{T}_<^{\vec{\sigma}}[z,\bar{z},(\partial_t+\mathrm{i}\Lambda_{\mathrm{KG}})z],\langle D\rangle^s z)_{L^2}\,d\tau\\ &\quad+\int_0^t (\mathscr{T}_<^{\vec{\sigma}}[z,(\partial_t+\mathrm{i}\Lambda_{\mathrm{KG}})z,z],\langle D\rangle^s z)_{L^2}\,d\tau+R, \end{aligned}\tag{7-74}$$

where

$$R=(\mathscr{T}_<^{\vec{\sigma}}[z(t),z(t),z(t)],\langle D\rangle^s z(t))_{L^2}-(\mathscr{T}_<^{\vec{\sigma}}[z(0),z(0),z(0)],\langle D\rangle^s z(0))_{L^2}.$$

The remainder R above is bounded from above by $N^\beta\|u\|^4_{L^\infty H^s}$ using Cauchy–Schwarz and (7-70). Let us now consider the first summand in the right-hand side of (7-74). Using that the operator $\langle D\rangle$ is self-adjoint and recalling (7-39) we have

$$\begin{aligned}(\mathscr{T}_<^{\vec{\sigma}}[z,z,z],\langle D\rangle^s(\partial_t+\mathrm{i}\Lambda_{\mathrm{KG}})z)_{L^2} &= (\langle D\rangle\mathscr{T}_<^{\vec{\sigma}}[z,z,z],\langle D\rangle^{s-1}(\partial_t+\mathrm{i}\Lambda_{\mathrm{KG}})z)_{L^2}\\ &= (\langle D\rangle\mathscr{T}_<^{\vec{\sigma}}[z,z,z],\langle D\rangle^{s-1}\mathrm{Op}^{\mathrm{BW}}(-\mathrm{i}\tilde{a}_2^+(x,\xi)\Lambda_{\mathrm{KG}}(\xi))z)_{L^2} \quad (7\text{-}75)\\ &\quad+(\mathscr{T}_<^{\vec{\sigma}}[z,z,z],\langle D\rangle^s(X^+_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z)+R_4^{(2,+)}(u)))_{L^2}. \quad (7\text{-}76)\end{aligned}$$

By the Cauchy–Schwarz inequality, estimate (7-70) with $p=1$, estimate (6-38) on the seminorm of the symbol $\tilde{a}_2^+(x,\xi)$, Lemma 3.1 and the equivalence (6-47), we get $|(7\text{-}75)|\lesssim\|u\|^7_{H^s}N^\beta$. Consider the term in (7-76). First of all notice that, by (4-31) and Lemma 3.1, and by (5-31) and Lemma 3.7, the field $X_{\mathrm{H}^{(4)}_{\mathrm{KG}}}(Z)$ in (6-49) satisfies the same estimates (4-32) as the field $X_{\mathscr{H}^{(4)}_{\mathrm{KG}}}$. Therefore, using (7-71) and (6-50), we obtain $|(7\text{-}76)|\lesssim\|u\|^6_{H^s}N^{\beta-1}$. Using that (see Hypothesis 6.8) $\|u\|_{H^s}\ll 1$, we conclude that the first summand in the right-hand side of (7-74) is bounded from above by $N^\beta\int_0^t\|u(\tau)\|^7\,d\tau+N^{\beta-1}\int_0^t\|u(\tau)\|^6\,d\tau$. The other terms in (7-74) are estimated in a similar way. We eventually obtain (7-63). □

Remark 7.15. In view of Remarks 4.6, 4.8, 6.14, 7.9, 7.11 and 7.14, if (KG) is semilinear we have the better (with respect to (7-63)) estimate

$$\left|\int_0^t\mathscr{B}(\sigma)\,d\sigma\right|\lesssim\|u\|^6_{L^\infty H^s}TN^{\beta-2}+\|u\|^4_{L^\infty H^s}TN^{-2}+N^{\beta-2}\|u\|^4_{L^\infty H^s}.\tag{7-77}$$

8. Proof of the main results

In this section we conclude the proof of our main theorems.

Proof of Theorem 1. Consider (NLS) and let u_0 be as in the statement of Theorem 1. By the result in [Feola and Iandoli 2022] we have that there is $T>0$ and a unique solution $u(t,x)$ of (NLS) with $V\equiv 0$

such that Hypothesis 6.1 is satisfied. To recover the result when $V \neq 0$ one can argue as done in [Feola and Iandoli 2019]. Consider a potential V as in (1-5), with $\vec{x} \in \mathscr{O} \setminus \mathscr{N}$, where $\mathscr{N}$ is the zero measure set given in Proposition 2.1. We claim that we have the following a priori estimate: Fix any $0 < N$. Then for any $t \in [0, T)$, with T as in Hypothesis 6.1, one has

$$\|u(t)\|_{H^s}^2 \leq 2\|u_0\|_{H^s}^2 + C(\|u\|_{L^\infty H^s}^{10} TN + \|u\|_{L^\infty H^s}^6 T + \|u\|_{L^\infty H^s}^4 TN^{-1} + \|u\|_{L^\infty H^s}^4) \tag{8-1}$$

for some $C > 0$ depending on s. To prove the claim we reason as follows. By Proposition 4.2 we have that (NLS) is equivalent to the system (4-12). By Propositions 6.4, 6.6 and Lemma 7.3 we can construct a function z_n with $2n = s$ such that if $u(t,x)$ solves the (NLS) then z_n solves (7-8). Moreover by Proposition 7.5 we have the equivalence (7-18), and we deduce

$$\|u(t)\|_{H^s}^2 \leq 2^{1/2}\|z_n(t)\|_{L^2}^2 \leq 2\|u_0\|_{H^s}^2 + 2\left|\int_0^t \mathscr{B}(\sigma)\, d\sigma\right| + 2\left|\int_0^t \mathscr{B}_{>5}(\sigma)\, d\sigma\right|. \tag{8-2}$$

Propositions 7.5 and 7.6 apply; therefore, by (7-26) and (7-21), we obtain (8-1). The thesis of Theorem 1 follows from the following lemma.

Lemma 8.1 (main bootstrap). *Let $u(t,x)$ be a solution of* (NLS) *with $t \in [0, T)$ and initial condition $u_0 \in H^s(\mathbb{T}^d; \mathbb{C})$. Then, for $s \gg 1$ large enough, there exist $\epsilon_0, c_0 > 0$ such that, for any $0 < \epsilon \leq \epsilon_0$, if*

$$\|u_0\|_{H^s} \leq \tfrac{1}{4}\epsilon, \qquad \sup_{t\in[0,T)} \|u(t)\|_{H^s} \leq \epsilon, \qquad T \leq c_0\epsilon^{-4}, \tag{8-3}$$

then we have the improved bound $\sup_{t\in[0,T)} \|u(t)\|_{H^s} \leq \frac{1}{2}\epsilon$.

Proof. For ϵ small enough the bound (8-1) holds true, and we fix $N := \epsilon^{-3}$. Therefore, there is $C = C(s) > 0$ such that, for any $t \in [0, T)$,

$$\begin{aligned} \|u(t)\|_{H^s}^2 &\leq 2\|u_0\|_{H^s}^2 + C(\|u\|_{L^\infty H^s}^4 + \|u\|_{L^\infty H^s}^{10} T\epsilon^{-3} + \|u\|_{L^\infty H^s}^6 T + \|u\|_{L^\infty H^s}^4 T\epsilon^3) \\ &\leq \tfrac{1}{8}\epsilon^2 + C(\epsilon^4 + 2\epsilon^7 T + \epsilon^6 T) \qquad\qquad \text{(by (8-3))} \\ &\leq \tfrac{1}{4}\epsilon^2\big(\tfrac{1}{2} + 4C(\epsilon^4 + 2\epsilon c_0 + c_0)\big) \leq \tfrac{1}{4}\epsilon^2, \end{aligned} \tag{8-4}$$

where in the last inequality we have chosen c_0 and ϵ sufficiently small. This implies the thesis. ☐

Proof of Theorem 2. One has to follow almost word by word the proof of Theorem 1. The only difference relies on the estimates on the small divisors, which in this case are given by Proposition 2.1(ii).

Proof of Theorem 3. Consider (KG) and let (ψ_0, ψ_1) be as in the statement of Theorem 3. Let $\psi(t,x)$ be a solution of (KG) satisfying the condition in Hypothesis 6.8. By Proposition 4.7, recall (3-77), the function

$$U := \begin{bmatrix} u \\ \bar{u} \end{bmatrix}$$

solves (4-12) with initial condition

$$u_0 = \frac{1}{\sqrt{2}}(\Lambda_{\mathrm{KG}}^{1/2}\psi_0 + \mathrm{i}\Lambda_{\mathrm{KG}}^{-1/2}\psi_1).$$

Moreover, by Hypothesis 6.8 and Remark 6.9 one has $\|u_0\|_{H^s} \le \frac{1}{16}\epsilon$. By Remark 6.9, in order to get (1-8), we have to show that the bound $\sup_{t\in[0,T)} \|u\|_{H^s} \le \frac{1}{4}\epsilon$ holds for time $T \gtrsim \epsilon^{-3^-}$ if $d=2$ and $T \gtrsim \epsilon^{-8/3^-}$ if $d \ge 3$. Fix β as in Proposition 2.2 and let $m \in \mathscr{C}_\beta$. By Propositions 6.11, 6.13 and Lemma 7.8 we can construct a function z_n with $n=s$ such that if $\psi(t,x)$ solves (KG) then z_n solves (7-41). By Proposition 7.10 and Remark 6.12 we get

$$\|u(t)\|_{H^s}^2 \le 2^{1/2}\|z_n(t)\|_{L^2}^2 \le 2\|u_0\|_{H^s}^2 + 2\left|\int_0^t \mathscr{B}(\sigma)\,d\sigma\right| + 2\left|\int_0^t \mathscr{B}_{>5}(\sigma)\,d\sigma\right|. \tag{8-5}$$

Propositions 7.10 and 7.12 apply, therefore, by (7-63) and (7-57), we obtain the following a priori estimate: Fix any $0 < N$. Then for any $t \in [0,T)$, with T as in Hypothesis 6.8, one has

$$\begin{aligned}\|u(t)\|_{H^s}^2 &\le 2\|u_0\|_{H^s}^2 \\ &\quad + C\big(\|u\|_{L^\infty H^s}^6 T N^{\beta-1} + \|u\|_{L^\infty H^s}^7 T N^{\beta} + \|u\|_{L^\infty H^s}^6 T + \|u\|_{L^\infty H^s}^4 T N^{-1} + N^{\beta-1}\|u\|_{L^\infty H^s}^4\big)\end{aligned} \tag{8-6}$$

for some $C > 0$ depending on s. The thesis of Theorem 3 follows from the lemma below.

Lemma 8.2 (main bootstrap). *Let $u(t,x)$ be a solution of* (4-44) *with $t \in [0,T)$ and initial condition $u_0 \in H^s(\mathbb{T}^d;\mathbb{C})$. Define $\mathtt{a} = 3$ if $d=2$ and $\mathtt{a} = \frac{8}{3}$ if $d \ge 3$. Then, for $s \gg 1$ large enough and any $\delta > 0$, there exists $\epsilon_0 = \epsilon_0(d,s,m,\delta) > 0$ such that, for any $0 < \epsilon \le \epsilon_0$, if*

$$\|u_0\|_{H^s} \le \tfrac{1}{16}\epsilon, \qquad \sup_{t\in[0,T)} \|u(t)\|_{H^s} \le \tfrac{1}{4}\epsilon, \qquad T \le \epsilon^{-\mathtt{a}+\delta}, \tag{8-7}$$

then we have the improved bound $\sup_{t\in[0,T)} \|u(t)\|_{H^s} \le \frac{1}{8}\epsilon$.

Proof. We start with $d \ge 3$. For ϵ small enough the bound (8-6) holds true. Let $\delta > 0$ and $0 < \sigma \ll \delta$. Define

$$\beta := 3 + \sigma, \qquad N := \epsilon^{-2/(3+\sigma)}. \tag{8-8}$$

By (8-6), (8-7), (8-8), there is $C = C(s) > 0$ such that, for any $t \in [0,T)$,

$$\|u(t)\|_{H^s}^2 \le 2\tfrac{1}{16^2}\epsilon^2 + C\epsilon^2\epsilon^{2/(3+\sigma)} + 2CT\epsilon^2(\epsilon^3 + \epsilon^{2+2/(3+\sigma)}) \le \tfrac{1}{64}\epsilon^2, \tag{8-9}$$

where in the last inequality we have chosen ϵ sufficiently small and we used the choice of T in (8-7) and that $\sigma \ll \delta$. This implies the thesis for $d \ge 3$. In the case $d=2$ the proof is similar setting $\beta = 2+\sigma$ and $N = \epsilon^{-2/(2+\sigma)}$. □

Proof of Theorem 4. Using Remarks 4.6, 4.8, 6.14, 7.9, 7.11, 7.14, 7.15 one deduces the result by reasoning as in the proof of Theorem 3 and using in particular the estimate (7-77).

Acknowledgements

We would like to warmly thank Prof. Fabio Pusateri, Joackim Bernier and the anonymous referees for their useful comments.

References

[Baldi 2013] P. Baldi, "Periodic solutions of fully nonlinear autonomous equations of Benjamin–Ono type", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **30**:1 (2013), 33–77. MR Zbl

[Baldi et al. 2018] P. Baldi, E. Haus, and R. Montalto, "Controllability of quasi-linear Hamiltonian NLS equations", *J. Differential Equations* **264**:3 (2018), 1786–1840. MR Zbl

[Bambusi 2003] D. Bambusi, "Birkhoff normal form for some nonlinear PDEs", *Comm. Math. Phys.* **234**:2 (2003), 253–285. MR Zbl

[Bambusi and Grébert 2006] D. Bambusi and B. Grébert, "Birkhoff normal form for partial differential equations with tame modulus", *Duke Math. J.* **135**:3 (2006), 507–567. MR Zbl

[Bambusi et al. 2007] D. Bambusi, J.-M. Delort, B. Grébert, and J. Szeftel, "Almost global existence for Hamiltonian semilinear Klein–Gordon equations with small Cauchy data on Zoll manifolds", *Comm. Pure Appl. Math.* **60**:11 (2007), 1665–1690. MR Zbl

[Bernier et al. 2020] J. Bernier, E. Faou, and B. Grébert, "Long time behavior of the solutions of NLW on the d-dimensional torus", *Forum Math. Sigma* **8** (2020), art. id. e12. MR Zbl

[Berti and Delort 2018] M. Berti and J.-M. Delort, *Almost global solutions of capillary-gravity water waves equations on the circle*, Lect. Notes Unione Mat. Ital. **24**, Springer, 2018. MR Zbl

[Berti et al. 2021a] M. Berti, R. Feola, and L. Franzoi, "Quadratic life span of periodic gravity-capillary water waves", *Water Waves* **3**:1 (2021), 85–115. MR Zbl

[Berti et al. 2021b] M. Berti, R. Feola, and F. Pusateri, "Birkhoff normal form for gravity water waves", *Water Waves* **3**:1 (2021), 117–126. MR Zbl

[Berti et al. 2021c] M. Berti, A. Maspero, and F. Murgante, "Local well posedness of the Euler–Korteweg equations on $\mathbb{T}^d$", *J. Dynam. Differential Equations* **33**:3 (2021), 1475–1513. MR Zbl

[Berti et al. 2023] M. Berti, R. Feola, and F. Pusateri, "Birkhoff normal form and long time existence for periodic gravity water waves", *Comm. Pure Appl. Math.* **76**:7 (2023), 1416–1494. MR

[Bony 1981] J.-M. Bony, "Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires", *Ann. Sci. École Norm. Sup.* (4) **14**:2 (1981), 209–246. MR Zbl

[Borovskiĭ and Galkin 1993] A. V. Borovskiĭ and A. L. Galkin, "Dynamic modulation of an ultrashort high-intensity laser pulse in matter", *J. Exp. Theor. Phys.* **77**:4 (1993), 562–573.

[de Bouard et al. 1997] A. de Bouard, N. Hayashi, and J.-C. Saut, "Global existence of small solutions to a relativistic nonlinear Schrödinger equation", *Comm. Math. Phys.* **189**:1 (1997), 73–105. MR Zbl

[Colin 2002] M. Colin, "On the local well-posedness of quasilinear Schrödinger equations in arbitrary space dimension", *Comm. Partial Differential Equations* **27**:1-2 (2002), 325–354. MR Zbl

[Colin 2003] M. Colin, "Stability of stationary waves for a quasilinear Schrödinger equation in space dimension 2", *Adv. Differential Equations* **8**:1 (2003), 1–28. MR Zbl

[Colin and Jeanjean 2004] M. Colin and L. Jeanjean, "Solutions for a quasilinear Schrödinger equation: a dual approach", *Nonlinear Anal.* **56**:2 (2004), 213–226. MR Zbl

[Delort 2009] J.-M. Delort, "On long time existence for small solutions of semi-linear Klein–Gordon equations on the torus", *J. Anal. Math.* **107** (2009), 161–194. MR Zbl

[Delort 2012] J.-M. Delort, *A quasi-linear Birkhoff normal forms method: application to the quasi-linear Klein–Gordon equation on* $\mathbb{S}^1$, Astérisque **341**, Soc. Math. France, Paris, 2012. MR Zbl

[Delort 2015] J.-M. Delort, *Quasi-linear perturbations of Hamiltonian Klein–Gordon equations on spheres*, Mem. Amer. Math. Soc. **1103**, Amer. Math. Soc., Providence, RI, 2015. MR Zbl

[Delort 2016] J.-M. Delort, "Semiclassical microlocal normal forms and global solutions of modified one-dimensional KG equations", *Ann. Inst. Fourier* (*Grenoble*) **66**:4 (2016), 1451–1528. MR Zbl

[Delort and Szeftel 2004] J.-M. Delort and J. Szeftel, "Long-time existence for small data nonlinear Klein–Gordon equations on tori and spheres", *Int. Math. Res. Not.* **2004**:37 (2004), 1897–1966. MR Zbl

[Delort and Szeftel 2006] J.-M. Delort and J. Szeftel, "Long-time existence for semi-linear Klein–Gordon equations with small Cauchy data on Zoll manifolds", *Amer. J. Math.* **128**:5 (2006), 1187–1218. MR Zbl

[Eliasson 2002] L. H. Eliasson, "Perturbations of linear quasi-periodic system", pp. 1–60 in *Dynamical systems and small divisors* (Cetraro, Italy, 1998), edited by S. Marmi and J.-C. Yoccoz, Lecture Notes in Math. **1784**, Springer, 2002. MR Zbl

[Eliasson et al. 2016] L. H. Eliasson, B. Grébert, and S. B. Kuksin, "KAM for the nonlinear beam equation", *Geom. Funct. Anal.* **26**:6 (2016), 1588–1715. MR Zbl

[Fang and Zhang 2010] D. Fang and Q. Zhang, "Long-time existence for semi-linear Klein–Gordon equations on tori", *J. Differential Equations* **249**:1 (2010), 151–179. MR Zbl

[Faou and Grébert 2013] E. Faou and B. Grébert, "A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus", *Anal. PDE* **6**:6 (2013), 1243–1262. MR Zbl

[Faou et al. 2013] E. Faou, L. Gauckler, and C. Lubich, "Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus", *Comm. Partial Differential Equations* **38**:7 (2013), 1123–1140. MR Zbl

[Feola and Iandoli 2019] R. Feola and F. Iandoli, "Local well-posedness for quasi-linear NLS with large Cauchy data on the circle", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **36**:1 (2019), 119–164. MR Zbl

[Feola and Iandoli 2020] R. Feola and F. Iandoli, "A non-linear Egorov theorem and Poincaré–Birkhoff normal forms for quasi-linear PDEs on the circle", preprint, 2020. arXiv 2002.12448

[Feola and Iandoli 2021] R. Feola and F. Iandoli, "Long time existence for fully nonlinear NLS with small Cauchy data on the circle", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **22**:1 (2021), 109–182. MR Zbl

[Feola and Iandoli 2022] R. Feola and F. Iandoli, "Local well-posedness for the quasi-linear Hamiltonian Schrödinger equation on tori", *J. Math. Pures Appl.* (9) **157** (2022), 243–281. MR Zbl

[Feola and Procesi 2015] R. Feola and M. Procesi, "Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations", *J. Differential Equations* **259**:7 (2015), 3389–3447. MR Zbl

[Grébert 2007] B. Grébert, "Birkhoff normal form and Hamiltonian PDEs", pp. 1–46 in *Partial differential equations and applications* (Lanzhou, China, 2004), edited by X. P. Wang and C. Zhong, Sémin. Congr. **15**, Soc. Math. France, Paris, 2007. MR Zbl

[Hasse 1980] R. W. Hasse, "A general method for the solution of nonlinear soliton and kink Schrödinger equations", *Z. Phys. B* **37**:1 (1980), 83–87. MR

[Ionescu and Pusateri 2015] A. D. Ionescu and F. Pusateri, "Global solutions for the gravity water waves system in 2D", *Invent. Math.* **199**:3 (2015), 653–804. MR Zbl

[Ionescu and Pusateri 2018] A. D. Ionescu and F. Pusateri, *Global regularity for 2D water waves with surface tension*, Mem. Amer. Math. Soc. **1227**, Amer. Math. Soc., Providence, RI, 2018. MR Zbl

[Ionescu and Pusateri 2019] A. D. Ionescu and F. Pusateri, "Long-time existence for multi-dimensional periodic water waves", *Geom. Funct. Anal.* **29**:3 (2019), 811–870. MR Zbl

[Kato 1975] T. Kato, "Quasi-linear equations of evolution, with applications to partial differential equations", pp. 25–70 in *Spectral theory and differential equations* (Dundee, Scotland, 1974), edited by W. N. Everitt, Lecture Notes in Math. **448**, Springer, 1975. MR Zbl

[Kenig et al. 2004] C. E. Kenig, G. Ponce, and L. Vega, "The Cauchy problem for quasi-linear Schrödinger equations", *Invent. Math.* **158**:2 (2004), 343–388. MR Zbl

[Laurie et al. 2010] J. Laurie, V. S. L'vov, S. Nazarenko, and O. Rudenko, "Interaction of Kelvin waves and nonlocality of energy transfer in superfluids", *Phys. Rev. B* **81**:10 (2010), art. id. 104526.

[Litvak and Sergeev 1978] A. G. Litvak and A. M. Sergeev, "One-dimensional collapse of plasma waves", *JETP Lett.* **27**:10 (1978), 517–520.

[Makhankov and Fedyanin 1984] V. G. Makhankov and V. K. Fedyanin, "Nonlinear effects in quasi-one-dimensional models of condensed matter theory", *Phys. Rep.* **104**:1 (1984), 1–86. MR

[Marzuola et al. 2021] J. L. Marzuola, J. Metcalfe, and D. Tataru, "Quasilinear Schrödinger equations, III: Large data and short time", *Arch. Ration. Mech. Anal.* **242**:2 (2021), 1119–1175. MR Zbl

[Métivier 2008] G. Métivier, *Para-differential calculus and applications to the Cauchy problem for nonlinear systems*, CRM Ser. **5**, Edizioni della Normale, Pisa, 2008. MR Zbl

[Poppenberg 2001] M. Poppenberg, "Smooth solutions for a class of fully nonlinear Schrödinger type equations", *Nonlinear Anal.* **45**:6 (2001), 723–741. MR Zbl

[Porkolab and Goldman 1976] M. Porkolab and M. V. Goldman, "Upper-hybrid solitons and oscillating-two-stream instabilities", *Phys. Fluids* **19**:6 (1976), 872–881. MR

[Stingo 2018] A. Stingo, "Global existence and asymptotics for quasi-linear one-dimensional Klein–Gordon equations with mildly decaying Cauchy data", *Bull. Soc. Math. France* **146**:1 (2018), 155–213. MR Zbl

[Taylor 2000] M. E. Taylor, *Tools for PDE: pseudodifferential operators, paradifferential operators, and layer potentials*, Math. Surv. Monogr. **81**, Amer. Math. Soc., Providence, RI, 2000. MR Zbl

Received 5 Oct 2020. Revised 14 Sep 2021. Accepted 7 Jan 2022.

ROBERTO FEOLA: roberto.feola@uniroma3.it
Dipartimento di Matematica e Fisica, Universitá degli studi Roma Tre, Rome, Italy

BENOÎT GRÉBERT: benoit.grebert@univ-nantes.fr
Laboratoire de Mathématiques Jean Leray, Université de Nantes, UMR CNRS 6629, Nantes, France

FELICE IANDOLI: felice.iandoli@unical.it
Dipartimento di Matematica e Informatica, Università della Calabria, Rende, Italy

ANALYSIS AND PDE
Vol. 16 (2023), No. 5, pp. 1205–1244
DOI: 10.2140/apde.2023.16.1205

AN EXTENSION PROBLEM, TRACE HARDY AND HARDY'S INEQUALITIES FOR THE ORNSTEIN–UHLENBECK OPERATOR

PRITAM GANGULY, RAMESH MANNA AND SUNDARAM THANGAVELU

We study an extension problem for the Ornstein–Uhlenbeck operator $L = -\Delta + 2x\cdot\nabla + n$, and we obtain various characterisations of the solution of the same. We use a particular solution of that extension problem to prove a trace Hardy inequality for L from which Hardy's inequality for fractional powers of L is obtained. We also prove an isometry property of the solution operator associated to the extension problem. Moreover, new $L^p - L^q$ estimates are obtained for the fractional powers of the Hermite operator.

1. Introduction and the main results 1205
2. The extension problem for the Ornstein–Uhlenbeck operator and fractional powers 1211
3. Characterisations of solutions of the extension problem 1217
4. Trace Hardy and Hardy's inequality 1229
5. Isometry property for the solution of the extension problem 1234
6. Hardy–Littlewood–Sobolev inequality for H_s 1238
Acknowledgements 1242
References 1242

1. Introduction and the main results

It is said that analysts are obsessed with inequalities. The usefulness of various weighted and unweighted inequalities in applications to problems in differential geometry, quantum mechanics, partial differential equations, etc., have made this a very attractive area of research. Hardy's inequality is one such inequality which finds its origin in an old paper of G. H. Hardy [1919] written more than a hundred years ago; see also [Hardy 1920]. In recent years, this has been intensively studied in different settings and various contexts. For a historical review of Hardy's inequality, we refer the reader to [Kufner et al. 2007].

We begin by recalling the classical Hardy's inequality which states that, given $f \in C_0^\infty(\mathbb{R}^n)$,

$$\frac{1}{4}(n-2)^2 \int_{\mathbb{R}^n} \frac{|f(x)|^2}{|x|^2}\,dx \le \int_{\mathbb{R}^n} |\nabla f(x)|^2\,dx, \quad n \ge 3,$$

where ∇ denotes the gradient in $\mathbb{R}^n$. This can be rephrased as follows in terms of the Euclidean Laplacian $\Delta := \sum_{j=1}^n \partial^2/\partial x_j^2$:

$$\frac{1}{4}(n-2)^2 \int_{\mathbb{R}^n} \frac{|f(x)|^2}{|x|^2}\,dx \le \langle(-\Delta)f, f\rangle,$$

MSC2020: primary 26D10, 35J15; secondary 26A33, 33C45, 35A23, 43A80.
Keywords: extension problem, trace Hardy inequality, Hardy's inequality, Ornstein–Uhlenbeck operator.

which has been generalised to fractional powers of the Laplacian. In fact, for $0 < s < \frac{1}{2}n$ and $f \in C_0^\infty(\mathbb{R}^n)$, we have

$$4^s \frac{\Gamma\big(\frac{1}{4}(n+2s)\big)^2}{\Gamma\big(\frac{1}{4}(n-2s)\big)^2} \int_{\mathbb{R}^n} \frac{|f(x)|^2}{|x|^{2s}}\, dx \le \langle (-\Delta)^s f, f\rangle. \tag{1-1}$$

The constant appearing on the left-hand side is known to be sharp (see [Beckner 2012; Yafaev 1999] for instance), but the equality is never achieved. Frank et al. [2008] used a ground state representation to give a new proof of (1-1) when $0 < s < \min\{1, \frac{1}{2}n\}$, improving the previous results. On the other hand, replacing the homogeneous weight $|x|^{-2s}$ by a nonhomogeneous weight we have the following version of Hardy's inequality:

$$4^s \frac{\Gamma\big(\frac{1}{4}(n+2s)\big)}{\Gamma\big(\frac{1}{4}(n-2s)\big)} \rho^{2s} \int_{\mathbb{R}^n} \frac{|f(x)|^2}{(\rho^2+|x|^2)^{2s}}\, dx \le \langle (-\Delta)^s f, f\rangle, \quad \rho > 0, \tag{1-2}$$

where the constant is sharp and the equality is achieved for the functions $(\rho^2+|x|^2)^{-(n-2s)/2}$. See [Boggarapu et al. 2019, Remark 2.6] for a proof of inequality (1-2). Note that proving such an inequality for fractional powers depends on how one views this type of operator. In fact, there are several ways of obtaining fractional powers of the Laplacian. Caffarelli and Silvestre [2007] first studied an extension problem associated to the Laplacian on $\mathbb{R}^n$ and obtained the fractional power as a mapping which takes Dirichlet data to the Neumann data. Motivated by this work, [Boggarapu et al. 2019] studied the extension problem in a more general setting of sums of squares of vector fields on certain stratified Lie groups. They used a solution of that extension problem to prove a trace Hardy inequality from which Hardy's inequality is obtained. Because of its several interesting features, the study of extension problems for various operators has received considerable attention in recent times, see e.g., [Roncal and Thangavelu 2020b; Stinga and Torrea 2010], etc.

Inspired by [Frank et al. 2015], Roncal and Thangavelu [2020a] considered a modified extension problem for the sub-Laplacian on the H-type groups which gives conformally invariant fractional powers of the sub-Laplacian, and they proved Hardy's inequality for the same. In this regard, we would also like to mention that Garofalo and Tralli [2021] recently used an extension problem for the heat operator associated to the sub-Laplacian on the H-type groups to study the usual and conformal fractional powers of the sub-Laplacian. See also [Garofalo and Tralli 2023] by the same authors in this direction.

Although this fractional Hardy-type inequality has been studied extensively in the setting of Euclidean harmonic analysis, not much has been studied in the framework of Gaussian harmonic analysis. As we know that the role of the Laplacian in Gaussian harmonic analysis is played by the Ornstein–Uhlenbeck operator defined by $\tilde{L} := -\Delta + 2x \cdot \nabla$, it is therefore natural to ask for such a fractional Hardy inequality for this operator. It is also convenient to work with $L := -\Delta + 2x \cdot \nabla + n$ instead of $\tilde{L}$. In fact, from the mathematical point of view, L is very closely related to the Hermite operator; see (1-3) below. Later in this article, this relationship will be discussed and exploited in some of our studies. Because of its various applications in probability theory, stochastic calculus, etc., the study of the Ornstein–Uhlenbeck operator experienced a lot of developments in the last couple of decades. We refer the reader to the book of Urbina-Romero [2019] in this regard.

Our aim in this article is to establish Hardy and trace Hardy inequalities for fractional powers of the Ornstein–Uhlenbeck operator L. Recall that $L = -\Delta + 2x \cdot \nabla + n$ can be defined on the Gaussian L^2 space: $L^2(\gamma) = L^2(\mathbb{R}^n, \gamma(x)\,dx)$ with $\gamma(x) = \pi^{-n/2}e^{-|x|^2}$ is a positive self-adjoint operator. We observe that $\sum_{j=1}^n \partial_j^* \partial_j = -\Delta + 2x \cdot \nabla$, where $\partial_j = \partial/\partial x_j$ and $\partial_j^* = 2x_j - \partial_j$ is its adjoint on $L^2(\gamma)$. The relation between L and the Hermite operator $H = -\Delta + |x|^2$ is given by

$$M_\gamma L M_\gamma^{-1} = H, \quad \text{where } M_\gamma f(x) = \gamma(x)^{1/2} f(x). \tag{1-3}$$

Hardy's inequality for the fractional powers H^s of the Hermite operator has been studied in [Ciaurri et al. 2018]. Here H^s is defined by spectral theory as

$$H^s = \sum_{k=0}^{\infty} (2k+n)^s P_k,$$

where $(2k+n)$, $k \in \mathbb{N}$ are the eigenvalues of H on $L^2(\mathbb{R}^n)$ and P_k is the orthogonal projection of $L^2(\mathbb{R}^n)$ onto the finite-dimensional eigenspace corresponding to the eigenvalue $(2k+n)$. However, there is another natural candidate for fractional powers of H, and hence of L, which will be treated here.

To motivate the new definition of fractional powers, denoted by H_s, it is better to recall the conformally invariant fractional powers of the sub-Laplacian $\mathcal{L}$ on the Heisenberg group $\mathbb{H}^n$. The connection between $\mathcal{L}$ and H is given by the relation $\pi_\lambda(\mathcal{L}f) = \pi_\lambda(f) H(\lambda)$, where the π_λ are the Schrödinger representations of $\mathbb{H}^n$ and $H(\lambda) = -\Delta + \lambda^2|x|^2$. The spectral decomposition of $H(\lambda)$ is given by

$$H(\lambda) = \sum_{k=0}^{\infty} (2k+n)|\lambda| P_k(\lambda).$$

The conformally invariant fractional powers of $\mathcal{L}$ are then defined, for $0 < s < (n+1)$, by the relation

$$\pi_\lambda(\mathcal{L}_s f) = \pi_\lambda(f) \sum_{k=0}^{\infty} (2|\lambda|)^s \frac{\Gamma\big(\frac{1}{2}(2k+n+1+s)\big)}{\Gamma\big(\frac{1}{2}(2k+n+1-s)\big)} P_k(\lambda).$$

The operator on the right-hand side which multiplies $\pi_\lambda(f)$ is the alternate candidate for fractional powers of $H(\lambda)$, which we denote by $H(\lambda)_s$. By defining $Q_k = M_\gamma^{-1} P_k M_\gamma$, the spectral decomposition of L becomes $L = \sum_{k=0}^{\infty} (2k+n) Q_k$, and hence the fractional powers we are interested in are given by

$$L_s f(x) = \sum_{k=0}^{\infty} 2^s \frac{\Gamma\big(\frac{1}{2}(2k+n+1+s)\big)}{\Gamma\big(\frac{1}{2}(2k+n+1-s)\big)} Q_k f(x).$$

Along with L we also consider $U = \frac{1}{2}L$ and the associated fractional powers

$$U_s f(x) = \sum_{k=0}^{\infty} 2^s \frac{\Gamma\big(\frac{1}{2}(k+n/2+1+s)\big)}{\Gamma\big(\frac{1}{2}(k+n/2+1-s)\big)} Q_k f(x).$$

For these operators, we prove the inequality in the following theorem. Letting A be either L or U, we define the trace norm of a suitable function $u(x,\rho)$ on $\mathbb{R}^n \times [0,\infty)$ as

$$a_s(A,u)^2 = \int_0^\infty \int_{\mathbb{R}^n} \big(|\nabla_A u(x,\rho)|^2 + \big(\tfrac{1}{2}n + \tfrac{1}{4}\rho^2\big) u(x,\rho)^2\big) \rho^{1-2s}\, d\gamma(x)\, d\rho,$$

where

$$\nabla_U u := (2^{-1/2}\partial_1 u, 2^{-1/2}\partial_2 u, \dots, 2^{-1/2}\partial_n u, \partial_\rho u)$$

and ∇_L is defined without the scaling factor $2^{-1/2}$.

Theorem 1.1 (general trace Hardy inequality). *Let $0 < s < 1$, and let A be either L or U. Suppose $\phi \in L^2(\gamma)$ is a real-valued function in the domain of A_s such that $\phi^{-1}A_s\phi$ is locally integrable. Then for any real-valued function $u(x,\rho)$ from the space $C_0^2([0,\infty), C_b^2(\mathbb{R}^n))$ we have*

$$a_s(A,u)^2 \ge 2^{1-2s}\frac{\Gamma(1-s)}{\Gamma(s)}\int_{\mathbb{R}^n} u(x,0)^2 \frac{A_s\phi(x)}{\phi(x)}\,d\gamma(x).$$

It would be nice if we could choose a function ϕ so that $A_s\phi$ can be calculated explicitly. It turns out that for $A = U$ we can do that. Indeed, with such a choice of ϕ we can prove an explicit trace Hardy inequality from which Hardy's inequality can be deduced.

Theorem 1.2 (Hardy's inequality for U_s). *Let $0 < s < 1$. Assume that $f \in L^2(\gamma)$ with $U_s f \in L^2(\gamma)$. Then for every $\rho > 0$ we have*

$$\langle U_s f, f\rangle_{L^2(\gamma)} \ge (2\rho)^s \frac{\Gamma\big(\frac12(n/2+1+s)\big)}{\Gamma\big(\frac12(n/2+1-s)\big)}\int_{\mathbb{R}^n}\frac{f(x)^2}{(\rho+|x|^2)^s}w_s(\rho+|x|^2)\,d\gamma(x)$$

for an explicit $w_s(t) \ge 1$. The inequality is sharp, and equality is attained for

$$f(x) = \sqrt{2}\frac{2^{-(n/2+1-s)/2}}{\Gamma\big(\frac12(n/2+1-s)\big)}e^{|x|^2/2}(\rho+|x|^2)^{-(n/2+1-s)/2}K_{(n/2+1-s)/2}(\rho+|x|^2),$$

where K_μ denotes the Macdonald's function.

We remark that since $w_s(t) \ge 1$, we have the following inequality which is slightly weaker:

$$\langle U_s f, f\rangle_{L^2(\gamma)} \ge (2\rho)^s \frac{\Gamma\big(\frac12(n/2+1+s)\big)}{\Gamma\big(\frac12(n/2+1-s)\big)}\int_{\mathbb{R}^n}\frac{f(x)^2}{(\rho+|x|^2)^s}\,d\gamma(x). \tag{1-4}$$

However, written in this form, we do not yet know if the constant appearing in the above inequality is sharp or not. Observe that the constant we have obtained is analogous to the sharp constant in the Euclidean case; see (1-2). It is worth pointing out that Hardy's inequality for the pure fractional powers U^s can be deduced from Theorem 1.2. Indeed, writing $R_s := U_s U^{-s}$, we see that R_s is a bounded operator on $L^2(\gamma)$ and its operator norm is given by

$$\|R_s\|_{\mathrm{op}} = \sup_{k\ge0}\big(\tfrac12(k+n/2)\big)^{-s}\frac{\Gamma\big(\frac12(k+n/2+1+s)\big)}{\Gamma\big(\frac12(k+n/2+1-s)\big)}.$$

To estimate this norm we use the fact that $x^{\beta-\alpha}\Gamma(x+\alpha)/\Gamma(x+\beta) \le (x+\beta)/(x+\alpha)$ for $\alpha > 0$ (see [Roncal and Thangavelu 2016]), which gives the estimate

$$\big(\tfrac12(k+n/2)\big)^{-s}\frac{\Gamma\big(\frac12(k+n/2+1+s)\big)}{\Gamma\big(\frac12(k+n/2+1-s)\big)} \le \frac{2k+n+2(1-s)}{2k+n+2(1+s)}.$$

The right-hand side of the above inequality being an increasing function of k, we obtain $\|R_s\|_{\mathrm{op}} \le 1$. Using this, Hardy's inequality for U^s reads as follows:

Corollary 1.3. *Let $0 < s < 1$. Assume that $f \in L^2(\gamma)$ with $U^s f \in L^2(\gamma)$. Then for any $\rho > 0$ we have*

$$\langle U^s f, f\rangle_{L^2(\gamma)} \geq (2\rho)^s \frac{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)} \int_{\mathbb{R}^n} \frac{f(x)^2}{(\rho+|x|^2)^s}\, d\gamma(x).$$

As a consequence of Hardy's inequality with nonhomogeneous weight, we obtain a Heisenberg-type uncertainty principle for the fractional powers of the Ornstein–Uhlenbeck operator. Indeed, an application of the Cauchy–Schwarz inequality yields

$$\int_{\mathbb{R}^n} |f(x)|^2\, d\gamma(x) \leq \left(\int_{\mathbb{R}^n} |f(x)|^2(\rho+|x|^2)^s\, d\gamma(x)\right)^{1/2} \left(\int_{\mathbb{R}^n} \frac{f(x)^2}{(\rho+|x|^2)^s}\, d\gamma(x)\right)^{1/2},$$

which along with Theorem 1.2 gives the following:

Corollary 1.4. *For any $f \in L^2(\gamma)$ with $U_s f \in L^2(\gamma)$, we have*

$$\left(\int_{\mathbb{R}^n} |f(x)|^2(\rho+|x|^2)^s\, d\gamma(x)\right)\langle U_s f, f\rangle_{L^2(\gamma)} \geq (2\rho)^s \frac{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)} \left(\int_{\mathbb{R}^n} |f(x)|^2\, d\gamma(x)\right)^2.$$

We must mention that one can use the L^2-boundedness of $U_s L^{-s}$ along with the inequality for U_s to derive an inequality for L^s. Indeed, the operator norm of $\mathfrak{R}_s := U_s L^{-s}$ is given by

$$\|\mathfrak{R}_s\|_{\mathrm{op}} = \sup_{k\geq 0} 2^s(2k+n)^{-s} \frac{\Gamma\big(\frac{1}{2}(k+n/2+1+s)\big)}{\Gamma\big(\frac{1}{2}(k+n/2+1-s)\big)},$$

which can be estimated as above to get $\|\mathfrak{R}_s\|_{\mathrm{op}} \leq 2^{-s}$. The fact that $\|\mathfrak{R}_s\|_{\mathrm{op}}\langle L^s f, f\rangle_{L^2(\gamma)} \geq \langle U_s f, f\rangle_{L^2(\gamma)}$ together with this estimate yields the following:

Theorem 1.5 (Hardy's inequality for L^s). *Let $0 < s < 1$. Assume that $f \in L^2(\gamma)$ with $L^s f \in L^2(\gamma)$. Then for any $\rho > 0$ we have*

$$\langle L^s f, f\rangle_{L^2(\gamma)} \geq (4\rho)^s \frac{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)} \int_{\mathbb{R}^n} \frac{f(x)^2}{(\rho+|x|^2)^s}\, d\gamma(x).$$

The main ingredient in proving the above mentioned trace Hardy and Hardy's inequality for fractional powers of L is a solution of the extension problem for L:

$$\Big(-L+\partial_\rho^2+\frac{1-2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)u(x,\rho)=0, \quad u(x,0)=f(x). \tag{1-5}$$

As we will see later, a solution of the above partial differential equation will play a very crucial role for our purpose. The second theme of this article is the study of general solutions of the extension problem for L under consideration. In fact, we prove a characterisation of the solution when the initial data is a tempered distribution. In order to state the result we need to introduce some more notations which will be briefly described here. More details can be found in Section 3. We introduce the following two operators. For any distribution f for which $M_\gamma f$ is tempered, we define

$$S_\rho^1 f(x) := \frac{\big(\frac{1}{2}\rho^2\big)^{(s-1)/2}}{\Gamma(s)} \sum_{k=0}^{\infty} \Gamma\big(\tfrac{1}{2}(2k+n+s+1)\big) W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big) Q_k f(x),$$

and for any function g for which $Q_k g$ has enough decay as a function of k, we define

$$S_\rho^2 g(x) := \left(\tfrac{1}{2}\rho^2\right)^{(s-1)/2} \sum_{k=0}^{\infty} M_{-(k+n/2),\, s/2}\left(\tfrac{1}{2}\rho^2\right) Q_k g(x),$$

where $W_{-(k+n/2),\, s/2}$ and $M_{-(k+n/2),\, s/2}$ are Whittaker functions.

In view of the asymptotic properties of the Whittaker functions stated in Lemma 3.2, it follows that the series defining $S_\rho^1 f$ converges for any tempered distribution $M_\gamma f$. Moreover, if we take g from $H^2_{\gamma,\rho}(\mathbb{R}^n)$, which is the image of $L^2(\mathbb{R}^n, \gamma)$ under the semigroup $e^{-\rho\sqrt{L}}$, then the series defining $S_\rho^2 g$ also converges and defines a smooth function. With these notations we prove the following characterisation:

Theorem 1.6. *Let f be a distribution such that $M_\gamma f$ is tempered. Then any function $u(x, \rho)$ for which $M_\gamma u(x, \rho)$ is tempered in x is a solution of the extension problem* (1-5) *with initial condition f if and only if $u(x, \rho) = S_\rho^1 f(x) + S_\rho^2 g(x)$ for some $g \in \bigcap_{t>0} H^2_{\gamma,t}(\mathbb{R}^n)$.*

We also prove another characterisation of the solution of the extension problem in terms of its holomorphic extendability. In order to state this we need to introduce some more notations. For any $t, \delta > 0$ we consider the positive weight function

$$w_t^\delta(x, y) = \frac{1}{\Gamma(\delta)} \int_{\mathbb{R}^n} e^{-2ux} \left(1 - \frac{|u|^2 + |y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, du.$$

For any $\rho > 0$, we let $H^2(\Omega_\rho, w_\rho^{2s})$ stand for the weighted Bergman space consisting of holomorphic functions on the tube domain $\Omega_\rho := \{z = x + iy \in \mathbb{C}^n : |y| < \rho\}$ belonging to $L^2(\Omega_\rho, w_\rho^{2s})$. Also for $m \in \mathbb{R}$, let $W_H^m(\mathbb{R}^n)$ stand for the Sobolev space associated to the Hermite operator H. This is a Hilbert space in which the norm is given by

$$\|f\|^2_{W_H^m} := \sum_{k=0}^{\infty} (2k+n)^{2m} \|P_k f\|_2^2.$$

Theorem 1.7. *A solution of the extension problem* (1-5) *is of the form $u(x, \rho) = S_\rho^1 f(x)$ for some distribution f such that $M_\gamma f \in W_H^{m_n}(\mathbb{R}^n)$, where $2m_n = -\frac{1}{4}(2n+1)$, if and only if for every $\rho > 0$, $M_\gamma u(\,\cdot\,, \rho)$ extends holomorphically to $\Omega_{\rho/2}$ and satisfies the uniform estimate*

$$\int_{\Omega_{\rho/2}} |M_\gamma u(z, \rho)|^2 w_{\rho/2}^{2s}(z)\, dz \le C\rho^{n-1/2}$$

for all $0 < \rho \le 1$.

We conclude the introduction by describing the plan of the paper. In Section 2, we study an extension problem for the Ornstein–Uhlenbeck operator. We provide two representations of solutions and their equivalence. In Section 3, we prove several characterisations of the solution of the extension problem under consideration. Using the results obtained in Section 2, we prove the trace Hardy and Hardy's inequality in Section 4. Then in Section 5, we prove an isometry property of the solution to the extension problem. Finally we end our discussion by proving an inequality of Hardy–Littlewood–Sobolev type for the fractional powers of the Hermite operator in Section 6.

2. The extension problem for the Ornstein–Uhlenbeck operator and fractional powers

The extension problem. Our strategy to prove Hardy's inequality for L_s is via the trace Hardy inequality which in turn requires the study of the following extension problem for the operator L:

$$\Big(-L+\partial_\rho^2+\frac{1-2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)u(x,\rho)=0,\quad u(x,0)=f(x). \tag{2-1}$$

If u is a solution of the above problem, it follows that $v(x,\rho)=M_\gamma u(x,\rho)$ solves the problem

$$\Big(-H+\partial_\rho^2+\frac{1-2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)v(x,\rho)=0,\quad v(x,0)=M_\gamma f(x). \tag{2-2}$$

A solution of the above problem can be obtained in terms of the solution of an extension problem for the sub-Laplacian on the Heisenberg group.

Let $\mathcal{L}$ be the sub-Laplacian on the Heisenberg group $\mathbb{H}^n$. Then a solution of the extension problem

$$\Big(-\mathcal{L}+\partial_\rho^2+\frac{1-2s}{\rho}\partial_\rho+\frac{1}{4}\rho^2\partial_t^2\Big)w(z,t,\rho)=0,\quad w(z,t,0)=f(z,t)$$

is given by $w(z,t,\rho)=\rho^{2s}f*\Phi_{s,\rho}(z,t)$, see [Roncal and Thangavelu 2020a], where $\Phi_{s,\rho}$ is an explicit function given by

$$\Phi_{s,\rho}(z,a)=\frac{2^{-(n+1+s)}}{\pi^{n+1}\Gamma(s)}\Gamma\big(\tfrac{1}{2}(n+1+s)\big)^2\big(\big(\tfrac{1}{4}\rho^2+\tfrac{1}{4}|z|^2\big)^2+a^2\big)^{-(n+1+s)/2}.$$

If we let π stand for the Schrödinger representation of $\mathbb{H}^n$ on $L^2(\mathbb{R}^n)$, then we have the following result.

Theorem 2.1. *For any $f\in L^2(\gamma)$ the function $v(x,\rho)$ defined by the equation*

$$v(x,\rho)=\rho^{2s}\int_{\mathbb{H}^n}\Phi_{s,\rho}(g)\pi(g)^*M_\gamma f(x)\,dg$$

solves the extension problem for the Hermite operator with initial condition $M_\gamma f$. Consequently, the extension problem for L is solved by $u(x,\rho)=e^{|x|^2/2}v(x,\rho)$.

Proof. For any X from the Heisenberg Lie algebra $\mathfrak{h}^n$ viewed as a left-invariant vector field on $\mathbb{H}^n$, we can easily check that

$$\pi(X)\int_{\mathbb{H}^n}\varphi(g)\pi(g)^*f(x)\,dg=-\int_{\mathbb{H}^n}X\varphi(g)\pi(g)^*f(x)\,dg.$$

This leads to

$$H\int_{\mathbb{H}^n}\varphi(g)\pi_\lambda(g)^*f(x)\,dg=\int_{\mathbb{H}^n}\mathcal{L}\varphi(g)\pi_\lambda(g)^*f(x)\,dg,$$

and consequently, as

$$\rho^{2s}\mathcal{L}\Phi_{s,\rho}(g)=\Big(\partial_\rho^2+\frac{1-2s}{\rho}\partial_\rho+\frac{1}{4}\rho^2\partial_t^2\Big)\rho^{2s}\Phi_{s,\rho}(g)=0,$$

we obtain

$$\Big(-H+\partial_\rho^2+\frac{1-2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)\Big(\rho^{2s}\int_{\mathbb{H}^n}\Phi_{s,\rho}(g)\pi(g)^*f(x)\,dg\Big)=0.$$

To check that $v(x,\rho)$ satisfies the initial condition, we make the change of variables $(z,t)\to(\rho z,\rho^2 t)$, so that

$$v(x,\rho)=\int_{\mathbb{H}^n}\Phi_{s,1}(z,t)\pi(\rho z,\rho^2 t)^* M_\gamma f(x)\,dz\,dt.$$

Since $\pi(\rho z,\rho^2 t)M_\gamma f$ converges to $M_\gamma f$ in $L^2(\mathbb{R}^n)$, we obtain $v(x,\rho)\to M_\gamma f$ as $\rho\to 0$ in view of the fact that $\int_{\mathbb{H}^n}\Phi_{s,1}(g)\,dg=1$. This completes the proof of the theorem. □

There is yet another convenient way of representing the solution of the extension problem for L. If we let $k_{t,s}(\rho)=(\sinh t)^{-s-1}e^{-(\coth t)\rho^2/4}$, then it is known that this function satisfies the equation

$$\partial_t k_{t,s}(\rho)=\Big(\partial_\rho^2+\frac{1+2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)k_{t,s}(\rho).$$

Theorem 2.2. *For $f\in L^p(\gamma)$ with $1\le p\le\infty$, a solution of the extension problem for L is given by*

$$u(x,\rho)=\frac{4^{-s}}{\Gamma(s)}\rho^{2s}\int_0^\infty k_{t,s}(\rho)e^{-tL}f(x)\,dt. \tag{2-3}$$

Moreover, as $\rho\to 0$, the solution $u(\,\cdot\,,\rho)$ converges to f in $L^p(\gamma)$ for any $1\le p<\infty$.

Proof. That u solves the extension problem follows easily from the fact that $e^{-tL}f(x)$ solves the heat equation associated to L, i.e., $-Le^{-tL}f(x)=\partial_t e^{-tL}f(x)$, and the definition of $k_{t,s}(\rho)$. Indeed, we have

$$-Lu(x,\rho)=\frac{4^{-s}}{\Gamma(s)}\rho^{2s}\int_0^\infty k_{t,s}(\rho)\partial_t v(x,t)\,dt,$$

which after an integration by parts in the t variable yields

$$Lu(x,\rho)=\frac{4^{-s}}{\Gamma(s)}\rho^{2s}\int_0^\infty \partial_t k_{t,s}(\rho)v(x,t)\,dt.$$

Since $k_{t,s}(\rho)$ is the heat kernel associated to the operator $\big(\partial_\rho^2+\frac{1+2s}{\rho}\partial_\rho-\frac14\rho^2\big)$, we have

$$Lu(x,\rho)=\frac{4^{-s}}{\Gamma(s)}\rho^{2s}\Big(\partial_\rho^2+\frac{1+2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)\int_0^\infty k_{t,s}(\rho)e^{-tL}f(x)\,dt.$$

Finally, an easy calculation shows that for any function $v(\rho)$ one has

$$\Big(\partial_\rho^2+\frac{1-2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)(\rho^{2s}v(\rho))=\rho^{2s}\Big(\partial_\rho^2+\frac{1+2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)v(\rho),$$

and hence it follows that $u(x,\rho)$ solves the extension problem.

Now to prove the $L^p(\gamma)$ convergence of the solution to the initial condition, we make use of the fact that e^{-tL} is a contraction semigroup on every $L^p(\gamma)$ and $e^{-tL}f$ converges to f in $L^p(\gamma)$ as $t\to 0$. We first make a change of variables $t\to\rho^2 t$ to get

$$u(x,\rho)=\frac{4^{-s}}{\Gamma(s)}\rho^{2s+2}\int_0^\infty k_{\rho^2 t,s}(\rho)e^{-\rho^2 tL}f(x)\,dt.$$

Note that

$$\begin{aligned}\rho^{2s+2}k_{\rho^2 t,s}(\rho) &= \rho^{2s+2}(\sinh\rho^2 t)^{-s-1}e^{-(\coth\rho^2 t)\rho^2/4}\\ &= t^{-s-1}\left(\frac{\rho^2 t}{\sinh\rho^2 t}\right)^{s+1}e^{-(\coth\rho^2 t)\rho^2 t/(4t)} \quad\to\quad t^{-s-1}e^{-1/(4t)} \qquad \text{as } \rho\to 0. \qquad (2\text{-}4)\end{aligned}$$

Here we have used the facts that $(\sinh y)/y \to 1$ and $y\coth y \to 1$ as $y\to 0$. Also we see that $t^{-s-1}e^{-1/(4t)} \in L^1(0,\infty)$, and an easy calculation yields

$$\int_0^\infty t^{-s-1}e^{-1/(4t)}\,dt = 4^s\Gamma(s).$$

Now using this result we can write, for any $x\in\mathbb{R}^n$,

$$\begin{aligned}u(x,\rho)-f(x) &= \frac{4^{-s}}{\Gamma(s)}\rho^{2s+2}\int_0^\infty k_{\rho^2 t,\,s}(\rho)e^{-\rho^2 tL}f(x)\,dt - \frac{4^{-s}}{\Gamma(s)}\int_0^\infty t^{-s-1}e^{-1/(4t)}f(x)\,dt\\ &= \frac{4^{-s}}{\Gamma(s)}\int_0^\infty (\rho^{2s+2}k_{\rho^2 t,\,s}(\rho)-t^{-s-1}e^{-1/(4t)})e^{-\rho^2 tL}f(x)\,dt\\ &\qquad\qquad + \frac{4^{-s}}{\Gamma(s)}\int_0^\infty t^{-s-1}e^{-1/(4t)}(e^{-\rho^2 tL}f(x)-f(x))\,dt.\end{aligned}$$

Therefore, using Minkowski's integral inequality and the fact that $\|e^{-\rho^2 tL}f\|_{L^p(\gamma)} \le \|f\|_{L^p(\gamma)}$, we have

$$\begin{aligned}\|u(\,\cdot\,,\rho)-f\|_{L^p(\gamma)} &\le \frac{4^{-s}}{\Gamma(s)}\int_0^\infty |\rho^{2s+2}k_{\rho^2 t,\,s}(\rho)-t^{-s-1}e^{-1/(4t)}|\,\|f\|_{L^p(\gamma)}\,dt\\ &\qquad + \frac{4^{-s}}{\Gamma(s)}\int_0^\infty t^{-s-1}e^{-1/(4t)}\|e^{-\rho^2 tL}f-f\|_{L^p(\gamma)}\,dt. \qquad (2\text{-}5)\end{aligned}$$

Note that using the asymptotics of the sine and cotangent hyperbolic functions, we have

$$|\rho^{2s+2}k_{\rho^2 t,\,s}(\rho)-t^{-s-1}e^{-1/(4t)}| \le C\rho^{2s+2}e^{-\rho^2 t(s+1)} + t^{-s-1}e^{-1/(4t)} := h_\rho(t), \quad t>M. \qquad (2\text{-}6)$$

It is not hard to see that for every $\rho>0$, we have $h_\rho\in L^1$ and $\lim_{\rho\to 0}\int_M^\infty h_\rho(t)\,dt = \int_M^\infty h(t)\,dt$, and also as $\rho\to 0$ we have $h_\rho(t)\to t^{-s-1}e^{-1/(4t)} =: h(t)$ pointwise. Hence by the generalised dominated convergence theorem (DCT) we have

$$\int_M^\infty |\rho^{2s+2}k_{\rho^2 t,s}(\rho)-t^{-s-1}e^{-1/(4t)}|\,\|f\|_{L^p(\gamma)}\,dt \to 0 \quad \text{as } \rho\to 0.$$

Now see that, similar to (2-4), one can show the function $h_\rho(t)$ goes to a finite limit as $t\to 0$, so there is no singularity of h_ρ at 0. Hence it is easy to see that

$$\int_0^M |\rho^{2s+2}k_{\rho^2 t,s}(\rho)-t^{-s-1}e^{-1/(4t)}|\,\|f\|_{L^p(\gamma)}\,dt$$

goes to zero as $\rho\to 0$. Hence it follows that the first integral in the right-hand side of (2-5) goes to zero. Also the integrand of the second integral is bounded above by an integrable function of t. Indeed,

$$t^{-s-1}e^{-1/(4t)}\|e^{-\rho^2 tL}f-f\|_{L^p(\gamma)} \le 2t^{-s-1}e^{-1/(4t)}\|f\|_{L^p(\gamma)}.$$

Hence by DCT the second integral goes to zero as $\rho \to 0$. Therefore we have

$$u(\,\cdot\,,\rho) = \frac{4^{-s}}{\Gamma(s)}\rho^{2s+2}\int_0^\infty k_{\rho^2 t,\,s}(\rho)e^{-\rho^2 tL} f\,dt \to f \quad \text{in } L^p(\gamma) \qquad \text{as } \rho \to 0. \qquad \square$$

We have thus given two representations for solutions of the extension problem, and we now claim they are the same. This is not obvious and needs a proof. It is convenient to work with the functions

$$\varphi_{s,\delta}(z,a) = \big((\delta + \tfrac{1}{4}|z|^2)^2 + a^2\big)^{-(n+1+s)/2},$$

in terms of which we can express $\Phi_{s,\rho}$ as follows: with $\delta = \frac{1}{4}\rho^2$,

$$\Phi_{s,\rho}(z,a) = \frac{2^{-(n+1+s)}}{\pi^{n+1}\Gamma(s)}\Gamma\big(\tfrac{1}{2}(n+1+s)\big)^2\varphi_{s,\delta}(z,a). \tag{2-7}$$

For a function $\varphi(z,t)$ on $\mathbb{H}^n$ we let $\varphi^\lambda(z)$ denote the inverse Fourier transform of φ in the t variable. Thus

$$\varphi^\lambda_{s,\delta}(z) = \int_{-\infty}^\infty \varphi_{s,\delta}(z,t)e^{i\lambda t}\,dt.$$

This is a radial function on $\mathbb{C}^n$ and hence has an expansion in terms of the Laguerre functions:

$$\varphi^\lambda_k(z) = L_k^{n-1}\big(\tfrac{1}{2}|\lambda|\,|z|^2\big)e^{-|\lambda|\,|z|^2/4}. \tag{2-8}$$

We let $c^\lambda_{k,\delta}(s)$ be the coefficients defined by

$$\varphi^\lambda_{s,\delta}(z) = (2\pi)^{-n}|\lambda|^n\sum_{k=0}^\infty c^\lambda_{k,\delta}(s)\varphi^\lambda_k(z). \tag{2-9}$$

These coefficients are given in terms of the auxiliary function $L(a,b,c)$ defined for $a, b \in \mathbb{R}_+$ and $c \in \mathbb{R}$ as follows:

$$L(a,b,c) = \int_0^\infty e^{-a(2x+1)}x^{b-1}(1+x)^{-c}\,dx. \tag{2-10}$$

The following proposition expresses the $c^\lambda_{k,\delta}(s)$ in terms of L; see [Cowling and Haagerup 1989].

Proposition 2.3 (Cowling–Haagerup). *For any $\delta > 0$ and $0 < s < \frac{1}{2}(n+1)$, we have*

$$c^\lambda_{k,\delta}(s) = \frac{(2\pi)^{n+1}|\lambda|^s}{\Gamma\big(\frac{1}{2}(n+1+s)\big)^2}L\big(\delta|\lambda|,\, \tfrac{1}{2}(2k+n+1+s),\, \tfrac{1}{2}(2k+n+1-s)\big).$$

Using this proposition we can compute the explicit formula for the group Fourier transform of $\Phi_{s,\rho}(g)$ on $\mathbb{H}^n$. Let $P_k(\lambda)$ stand for the projections associated to $H(\lambda) = -\Delta + \lambda^2|x|^2$. Then making use of the fact that

$$\int_{\mathbb{C}^n}\varphi^\lambda_k(z)\pi_\lambda(z,0)\,dz = (2\pi)^{-n}|\lambda|^{-n}P_k(\lambda),$$

we obtain the following formula: with $\delta = \frac{1}{4}\rho^2$, as before,

$$\int_{\mathbb{H}^n}\Phi_{s,\rho}(g)\pi_\lambda(g)^*\,dg = \frac{2^{-(n+1+s)}}{\pi^{n+1}\Gamma(s)}\Gamma\big(\tfrac{1}{2}(n+1+s)\big)^2\sum_{k=0}^\infty c^\lambda_{k,\delta}(s)P_k(\lambda).$$

As the projections associated to L are given by $Q_k = M_\gamma^{-1} P_k M_\gamma$, we see that the solution defined in Theorem 2.1 is given by

$$u(x,\rho) = \frac{2^{-(n+1+s)}}{\pi^{n+1}\Gamma(s)} \Gamma\big(\tfrac{1}{2}(n+1+s)\big)^2 \rho^{2s} \sum_{k=0}^{\infty} c_{k,\delta}^1(s) Q_k f(x).$$

Therefore, in order to prove our claim, we only need to check that

$$\frac{4^{-s}}{\Gamma(s)}\rho^{2s}\int_0^\infty k_{t,s}(\rho)e^{-tL}f(x)\,dt = \frac{2^{-(n+1+s)}}{\pi^{n+1}\Gamma(s)} \Gamma\big(\tfrac{1}{2}(n+1+s)\big)^2 \rho^{2s} \sum_{k=0}^{\infty} c_{k,\delta}^1(s) Q_k f(x),$$

where $\delta = \frac{1}{4}\rho^2$. Equivalently, we need to check that

$$\int_0^\infty k_{t,s}(\rho)e^{-t(2k+n)}\,dt = L\big(\tfrac{1}{4}\rho^2, \tfrac{1}{2}(2k+n+1+s), \tfrac{1}{2}(2k+n+1-s)\big).$$

In order to compute the above integral, we make the change of variable $\coth t = 2z+1$ and note that $-(\sinh^2 t)^{-1}\,dt = 2\,dz$ and $\sinh t = (2z(2z+2))^{-1/2}$. We get

$$\begin{aligned}
\int_0^\infty &(\sinh t)^{-s-1} e^{-(\coth t)\rho^2/4} e^{-t(2k+n)}\,dt \\
&= 2\int_0^\infty (2z(2z+2))^{(s-1)/2} e^{-(2z+1)\rho^2/4}\Big(\frac{2z+2}{2z}\Big)^{-(2k+n)/2}\,dz \\
&= 2\int_0^\infty e^{-(2z+1)\rho^2/4}(2z)^{[(s-1)+(2k+n)]/2}(2z+2)^{-[(1-s)+(2k+n)]/2}\,dz \\
&= 2^s\int_0^\infty e^{-(2z+1)\rho^2/4}(z)^{[(s+1)+(2k+n)]/2-1}(z+1)^{-[(1-s)+(2k+n)]/2}\,dz \\
&= 2^s L\big(\tfrac{1}{4}\rho^2, \tfrac{1}{2}(2k+n+1+s), \tfrac{1}{2}(2k+n+1-s)\big).
\end{aligned}$$

This proves our claim that Theorems 2.1 and 2.2 define the same solution of the extension problem.

The above proof also shows that the function $u(x,\rho)$ defined by the integral (using U in place of L)

$$u(x,\rho) = \frac{4^{-s}}{\Gamma(s)}\rho^{2s}\int_0^\infty k_{t,s}(\rho)e^{-tU}f(x)\,dt,$$

solves the extension problem for U and the following expansion for the solution u is valid.

Proposition 2.4. *For $0 < s < \frac{1}{2}(n+1)$ and $f \in L^2(\gamma)$, the solution of the extension problem associated to U is given by*

$$u(\cdot,\rho) = \frac{2^{-s}}{\Gamma(s)}\rho^{2s}\sum_{k=0}^{\infty} L\big(\tfrac{1}{4}\rho^2, \tfrac{1}{2}(k+n/2+1+s), \tfrac{1}{2}(k+n/2+1-s)\big)Q_k f. \tag{2-11}$$

We let $T_{s,\rho}$ stand for the solution operator which takes f into the solution $u(x,\rho)$ of the extension problem. Thus

$$T_{s,\rho}f(x) = \frac{4^{-s}}{\Gamma(s)}\rho^{2s}\int_0^\infty k_{t,s}(\rho)e^{-tL}f(x)\,dt,$$

which is also given by the expansion in the above proposition. In what follows we make use of the transformation property

$$\frac{(2\lambda)^a}{\Gamma(a)} L(\lambda, a, b) = \frac{(2\lambda)^b}{\Gamma(b)} L(\lambda, b, a) \tag{2-12}$$

satisfied by the L function for all admissible values of (a, b, c); see Cowling and Haagerup [1989].

Fractional powers of the operators L and U. In what follows let A stand for either L or U. Note that the associated eigenvalues λ_k are given by $(2k+n)$ and $\big(k+\frac{1}{2}n\big)$, respectively. The above representation of the solution of the extension problem allows us to define A_s as the Neumann boundary data associated to the extension problem. More precisely we have the following result:

Theorem 2.5. *Assume that* $0 < s < 1$. *Let* $f \in L^2 \cap L^p(\gamma)$ *with* $1 \le p < \infty$ *be such that* $A_s f \in L^p(\gamma)$. *Then the solution of the extension problem* $u(x, \rho) = T_{s,\rho} f(x)$ *satisfies*

$$\lim_{\rho \to 0} \rho^{1-2s} \partial_\rho u(x, \rho) = -2^{1-2s} \frac{\Gamma(1-s)}{\Gamma(s)} A_s f,$$

where the convergence is understood in the $L^p(\gamma)$ *sense.*

Proof. The expansion of $T_{s,\rho} f$ given in Proposition 2.4 and the transformation property (2-12) of the L function allows us to verify the identity

$$\rho^{2s} T_{-s,\rho}(A_s f)(x) = \frac{4^s \Gamma(s)}{\Gamma(-s)} T_{s,\rho} f(x), \tag{2-13}$$

which when expanded reads as

$$\frac{4^s \Gamma(s)}{\Gamma(-s)} u(x, \rho) = \frac{4^s}{\Gamma(-s)} \int_0^\infty (\sinh t)^{s-1} e^{-(\coth t)\rho^2/4} e^{-tA} A_s f(x)\, dt.$$

Differentiating with respect to ρ and multiplying both sides by $-\rho^{1-2s}$, we get

$$-\rho^{1-2s} \partial_\rho u(x, \rho) = \frac{1}{2\Gamma(s)} \rho^{2(1-s)} \int_0^\infty (\sinh t)^{s-1} (\coth t) e^{-(\coth t)\rho^2/4} e^{-tA} A_s f(x)\, dt.$$

Now we make the change of variable $t \to t\rho^2$ to get

$$\begin{aligned} -\rho^{1-2s} \partial_\rho u(x, \rho) &= \frac{1}{2\Gamma(-s)} \rho^{4-2s} \int_0^\infty (\sinh(t\rho^2))^{s-1} \coth(t\rho^2) e^{-\coth(t\rho^2)\rho^2/4} e^{-t\rho^2 A} A_s f(x)\, dt \\ &= \frac{1}{2\Gamma(-s)} \int_0^\infty t^{s-2} \left(\frac{\sinh(t\rho^2)}{t\rho^2}\right)^{s-1} \coth(t\rho^2)(t\rho^2) e^{-\coth(t\rho^2)\rho^2/4} e^{-t\rho^2 A} A_s f(x)\, dt. \end{aligned}$$

Under the extra assumption that $A_s f \in L^p(\gamma)$ with $1 \le p < \infty$, we know that $\lim_{\rho \to 0} e^{-\rho^2 t A} A_s f = A_s f$, in $L^p(\gamma)$. So as $\rho \to 0$, we can argue as in the proof of Theorem 2.2 to obtain

$$\lim_{\rho \to 0} (-\rho^{1-2s} \partial_\rho u(x, \rho)) = \frac{1}{2\Gamma(s)} A_s f \left(\int_0^\infty t^{s-2} e^{-1/(4t)}\, dt \right).$$

Computing the last integral and simplifying we obtain

$$\lim_{\rho \to 0} (\rho^{1-2s} \partial_\rho u(x, \rho)) = -2^{(1-2s)} \frac{\Gamma(1-s)}{\Gamma(s)} A_s f. \qquad \square$$

3. Characterisations of solutions of the extension problem

In this section we prove several characterisations of solutions of the extension problem for L. Recall that the extension problem for L reads as

$$\Big(-L+\partial_\rho^2+\frac{1-2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)u(x,\rho)=0,\quad u(x,0)=f(x).$$

Now given $\alpha\in\mathbb{N}^n$ and $\rho>0$ we define the Fourier–Hermite coefficients associated to the expansion in terms of the normalised Hermite polynomials H_α as

$$\tilde{u}(\alpha,\rho):=\int_{\mathbb{R}^n}u(x,\rho)H_\alpha(x)\,d\gamma(x).$$

Now letting $v_\alpha(\rho):=\tilde{u}(\alpha,2\sqrt{\rho})$, we see that

$$(-(2|\alpha|+n)+\rho\partial_\rho^2+(1-s)\partial_\rho-\rho)v_\alpha(\rho)=0,\quad v_\alpha(0)=(f,H_\alpha)_{L^2(\gamma)}.$$

Again if we write $v_\alpha(\rho)=e^{-\rho}g_\alpha(2\rho)$, then it can be easily checked that the above equation becomes

$$rg_\alpha''(r)+(1-s-r)g_\alpha'(r)-\tfrac{1}{2}(2|\alpha|+n+1-s)g_\alpha(r)=0,$$

where $r=2\rho$. Now we let $g_\alpha(r)=r^sh_\alpha(r)$, which leads to

$$rh_\alpha''(r)+(1+s-r)h_\alpha'(r)-\tfrac{1}{2}(2|\alpha|+n+1+s)h_\alpha(r)=0. \tag{3-1}$$

Note that this is in the form of Kummer's equation: $xh''(x)+(b-x)h'-ah(x)=0$. The solutions of Kummer's equation are given by the functions $M(a,b,x)$ and $V(a,b,x)$, which are known as the confluent hypergeometric functions. The function M, given by $M(a,b,x)=\sum_{m=0}^\infty((a)_m/(b)_m m!)x^m$, is analytic, and

$$V(a,b,x)=\frac{\pi}{\sin\pi b}\bigg(\frac{M(a,b,x)}{\Gamma(1+a-b)\Gamma(b)}-x^{1-b}\frac{M(1+a-b,2-b,x)}{\Gamma(a)\Gamma(2-b)}\bigg),\quad x>0.$$

Also, V has the integral representation given by

$$V(a,b,x)=\frac{1}{\Gamma(a)}\int_0^\infty e^{-tx}t^{a-1}(1+t)^{b-a-1}\,dt,\quad x>0.$$

For more details, see for instance [Abramowitz and Stegun 1964, Chapter 13] and also [Frank et al. 2015, Lemma 5.2].

Finally, writing $\mu=\frac{1}{2}s$ and $\kappa=|\alpha|+\frac{1}{2}n$, performing another substitution $w_\alpha(r)=e^{-1/2r}r^{1/2+\mu}h_\alpha(r)$, transforms (3-1) to

$$w_\alpha''(r)+\bigg(-\frac{1}{4}-\frac{\kappa}{r}+\frac{1/4-\mu^2}{r^2}\bigg)w_\alpha(r)=0, \tag{3-2}$$

which is in the form of a Whittaker equation. This warrants the following lemma which describes the properties of solutions of Whittaker equations.

Lemma 3.1 [Olver and Maximon 2010]. *Let $\kappa \in \mathbb{R}$ and $-2\mu \notin \mathbb{N}$. The two linearly independent solutions of the ordinary differential equation*

$$w''(x) + \left(-\frac{1}{4} + \frac{\kappa}{x} + \frac{1/4 - \mu^2}{x^2}\right) w(x) = 0$$

are given by the functions $M_{\kappa,\mu}(x)$ and $W_{\kappa,\mu}(x)$, where

$$M_{\kappa,\mu}(x) = e^{-x/2} x^{1/2+\mu} \sum_{p=0}^{\infty} \frac{1/2 + \mu - \kappa}{(1+2\mu)_p \, p!} x^p,$$

and when 2μ is not an integer,

$$W_{\kappa,\mu}(x) = \frac{\Gamma(-2\mu)}{\Gamma(1/2 - \mu - \kappa)} M_{\kappa,\mu}(x) + \frac{\Gamma(+2\mu)}{\Gamma(1/2 + \mu - \kappa)} M_{\kappa,-\mu}(x). \tag{3-3}$$

Moreover, we have the following asymptotic properties of these Whittaker functions:

For large x,

$$M_{\kappa,\mu}(x) \sim \frac{\Gamma(1+2\mu)}{\Gamma(1/2+\mu-\kappa)} e^{x/2} x^{-\kappa}, \quad \mu - \kappa \neq -\tfrac{1}{2}, -\tfrac{3}{2}, \dots \quad \textit{and} \quad W_{\kappa,\mu}(x) \sim e^{-x/2} x^{\kappa}. \tag{3-4}$$

Also as $x \to 0$ we have

$$M_{\kappa,\mu}(x) = x^{\mu+1/2}(1 + O(x)), \quad 2\nu \neq -1, -2, -3, \dots, \tag{3-5}$$

$$W_{\kappa,\mu}(x) = \frac{\Gamma(2\mu)}{\Gamma(1/2+\mu-\kappa)} x^{1/2-\mu} + \frac{\Gamma(-2\mu)}{\Gamma(1/2-\mu-\kappa)} x^{1/2+\mu} + O(x^{3/2-\mu}), \quad 0 < \mu < \tfrac{1}{2}. \tag{3-6}$$

In view of the above lemma, generic solutions of (3-2) are given by

$$w_\alpha(r) = C_1(|\alpha|) M_{-(|\alpha|+n/2),\, s/2}(r) + C_2(|\alpha|) W_{-(|\alpha|+n/2),\, s/2}(r).$$

But we know $v_\alpha(\rho) = e^{-\rho} g_\alpha(2\rho) = e^{-\rho}(2\rho)^s h_\alpha(\rho) = e^{-\rho}(2\rho)^s e^{\rho/2} \rho^{-1/2-\mu} w_\alpha(\rho)$, and by definition $v_\alpha(\rho) = \tilde{u}(\alpha, 2\sqrt{\rho})$. Hence we have

$$\tilde{u}(\alpha, \rho) = \left(\tfrac{1}{2}\rho^2\right)^{(s-1)/2} \left(C_1(|\alpha|) W_{-(|\alpha|+n/2),\, s/2}\left(\tfrac{1}{2}\rho^2\right) + C_2(|\alpha|) M_{-(|\alpha|+n/2),\, s/2}\left(\tfrac{1}{2}\rho^2\right)\right). \tag{3-7}$$

The initial condition on the solution along with the behaviour of the Whittaker functions stated in the previous lemma allows us to conclude that

$$C_1(|\alpha|) = \frac{1}{\Gamma(s)} \Gamma\left(\tfrac{1}{2}(2k+n+s+1)\right) (f, H_\alpha)_{L^2(\gamma)}.$$

Thus the solution of the extension problem can be written as a sum of two functions, namely

$$\left(\tfrac{1}{2}\rho^2\right)^{(s-1)/2} \frac{1}{\Gamma(s)} \sum_{k=0}^{\infty} \Gamma\left(\tfrac{1}{2}(2k+n+s+1)\right) W_{-(k+n/2),\, s/2}\left(\tfrac{1}{2}\rho^2\right) Q_k f,$$

$$\left(\tfrac{1}{2}\rho^2\right)^{(s-1)/2} \sum_{\alpha \in \mathbb{N}^n} C_2(|\alpha|) M_{-(|\alpha|+n/2),\, s/2}\left(\tfrac{1}{2}\rho^2\right) H_\alpha(x).$$

The second series above converges under some decay conditions on the coefficients $C_2(|\alpha|)$ as we will see soon. We make use of these considerations in the proof of Theorem 3.3 below.

To proceed further with our description of solutions of the extension problem, we need the following asymptotic properties of the Whittaker functions appearing in the above expressions for large values of the parameter k.

Lemma 3.2. *For any $\rho \in (0, \infty)$, we have the following asymptotic properties, as k tends to infinity*:

$$\left(\tfrac{1}{2}\rho^2\right)^{(s-1)/2} M_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right) \sim (\rho)^{s-1/2}(\sqrt{2k+n})^{-s-1/2} \exp\left(2(2k+n)\zeta\left(\frac{\rho^2}{4(2k+n)}\right)^{1/2}\right), \quad (3\text{-}8)$$

$$\left(\tfrac{1}{2}\rho^2\right)^{(s-1)/2} W_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right) \sim \frac{(\rho\sqrt{2k+n})^{s-1/2}}{\Gamma\left(\frac{1}{2}(2k+n+1+s)\right)} \exp\left(-2(2k+n)\zeta\left(\frac{\rho^2}{4(2k+n)}\right)^{1/2}\right), \quad (3\text{-}9)$$

where $2\sqrt{\zeta(x)} = \sqrt{x+x^2} + \ln(\sqrt{x}+\sqrt{x+1})$ *for* $x > 0$.

Proof. For large values of κ and for any $x \in (0, \infty)$, the following asymptotic properties can be found in [Olver and Maximon 2010, 13.21.6, 13.21.7]:

$$M_{-\kappa,\mu}(4\kappa x) = \frac{2\Gamma(2\mu+1)}{\kappa^{\mu-1/2}}\left(\frac{x\zeta(x)}{1+x}\right)^{1/4} I_{2\mu}(4\kappa\zeta(x)^{1/2})(1+O(\kappa^{-1})), \quad (3\text{-}10)$$

$$W_{-\kappa,\mu}(4\kappa x) = \frac{\sqrt{8/\pi}\,e^{\kappa}}{\kappa^{\kappa-1/2}}\left(\frac{x\zeta(x)}{1+x}\right)^{1/4} K_{2\mu}(4\kappa\zeta(x)^{1/2})(1+O(\kappa^{-1})), \quad (3\text{-}11)$$

where $I_{2\mu}$ is the modified Bessel function of the first kind and $K_{2\mu}$ denotes the Macdonald function of order 2μ. Taking $x = \frac{1}{2}\rho^2$, $\kappa = k + \frac{1}{2}n$ and $\mu = \frac{1}{2}s$, for large values of k, from (3-10) we have

$$M_{-(k+n/2),\,s/2}(x) = \frac{2\Gamma(2\mu+1)}{(k+n/2)^{\mu-1/2}}\left(\frac{x\zeta(x/2(2k+n))}{2(2k+n)+x}\right)^{1/4} I_{2\mu}\left(2(2k+n)\zeta\left(\frac{x}{2(2k+n)}\right)^{1/2}\right)(1+O(k^{-1})).$$

Recall that the modified Bessel function of the first kind has the following asymptotic property:

$$I_{2\mu}(x) \sim \frac{1}{\sqrt{2\pi x}}e^x \quad \text{when } x \text{ is real and } x \to \infty. \quad (3\text{-}12)$$

But it is easy to see that $2(2k+n)\zeta(x/(2(2k+n)))^{1/2}$ goes to infinity as $k \to \infty$, which by the above asymptotic property yields

$$I_{2\mu}\left(2(2k+n)\zeta\left(\frac{x}{2(2k+n)}\right)^{1/2}\right) \sim \left(2(2k+n)\zeta\left(\frac{x}{2(2k+n)}\right)^{1/2}\right)^{-1/2} \exp\left(2(2k+n)\zeta\left(\frac{x}{2(2k+n)}\right)^{1/2}\right), \quad (3\text{-}13)$$

valid for large values of k. It can be easily checked that for any $x > 0$ and large k,

$$\left(\frac{1}{4}\right)^{1/4}\left(\frac{x}{2k+n}\right)^{1/4} \le \left(\frac{x}{2(2k+n)+x}\right)^{1/4} \le \left(\frac{3}{4}\right)^{1/4}\left(\frac{x}{2k+n}\right)^{1/4}. \quad (3\text{-}14)$$

This, along with (3-13), proves the result for the function $M_{-(k+n/2),\,s/2}$.

We obtain the asymptotic property for the other function similarly: for large k, from (3-11) we have

$$W_{-(k+n/2),\,s/2}(x) = \frac{\sqrt{8/\pi}\,e^{k+n/2}}{(k+n/2)^{k+n/2-1/2}}\left(\frac{x\zeta(x/(2(2k+n)))}{2(2k+n)+x}\right)^{1/4} K_{2\mu}\left(2(2k+n)\zeta\left(\frac{x}{2(2k+n)}\right)^{1/2}\right)(1+O(k^{-1})).$$

Now the Macdonald's function $K_{2\mu}(z)$ has the following asymptotic property:

$$K_{2\mu}(x) \sim \sqrt{\pi/(2x)}\,e^{-x}, \quad \text{when } x \text{ is real and } x \to \infty. \tag{3-15}$$

Again for the same reason as above, as $k \to \infty$, using (3-15) we have

$$K_{2\mu}\left(2(2k+n)\zeta\left(\frac{x}{2(2k+n)}\right)^{1/2}\right) \sim \left(2(2k+n)\zeta\left(\frac{x}{2(2k+n)}\right)^{1/2}\right)^{-1/2} \exp\left(-2(2k+n)\zeta\left(\frac{x}{2(2k+n)}\right)^{1/2}\right). \tag{3-16}$$

Using Stirling's formula, $\Gamma(x) = \sqrt{2\pi}\,x^{x-1/2}e^{-x}e^{\theta(x)/12x}$ for $0 < \theta(x) < 1$ which is true for $x > 0$, see [Ahlfors 1953], we have

$$\frac{\Gamma\left(\frac{1}{2}(2k+n+1+s)\right)e^{(k+n/2)}}{(k+n/2)^{(k+n/2)-1/2}} = \frac{\Gamma\left(\frac{1}{2}(2k+n+1+s)\right)}{e^{-\theta(k+n/2)/6(2k+n)}\,\Gamma\left(\frac{1}{2}(2k+n)\right)} \sim \left(\tfrac{1}{2}(2k+n)\right)^{(1+s)/2},$$

as $k \to \infty$. This observation along with (3-14) and the asymptotic property (3-16) yields

$$\Gamma\left(\tfrac{1}{2}(2k+n+1+s)\right)W_{-(k+n/2),\,s/2}(x) \sim (2k+n)^{s/2-1/4}x^{1/4}\exp\left(-2(2k+n)\zeta\left(\frac{x}{2(2k+n)}\right)^{1/2}\right). \qquad \square$$

Remark. It can be easily checked that for large κ the following inequality is valid for any $x > 0$:

$$\tfrac{1}{2}\sqrt{x\kappa} \le \kappa\sqrt{\zeta(x/\kappa)} \le \tfrac{3}{2}\sqrt{x\kappa}, \tag{3-17}$$

which can be used to further simplify the exponential part in the above estimates.

The analysis preceding Lemma 3.2 motivates us to define the following two operators. Given a distribution f such that $M_\gamma f$ is a tempered distribution, we define

$$S^1_\rho f = \frac{\left(\frac{1}{2}\rho^2\right)^{(s-1)/2}}{\Gamma(s)} \sum_{k=0}^{\infty} \Gamma\left(\tfrac{1}{2}(2k+n+s+1)\right)W_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right)Q_k f. \tag{3-18}$$

Recall that h is a tempered distribution on $\mathbb{R}^n$ if and only if the Hermite coefficients satisfy the estimate $|(h, \Phi_\alpha)| \le C(2|\alpha|+n)^m$ for some integer m. So $M_\gamma f$ being a tempered distribution, its Hermite coefficients have at most polynomial growth, and consequently $Q_k f$ has polynomial growth in k. So because of the exponential decay in (3-9), the above series defining $S^1_\rho f$ converges uniformly. Consequently, in view of (3-7), $S^1_\rho f$ defines a solution of the extension problem.

For the other solution of the Whittaker equation we define the operator S^2_ρ for nice functions g by

$$S^2_\rho g = \left(\tfrac{1}{2}\rho^2\right)^{(s-1)/2} \sum_{k=0}^{\infty} M_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right)Q_k g. \tag{3-19}$$

It is not hard to see that as the Whittaker function $M_{-(k+n/2),\,s/2}\big(\frac{1}{2}\rho^2\big)$ has exponential growth as $k\to\infty$, Q_kg must have enough decay for the series in (3-19) to converge. This encourages us to determine a condition on the function g so that the projections Q_kg have enough decay. Now as can be seen in the above lemma, the function $M_{-(k+n/2),\,s/2}\big(\frac{1}{2}\rho^2\big)$ is growing like $e^{c\rho\sqrt{2k+n}}$ for large values of k which leads us to consider the image of $L^2(\gamma)$ under the semigroup $e^{-tL^{1/2}}$, which we denote by $H^2_{\gamma,t}(\mathbb{R}^n)$. Clearly, if $g\in\bigcap_{t>0}H^2_{\gamma,t}(\mathbb{R}^n)$, the series in (3-19) converges and defines a smooth function. But in view of the connection between L and the Hermite operator H, we note that a function g is in $H^2_{\gamma,t}(\mathbb{R}^n)$ if and only if $ge^{-|\cdot|^2/2}$ is in the image of $L^2(\mathbb{R}^n)$ under the Poisson semigroup $e^{-tH^{1/2}}$. Let us write $H^2_t(\mathbb{R}^n):=e^{-tH^{1/2}}(L^2(\mathbb{R}^n))$. We are ready to prove the following characterisation for the solution of the extension problem.

Theorem 3.3. *Let f be a distribution such that $M_\gamma f$ is tempered. Then any function $u(x,\rho)$ for which $M_\gamma u(x,\rho)$ is tempered in x is a solution of the extension problem* (1-5) *with initial condition f if and only if $u(x,\rho)=S^1_\rho f(x)+S^2_\rho g(x)$ for some $g\in\bigcap_{t>0}H^2_{\gamma,t}(\mathbb{R}^n)$.*

Proof. First suppose $u(x,\rho)=S^1_\rho f(x)+S^2_\rho g(x)$ for some g such that $g\in\bigcap_{t>0}H^2_{\gamma,t}(\mathbb{R}^n)$. Consequently, for every $t>0$, we have $\|Q_kg\|^2_{L^2(\gamma)}\le Ce^{-2t\sqrt{2k+n}}$ for large k. So the expression (3-19) defining $S^2_\rho g$ is well defined and solves the extension problem.

Now since $M_\gamma f$ is a tempered distribution, as mentioned above, the Fourier–Hermite coefficients associated to Hermite polynomials of f satisfy

$$|\tilde{f}(\alpha)|=|(f,H_\alpha)_{L^2(\gamma)}|\le C(2|\alpha|+n)^m\quad\text{for some integer } m.$$

But in view of the fact that $\sum_{|\alpha|=k}1=(k+n-1)!/(k!(n-1)!)\le C(2k+n)^{n-1}$, we must have that $\|Q_kf\|^2_{L^2(\gamma)}\le C(2k+n)^{2m+n-1}$. Now the asymptotic property (3-8) in Lemma 3.2 along with estimate (3-17) gives

$$\big(\tfrac{1}{2}\rho^2\big)^{(s-1)/2}\Gamma\big(\tfrac{1}{2}(2k+n+1+s)\big)W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)\le(\rho\sqrt{2k+n})^{s-1/2}e^{-\rho\sqrt{2k+n}/2},$$

which allows us to conclude that

$$\sum_{k=0}^{\infty}\Big(\Gamma\big(\tfrac{1}{2}(2k+n+1+s)\big)W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)\Big)^2(2k+n)^{2m+n-1}<\infty.$$

Consequently, $S^1_\rho f$ make sense and hence solves the extension problem. Now we observe that an easy calculation yields

$$\begin{aligned}\frac{\big(\frac{1}{2}\rho^2\big)^{(s-1)/2}}{\Gamma(s)}&\Gamma\big(\tfrac{1}{2}(2k+n+1+s)\big)W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)\\&=\frac{2^{-s}}{\Gamma(s)}\rho^{2s}L\big(\tfrac{1}{4}\rho^2,\tfrac{1}{2}(2|\alpha|+n+1+s),\tfrac{1}{2}(2|\alpha|+n+1-s)\big),\end{aligned}\tag{3-20}$$

which together with the expression (3-18) yields that $S^1_\rho f$ is in the form (2-3) and, as discussed in the previous subsection, this converges to f as $\rho\to0$. Also note that from the asymptotic property in (3-5), we have $\big(\frac{1}{2}\rho^2\big)^{(s-1)/2}M_{-(k+n/2),\,s/2}\big(\frac{1}{2}\rho^2\big)$ approaches zero as $\rho\to0$. So $S^2_\rho g\to0$ as $\rho\to0$. Therefore $u=S^1_\rho f+S^2_\rho g$ solves the extension problem with initial condition f.

Conversely, suppose $u(x,\rho)$ is a solution of the extension equation (2-1) with initial condition f whose Fourier–Hermite coefficients associated to the Hermite polynomials have tempered growth. Then as discussed in the beginning of this subsection we have

$$\tilde{u}(\alpha,\rho)=\big(\tfrac{1}{2}\rho^2\big)^{(s-1)/2}\big(C_1(|\alpha|)W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)+C_2(|\alpha|)M_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)\big).$$

Now using $\tilde{u}(\alpha,0)=(f,H_\alpha)$ and the behaviour of $\big(\tfrac{1}{2}\rho^2\big)^{(s-1)/2}W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)$ near $\rho=0$, see (3-5), we have

$$C_1(|\alpha|)=\frac{\Gamma\big(\tfrac{1}{2}(2|\alpha|+n+1+s)\big)}{\Gamma(s)}(f,H_\alpha)_{L^2(\gamma)}.$$

Also since $M_\gamma u(x,\rho)$ is tempered, $\tilde{u}(\alpha,\rho)$ has at most polynomial growth in $|\alpha|$. But estimate (3-17) along with the asymptotic property (3-8) yields

$$\big(\tfrac{1}{2}\rho^2\big)^{(s-1)/2}M_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)\le C(\rho)^{s-1/2}(\sqrt{2k+n})^{-s-1/2}e^{3\rho\sqrt{2k+n}/2}$$

for large k. Hence we must have $C_2(|\alpha|)$ decaying as $e^{-3\rho\sqrt{2|\alpha|+n}/2}$ for every $\rho>0$. So let us take $g=\sum_{\alpha\in\mathbb{N}^n}C_2(|\alpha|)H_\alpha$. Then the function g satisfies $\|Q_kg\|^2_{L^2(\gamma)}\le Ce^{-3\rho\sqrt{2k+n}/2}$ for every $\rho>0$. This ensures that $g\in H^2_{\gamma,3\rho/2}(\mathbb{R}^n)$ for every $\rho>0$, which completes the proof. □

Remark. For any $\rho>0$, the space $H^2_\rho(\mathbb{R}^n)$ has an interesting characterisation. It is well known that any g from this space has a holomorphic extension to the tube domain $\Omega_\rho=\{z=x+iy\in\mathbb{C}^n:|y|<\rho\}$ in $\mathbb{C}^n$ which belongs to $L^2(\Omega_\rho,w_\rho)$ for an explicit positive weight function w_ρ given by

$$w_\rho(z)=(\rho^2-|y|^2)^{n/2}\frac{J_{n/2-1}(2i(\rho^2-|x|^2)^{1/2}|x|)}{(2i(\rho^2-|x|^2)|x|)^{n/2-1}},\quad z=x+iy\in\mathbb{C}^n,$$

where $J_{n/2-1}$ denotes the Bessel function of order $(n/2-1)$. We denote this weighted Bergman space by $H^2_\rho(\mathbb{C}^n)$. Thangavelu [2010] proved that for any holomorphic function F on Ω_ρ,

$$\int_{\Omega_\rho}|F(z)|^2w_\rho(z)\,dz=c_n\sum_{k=0}^{\infty}\|P_kf\|^2\frac{k!(n-1)!}{(k+n-1)!}L_k^{n-1}(-2\rho^2)e^{\rho^2},\tag{3-21}$$

where f is the restriction of F to $\mathbb{R}^n$. In view of this identity we see that $g\in H^2_\rho(\mathbb{R}^n)$ if and only if the function $M_\gamma g$ extends holomorphically to Ω_ρ and belongs to $H^2_\rho(\mathbb{C}^n)$. We refer the reader to [Thangavelu 2010] for more details in this regard. From this observation we infer that the condition $g\in\bigcap_{t>0}H^2_{\gamma,t}(\mathbb{R}^n)$ in the above theorem can be replaced by the requirement that $M_\gamma g$ extends holomorphically and belongs to $\bigcap_{t>0}H^2_t(\mathbb{C}^n)$.

We also have the following characterisation of the solution $u(x,\rho)$ when $M_\gamma u(x,\rho)$ has tempered growth in both the variables.

Theorem 3.4. *Suppose $u(x,\rho)$ is a solution of the extension problem* (2-1), *where $M_\gamma u$ is tempered* (*in both variables*). *Then $u=S^1_\rho f$ for some $f\in L^p(\gamma)$ if and only if* $\sup_{\rho>0}\|u(\cdot,\rho)\|_{L^p(\gamma)}\le C$.

Proof. Suppose $f \in L^p(\gamma)$, and let $u = S^1_\rho f$. Then, as mentioned earlier,

$$u(x,\rho) = \frac{4^{-s}}{\Gamma(s)}\rho^{2s}\int_0^\infty k_{t,s}(\rho)e^{-tL}f(x)\,dt.$$

Now since e^{-tL} is a contraction semigroup on $L^p(\gamma)$, we have

$$\|u(\,\cdot\,,\rho)\|_{L^p(\gamma)} \le \|f\|_{L^p(\gamma)}\frac{4^{-s}}{\Gamma(s)}\rho^{2s}\int_0^\infty k_{t,s}(\rho)\,dt.$$

Proceeding in a similar way as before, one can easily see that

$$\int_0^\infty k_{t,s}(\rho)\,dt = 2^{s+1}L\big(\tfrac{1}{4}\rho^2, \tfrac{1}{2}(1+s), \tfrac{1}{2}(1-s)\big).$$

So we have

$$\|u(\,\cdot\,,\rho)\|_{L^p(\gamma)} \le C_s\|f\|_{L^p(\gamma)}\rho^{2s}L\big(\tfrac{1}{4}\rho^2, \tfrac{1}{2}(1+s), \tfrac{1}{2}(1-s)\big).$$

Now we make use of an estimate for the L function, see [Roncal and Thangavelu 2020b, p. 18], to get

$$\|u(\,\cdot\,,\rho)\|_{L^p(\gamma)} \le C_s\|f\|_{L^p(\gamma)}\rho^{2s}\Gamma(s)\big(\tfrac{1}{2}\rho^2\big)^{-s}e^{-\rho^2/4} = 2\|f\|_{L^p(\gamma)}e^{-\rho^2/4},$$

which gives the required boundedness.

Conversely, let $\sup_{\rho>0}\|u(\,\cdot\,,\rho)\|_{L^p(\gamma)} \le C$. This condition allows us to extract a subsequence ρ_j along which $u(\,\cdot\,,\rho)$ converges weakly to a function $f \in L^p(\gamma)$. Letting ρ go to zero along ρ_j, from (3-7) we have

$$\tilde{u}(\alpha,\rho)\big(\tfrac{1}{2}\rho^2\big)^{(s-1)/2} = \frac{\Gamma\big(\frac{1}{2}(2|\alpha|+n+1+s)\big)}{\Gamma(s)}(f, H_\alpha)_{L^2(\gamma)}W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big) + C_2(|\alpha|)M_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big).$$

Now as $\rho \to \infty$ we have

$$M_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big) \sim \frac{\Gamma(1+2\mu)}{\Gamma(1/2+\mu+(k+n/2))}e^{\rho^2/4}\rho^{(2k+n)}.$$

But it is given that $\tilde{u}(\alpha,\rho)$ has polynomial growth in the ρ variable, so we must have $C_2(\alpha) = 0$, and hence we are done. □

Now we turn our attention to the holomorphic extendability of solutions of the extension problem under consideration. To motivate what we plan to do, we first recall a result about holomorphic extendability of solutions of the following extension problem for the Laplacian on $\mathbb{R}^n$:

$$\Big(\Delta + \partial_\rho^2 + \frac{1-s}{\rho}\partial_\rho\Big)u(x,\rho) = 0, \quad u(x,0) = f(x), \quad x \in \mathbb{R}^n, \quad \rho > 0.$$

After the remarkable work of Caffarelli and Silvestre [2007], this problem has been extensively studied in the literature. See, for example, the work of Stinga and Torrea [2010]. It is known that for $f \in L^2(\mathbb{R}^n)$, the function $u(x,\rho) = \rho^s f * \varphi_{s,\rho}(x)$, where $\varphi_{s,\rho}$ is the generalised Poisson kernel given by

$$\varphi_{s,\rho}(x) = \pi^{-n/2}\frac{\Gamma\big(\frac{1}{2}(n+s)\big)}{|\Gamma(s)|}(\rho^2+|x|^2)^{-(n+s)/2}, \quad x \in \mathbb{R}^n,$$

is a solution of the extension problem. Recently in [Roncal and Thangavelu 2020b], the authors proved that a necessary and sufficient condition for the solution of the above problem to be of the form $u(x,\rho) = \rho^s f * \varphi_{s,\rho}(x)$ for some $f \in L^2(\mathbb{R}^n)$ is that $u(\cdot,\rho)$ extends holomorphically to the tube domain Ω_ρ in $\mathbb{C}^n$, belongs to a weighted Bergman space $B_s(\Omega_\rho)$ and satisfies the uniform estimate $\|u(\cdot,\rho)\|_{B_s} \le C$ for all $\rho > 0$, where the norm $\|\cdot\|_{B_s}$ is given by

$$\|F\|_{B_s}^2 := \rho^{-n} \int_{\Omega_\rho} |F(x+iy)|^2 \left(1 - \frac{|y|^2}{\rho^2}\right)_+^{s-1} dx\, dy.$$

Our aim in the rest of this section is to prove an analogous result for the extension problem we considered for the Ornstein–Uhlenbeck operator L. In order to do so, we require the following Gutzmer's formula for the Hermite expansions. In order to state the same, we need to introduce some more notations.

Let $\mathrm{Sp}(n,\mathbb{R})$ denote the symplectic group consisting of $2n \times 2n$ real matrices which preserves the symplectic form $[(x,u),(y,v)] = (u \cdot y - v \cdot x)$ on $\mathbb{R}^{2n}$ with determinant 1. Recall that $O(2n,\mathbb{R})$ stands for the orthogonal group, and let $K := \mathrm{Sp}(n,\mathbb{R}) \cap O(2n,\mathbb{R})$. For a complex matrix $\sigma = a + ib$, it is known that σ is unitary if and only if the matrix $\sigma_A := \left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ belongs to the group K which yields a one to one correspondence between K and the unitary group $U(n)$. A proof of this can be found in [Folland 1989]. We let $\sigma \cdot (x,u)$ stand for the action of σ_A on (x,u), which clearly has a natural extension to $\mathbb{C}^n \times \mathbb{C}^n$. Also given $(x,u) \in \mathbb{R}^n \times \mathbb{R}^n$, let $\pi(x,u)$ be the unitary operator acting on $L^2(\mathbb{R}^n)$ defined by

$$\pi(x,u)\phi(\xi) = e^{i(x\cdot\xi + x\cdot u/2)}\phi(\xi + u), \quad \xi \in \mathbb{R}^n.$$

Clearly for $(z,w) \in \mathbb{C}^n \times C^n$, as long as ϕ is holomorphic, $\pi(z,w)\phi(\xi)$ makes perfect sense. Also note that Laguerre functions of type $(n-1)$, defined earlier in (2-8), can be considered as a function on $\mathbb{R}^n \times \mathbb{R}^n$ which can be holomorphically extended to $\mathbb{C}^n \times \mathbb{C}^n$ as follows:

$$\varphi_k(z,w) := L_k^{n-1}\left(\tfrac{1}{2}(z^2 + w^2)\right) e^{-(z^2+w^2)/4}, \quad z, w \in \mathbb{C}^n.$$

We have the following very useful identity proved in [Thangavelu 2008]:

Theorem 3.5 (Gutzmer's formula). *For a holomorphic function f on $\mathbb{C}^n$, we have*

$$\int_{\mathbb{R}^n}\int_K |\pi(\sigma\cdot(z,w))f(\xi)|^2\, d\sigma\, d\xi = e^{(u\cdot y - v\cdot x)} \sum_{k=0}^{\infty} \frac{k!(n-1)!}{(k+n-1)!} \varphi_k(2iy, 2iv) \|P_k f\|_2^2,$$

where $z = x + iy,\ w = u + iv \in \mathbb{C}^n$.

We use this to prove the following result:

Proposition 3.6. *Let $\delta > 0$. For a holomorphic function F on Ω_t, we have the identity*

$$\int_{\mathbb{R}^n}\int_{|y|<t} |F(x+iy)|^2 w_t^\delta(x,y)\, dx\, dy = C_n \sum_{k=0}^{\infty} \|P_k f\|_2^2 \frac{\Gamma(k+1)\Gamma(n+\delta)}{\Gamma(k+n+\delta)} L_k^{n+\delta-1}(-2t^2) t^{2n},$$

where f denotes the restriction of F to $\mathbb{R}^n$ and the weight $w_t^\delta > 0$ is given by

$$w_t^\delta(x,y) = \frac{1}{\Gamma(\delta)} \int_{\mathbb{R}^n} e^{-2u\cdot x} \left(1 - \frac{|u|^2 + |y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, du.$$

Proof. Let F be holomorphic in the tube domain $\Omega_t = \{z = x + iy : |y| < t\}$ of $\mathbb{C}^n$. Now since the Lebesgue measure is rotationally invariant, $(1-(|u|^2+|y|^2)/t^2)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, dy\, du$ is a rotation-invariant measure. So, using Gutzmer's formula, we have

$$\int_{\mathbb{R}^{2n}} \left(\int_{\mathbb{R}^n} |\pi(iy, iv) F(\xi)|^2 \, d\xi \right) \left(1 - \frac{|u|^2+|y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, dy\, du$$
$$= c_n \sum_{k=0}^{\infty} \|P_k f\|_2^2 \frac{k!\,(n-1)!}{(k+n-1)!} \int_{\mathbb{R}^{2n}} \varphi_k(2iy, 2iu) \left(1 - \frac{|u|^2+|y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, dy\, du. \quad (3\text{-}22)$$

Integrating in polar coordinates, the integral on the right-hand side becomes

$$\int_{\mathbb{R}^{2n}} \varphi_k(2iy, 2iu) \left(1 - \frac{|u|^2+|y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, dy\, du = \omega_{2n} \int_0^\infty L_k^{n-1}(-2r^2) \left(1 - \frac{r^2}{t^2}\right)_+^{\delta-1} r^{2n-1}\, dr.$$

Now using a change of variable $r \to rt$ followed by another change of variable $r \to \sqrt{r}$ in the integral in the right-hand side of the above equation, we have

$$\int_0^\infty L_k^{n-1}(-2r^2) \left(1 - \frac{r^2}{t^2}\right)_+^{\delta-1} r^{2n-1}\, dr = \frac{1}{2} t^{2n} \int_0^1 L_k^{n-1}(r(-2t^2))(1-r)^{\delta-1} r^{n-1}\, dr.$$

By making use of the following identity (see [Szegő 1967]),

$$L_k^{\alpha}(t) = \frac{\Gamma(k+\alpha+1)}{\Gamma(\alpha-\beta)\Gamma(k+\beta+1)} \int_0^1 (1-r)^{\alpha-\beta-1} r^{\beta} L_k^{\beta}(rt)\, dr,$$

the above yields

$$\frac{1}{\Gamma(\delta)} \int_{\mathbb{R}^{2n}} \varphi_k(2iy, 2iu) \left(1 - \frac{|u|^2+|y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, dy\, du$$
$$= \frac{1}{2} t^{2n} \omega_{2n} \frac{\Gamma(k+n)}{\Gamma(k+n+\delta)} L_k^{n+\delta-1}(-2t^2). \quad (3\text{-}23)$$

Now we simplify the left-hand side of (3-22):

$$\frac{1}{\Gamma(\delta)} \int_{\mathbb{R}^{2n}} \left(\int_{\mathbb{R}^n} |\pi(iy, iv) F(\xi)|^2 \, d\xi \right) \left(1 - \frac{|u|^2+|y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, dy\, du$$
$$= \frac{1}{\Gamma(\delta)} \int_{\mathbb{R}^{2n}} \left(\int_{\mathbb{R}^n} |e^{i(iy\cdot\xi + iy\cdot iu/2)} F(\xi+iu)|^2 \, d\xi \right) \left(1 - \frac{|u|^2+|y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, dy\, du$$
$$= \frac{1}{\Gamma(\delta)} \int_{\mathbb{R}^{2n}} \left(\int_{\mathbb{R}^n} |e^{-2y\cdot\xi} F(\xi+iu)|^2 \, d\xi \right) \left(1 - \frac{|u|^2+|y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, dy\, du$$
$$= \int_{\mathbb{R}^{2n}} |F(\xi+iu)|^2 \left(\frac{1}{\Gamma(\delta)} \int_{\mathbb{R}^n} e^{-2y\cdot\xi} \left(1 - \frac{|u|^2+|y|^2}{t^2}\right)_+^{\delta-1} e^{-(|u|^2+|y|^2)}\, dy \right) d\xi\, du$$
$$= \int_{\mathbb{R}^{2n}} |F(\xi+iu)|^2 w_t^{\delta}(u, \xi)\, d\xi\, du.$$

Now, when $|u| \geq t$, we see that $(1-(|u|^2+|y|^2)/t^2)_+^{\delta-1} = 0$ for all $y \in \mathbb{R}^n$. Thus,

$$\int_{\mathbb{R}^{2n}} |F(\xi+iu)|^2 w_t^\delta(u,\xi)\,d\xi\,du = \int_{\mathbb{R}^n}\int_{|u|<t} |F(\xi+iu)|^2 w_t^\delta(u,\xi)\,d\xi\,du.$$

Finally, we have

$$\int_{\mathbb{R}^n}\int_{|u|<t} |F(\xi+iu)|^2 w_t^\delta(u,\xi)\,d\xi\,du = c_n \sum_{k=0}^{\infty} \|P_k f\|_2^2 \frac{\Gamma(k+1)\Gamma(n+\delta)}{\Gamma(k+n+\delta)} L_k^{n+\delta-1}(-2t^2)t^{2n}. \qquad \square$$

For $s > 0$, we consider the following positive weight function $\widetilde{w}_\rho(k)$ on $\mathbb{N}$ given by the sequence

$$\left(\tfrac{1}{2}\rho^2\right)^{s-1}\left(\Gamma\left(\tfrac{1}{2}(2k+n+1+s)\right)W_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right)\right)^2 \frac{\Gamma(k+1)\Gamma(n+2s)}{\Gamma(k+n+2s)} L_k^{n+2s-1}\left(-\tfrac{1}{2}\rho^2\right).$$

We define $W_\rho^s(\mathbb{R}^n)$ to be the space of all tempered distributions f for which

$$\|f\|_{s,\rho}^2 := \sum_{k=0}^{\infty} \widetilde{w}_\rho(k)\|P_k f\|_2^2 < \infty.$$

Remark. For $r < 0$, the following asymptotic property of Laguerre functions is well known (see [Szegő 1967, Theorem 8.22.3]) and is valid for large k, for $r \leq -c$ and for $c > 0$:

$$L_k^\alpha(r) = \frac{1}{2\sqrt{\pi}} e^{r/2}(-r)^{-\alpha/2-1/4} k^{\alpha/2-1/4} e^{2\sqrt{-kr}}(1+O(k^{-1/2})). \tag{3-24}$$

The asymptotic property (3-9) together with (3-17) gives

$$\left(\tfrac{1}{2}\rho^2\right)^{s-1}\left(\Gamma\left(\tfrac{1}{2}(2k+n+1+s)\right)W_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right)\right)^2 \leq c_1(\rho\sqrt{2k+n})^{2s-1}e^{-\rho\sqrt{2k+n}},$$

and from (3-24) we have

$$L_k^{n+2s-1}\left(-\tfrac{1}{2}\rho^2\right) \leq c e^{\rho^2/4}\rho^{-n-2s+1/2}(2k+n)^{(n+2s-1)/2-1/4}e^{\rho\sqrt{2k+n}}.$$

Now using the fact that $\Gamma(k+1)\Gamma(n+2s)/\Gamma(k+n+2s) \sim (2k+n)^{-(n+2s-1)}$, we have

$$\widetilde{w}_\rho(k) \leq c_1 e^{\rho^2/4}(\rho^2(2k+n))^{-(2n+1)/4}.$$

On the other hand, using (3-9) and (3-24), for large k, we have

$$\widetilde{w}_\rho(k) \geq c_2 e^{\rho^2/4}(\rho^2(2k+n))^{-(2n+1)/4}e^{-\psi_\rho(k)},$$

where $\psi_\rho(k) = 4(2k+n)\zeta(\rho^2/(4(2k+n)))^{1/2} - \rho\sqrt{2k}$. It can be checked that for $0 < \rho \leq 1$, the function $\psi_\rho(k)$ is decreasing in k, whence $\psi_\rho(k) \leq c$ for some constant c depending on ρ. So finally we have

$$c_2 e^{\rho^2/4}(\rho^2(2k+n))^{-(2n+1)/4} \leq \widetilde{w}_\rho(k) \leq c_1 e^{\rho^2/4}(\rho^2(2k+n))^{-(2n+1)/4}. \tag{3-25}$$

By letting $m_n = -\frac{1}{8}(2n+1)$, we clearly see that $f \in W_\rho^s(\mathbb{R}^n)$ if and only if $f \in W_H^{m_n}(\mathbb{R}^n)$ whenever $0 < \rho \leq 1$. Here $W_H^m(\mathbb{R}^n)$ denotes the Hermite Sobolev spaces.

In view of the connection between the operators H and L, to prove Theorem 1.7 it suffices to prove the following characterisation for the solution of the extension problem for H. Note that the extension problem for the Hermite operator H we are talking about reads as

$$\Big(-H+\partial_\rho^2+\frac{1-2s}{\rho}\partial_\rho-\frac{1}{4}\rho^2\Big)u(x,\rho)=0,\quad u(x,0)=f(x).$$

For $\rho>0$, let T_ρ stand for the operator defined for reasonable f by

$$T_\rho f(x):=\big(\tfrac{1}{2}\rho^2\big)^{(s-1)/2}\frac{1}{\Gamma(s)}\sum_{k=0}^{\infty}\Gamma\big(\tfrac{1}{2}(2k+n+s+1)\big)W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)P_kf(x).$$

Using similar reasoning as in the case of L, we point out that for a tempered distribution f, the above expression makes sense and solves the extension problem for H. Moreover, in view of the relation $Q_k=M_\gamma^{-1}P_kM_\gamma$, we have $T_\rho f=M_\gamma^{-1}S_\rho^1M_\gamma^{-1}f$. Thus Theorem 1.7 easily follows from the following:

Theorem 3.7. *A solution of the extension problem for H is of the form $u(x,\rho)=T_\rho f(x)$ for some $f\in W_H^{m_n}(\mathbb{R}^n)$ if and only if for every $\rho>0$, $u(\,\cdot\,,\rho)$ extends holomorphically to $\Omega_{\rho/2}$ and satisfies the estimate*

$$\int_{\Omega_{\rho/2}}|u(z,\rho)|^2w_{\rho/2}^{2s}(z)\,dz\le C\rho^{n-1/2},\tag{3-26}$$

for all $0<\rho\le 1$.

Proof. First suppose $u(x,\rho)=T_\rho f(x)$ for some f such that $f\in W_H^{m(s)}(\mathbb{R}^n)$. So clearly

$$u(x,\rho)=\frac{\big(\tfrac{1}{2}\rho^2\big)^{(s-1)/2}}{\Gamma(s)}\sum_{k=0}^{\infty}\Gamma\big(\tfrac{1}{2}(2k+n+s+1)\big)W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)P_kf(x).$$

But the Hermite function $\Phi_\alpha(x)=H_\alpha(x)e^{-|x|^2/2}$ has holomorphic extension to $\mathbb{C}^n$. Let $\Phi_k(z,w):=\sum_{|\alpha|=k}\Phi_\alpha(z)\Phi_\alpha(w)$. Then using the estimate (see [Thangavelu 2010])

$$|\Phi_k(z,\bar z)|\le C(y)(2k+n)^{3(n-1)/4}e^{2\sqrt{2k+n}|y|}$$

along with the asymptotic property (3-9), we conclude that the series

$$\sum_{k=0}^{\infty}\Gamma\big(\tfrac{1}{2}(2k+n+s+1)\big)W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)P_kf(z)$$

converges uniformly over compact subsets of $\Omega_{\rho/2}$ and hence defines a holomorphic function in the domain $\Omega_{\rho/2}$. Now noting that

$$\|P_ku(\,\cdot\,,\rho)\|_2^2=\rho^{2s-2}c_s^2\Big(\Gamma\big(\tfrac{1}{2}(2k+n+s+1)\big)W_{-(k+n/2),\,s/2}\big(\tfrac{1}{2}\rho^2\big)\Big)^2\|P_kf\|_2^2,$$

in view of Proposition 3.6 we obtain

$$\int_{\mathbb{R}^n}\int_{|y|<\rho/2}|u(x+iy,\rho)|^2w_{\rho/2}^{2s}(x,y)\,dx\,dy=c_n\rho^{2n}\sum_{k=0}^{\infty}\widetilde{w}_\rho(k)\|P_kf\|_2^2.$$

But in view of (3-25),

$$\|f\|_{s,\rho}^2 \le Ce^{\rho^2/4}\rho^{-(2n+1)/2}\sum_{k=0}^{\infty}(2k+n)^{2m_n}\|P_kf\|_2^2,$$

which gives

$$\int_{\mathbb{R}^n}\int_{|y|<\rho/2}|u(x+iy,\rho)|^2 w_{\rho/2}^{2s}(x,y)\,dx\,dy \le Ce^{\rho^2/4}\rho^{n-1/2}\|f\|^2_{W_H^{m_n}},$$

proving the first part of the theorem.

Conversely, let $u(z,\rho)$ be holomorphic on $\Omega_{\rho/2}$ for every $\rho>0$ satisfying the estimate (3-26). Let g_ρ be a tempered distribution such that

$$P_ku(\cdot,\rho) = \left(\tfrac{1}{2}\rho^2\right)^{(s-1)/2}\Gamma\left(\tfrac{1}{2}(2k+n+s+1)\right)W_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right)P_kg_\rho. \tag{3-27}$$

Now for $0<\rho\le 1$, using (3-25) we have

$$\|g_\rho\|^2_{W_H^{m_n}} \le Ce^{-\rho^2/4}\rho^{(2n+1)/2}\sum_{k=0}^{\infty}\widetilde{w}_\rho(k)\|P_kg_\rho\|_2^2.$$

Note that using Proposition 3.6 we obtain

$$\int_{\Omega_{\rho/2}}|u(z,\rho)|^2 w_{\rho/2}^{2s}(z)\,dz = c_n\rho^{2n}\sum_{k=0}^{\infty}\widetilde{w}_\rho(k)\|P_kg_\rho\|_2^2,$$

which by the hypothesis yields $\|g_\rho\|^2_{W_H^{m_n}} \le C$ for all $0<\rho\le 1$. Now by the Banach–Alaoglu theorem, we choose a sequence $\{\rho_m\}$ going to 0 such that g_{ρ_m} converges weakly in $W_H^{m_n}(\mathbb{R}^n)$ as $k\to\infty$. Let f be the weak limit in this case. Now given $\varphi\in S(\mathbb{R}^n)$, we have

$$\int_{\mathbb{R}^n}u(x,\rho_m)\varphi(x)\,dx = \sum_{k=0}^{\infty}\int_{\mathbb{R}^n}P_ku(x,\rho_m)\overline{P_k\varphi(x)}\,dx.$$

But using (3-27), the above integral equals

$$\sum_{k=0}^{\infty}\left(\tfrac{1}{2}\rho_m^2\right)^{(s-1)/2}\Gamma\left(\tfrac{1}{2}(2|\alpha|+n+s+1)\right)W_{-(|\alpha|+n/2),\,s/2}\left(\tfrac{1}{2}\rho_m^2\right)\int_{\mathbb{R}^n}P_kg_{\rho_m}(x)\overline{P_k\varphi(x)}\,dx.$$

This allows us to conclude that $u(\cdot,\rho_m)$ converges to f in the sense of distribution. Now under the assumption that u solves the extension problem for H, the exact same argument as in the beginning of this subsection gives

$$\hat{u}(\alpha,\rho) = \left(\tfrac{1}{2}\rho^2\right)^{(s-1)/2}\left(C_1(|\alpha|)W_{-(|\alpha|+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right)+C_2(|\alpha|)M_{-(|\alpha|+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right)\right),$$

where $\hat{u}(\alpha,\rho)$ denotes the Hermite coefficients. But the estimate (3-26) gives

$$\sum_{k=0}^{\infty}\left(C_2(k)M_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right)\right)^2 L_k^{n+2s-1}\left(-\tfrac{1}{2}\rho^2\right) \le C(\rho).$$

But since both $M_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho^2\right)$ and $L_k^{n+2s-1}\left(-\tfrac{1}{2}\rho^2\right)$ have exponential growth in k (see (3-8) and (3-24)), the above inequality forces $C_2(k)$ to be zero. Now, as $u(\cdot,\rho_m)$ converges to f and as ρ_m tends to zero, $\left(\tfrac{1}{2}\rho_m^2\right)^{(s-1)/2}W_{-(k+n/2),\,s/2}\left(\tfrac{1}{2}\rho_m^2\right)$ goes to a constant $\Gamma(s)/\Gamma\left(\tfrac{1}{2}(2k+n+s+1)\right)$ (see (3-5)), and the theorem follows. □

4. Trace Hardy and Hardy's inequality

Trace Hardy inequality. We prove the following trace Hardy inequality only for the operator U as the case of L is similar. We shall work with the gradient on $\mathbb{R}^n \times [0, \infty)$ defined by

$$\nabla_U u := (2^{-1/2}\partial_1 u, 2^{-1/2}\partial_2 u, \ldots, 2^{-1/2}\partial_n u, \partial_\rho u).$$

We also let $P_s(\partial_x, \partial_\rho) = \left(-U + \partial_\rho^2 + \frac{1-2s}{\rho}\partial_\rho - \frac{1}{4}\rho^2\right)$ stand for the extension operator.

Lemma 4.1. *Let u and v be two real-valued functions on $\mathbb{R}^n \times [0, \infty)$ such that $u, v \in C_0^2([0, \infty), C^2(\mathbb{R}^n))$. Then for $0 < s < 1$ we have*

$$\begin{aligned}\int_0^\infty \int_{\mathbb{R}^n} \left| \nabla_U u(x, \rho) - \frac{u(x, \rho)}{v(x, \rho)} \nabla_U v(x, \rho) \right|^2 \rho^{1-2s} \, d\gamma(x) \, d\rho \\ = \int_0^\infty \int_{\mathbb{R}^n} \left(|\nabla_U u(x, \rho)|^2 + \left(\tfrac{1}{2}n + \tfrac{1}{4}\rho^2\right) u(x, \rho)^2 \right) \rho^{1-2s} \, d\gamma(x) \, d\rho \\ + \int_0^\infty \int_{\mathbb{R}^n} \frac{u(x, \rho)^2}{v(x, \rho)} (P_s(\partial_x, \partial_\rho) v(x, \rho)) \rho^{1-2s} \, d\gamma(x) \, d\rho \\ + \int_{\mathbb{R}^n} \frac{u(x, 0)^2}{v(x, 0)} \lim_{\rho \to 0} (\rho^{1-2s} \partial_\rho v)(x, \rho) \, d\gamma(x).\end{aligned} \tag{4-1}$$

Proof. For any $1 \le j \le n$, we consider the integral

$$\int_{\mathbb{R}^n} \left(\partial_j u - \frac{u}{v} \partial_j v \right)^2 d\gamma(x) = \int_{\mathbb{R}^n} \left((\partial_j u)^2 - 2\frac{u}{v} \partial_j u \partial_j v + \frac{u^2}{v^2} (\partial_j v)^2 \right) d\gamma(x). \tag{4-2}$$

Now by the definition of adjoint we get

$$\int_{\mathbb{R}^n} \frac{u}{v} \partial_j u \partial_j v \, d\gamma(x) = \int_{\mathbb{R}^n} u \partial_j^* \left(\frac{u}{v} \partial_j v \right) d\gamma(x).$$

Using the fact that $\partial_j^* = 2x_j - \partial_j$ on $L^2(\gamma)$, we have

$$u \partial_j^* \left(\frac{u}{v} \partial_j v \right) = 2x_j \frac{u^2}{v} \partial_j v - \frac{u}{v} \partial_j u \partial_j v - u^2 \partial_j \left(\frac{1}{v} \partial_j v \right),$$

which together with the above equation yields

$$\begin{aligned}2 \int_{\mathbb{R}^n} \frac{u}{v} \partial_j u \partial_j v \, d\gamma(x) &= \int_{\mathbb{R}^n} \left(2x_j \frac{u^2}{v} \partial_j v - u^2 \partial_j \left(\frac{1}{v} \partial_j v \right) \right) d\gamma(x) \\ &= \int_{\mathbb{R}^n} \left(2x_j \frac{u^2}{v} \partial_j v - \frac{u^2}{v} \partial_j^2 v + \frac{u^2}{v^2} (\partial_j v)^2 \right) d\gamma(x).\end{aligned}$$

Hence we have

$$\int_{\mathbb{R}^n} \left(\frac{u^2}{v^2} (\partial_j v)^2 - 2\frac{u}{v} \partial_j u \partial_j v \right) d\gamma(x) = -\int_{\mathbb{R}^n} \frac{u^2}{v} \partial_j^* \partial_j v \, d\gamma(x).$$

Similarly, for any $x \in \mathbb{R}^n$ one can obtain

$$\int_0^\infty \left(\frac{u^2}{v^2}(\partial_\rho v)^2 - 2\frac{u}{v}\partial_\rho u \partial_\rho v\right)\rho^{1-2s}\,d\rho = \int_0^\infty \frac{u^2}{v^2}\partial_\rho(\rho^{1-2s}\partial_\rho v)\,d\rho + \frac{u(x,0)^2}{v(x,0)}\lim_{\rho\to 0}(\rho^{1-2s}\partial_\rho v)(x,\rho).$$

Multiplying both side of (4-2) by $\frac{1}{2}$ and summing over j we get the required result. □

Theorem 4.2 (general trace Hardy inequality). *Let $0 < s < 1$. Suppose $\phi \in L^2(\gamma)$ is a real-valued function in the domain of U_s such that $\phi^{-1}U_s\phi$ is locally integrable. Then for any real-valued function $u(x,\rho)$ from the space $C_0^2([0,\infty), C_b^2(\mathbb{R}^n))$ we have*

$$\int_0^\infty\int_{\mathbb{R}^n}\big(|\nabla_U u(x,\rho)|^2 + \big(\tfrac{1}{2}n + \tfrac{1}{4}\rho^2\big)u(x,\rho)^2\big)\rho^{1-2s}\,d\gamma(x)\,d\rho \ge C_{n,s}\int_{\mathbb{R}^n} u(x,0)^2\frac{L_s\phi(x)}{\phi(x)}\,d\gamma(x).$$

Proof. To prove this result, we make use of Lemma 4.1. Since the left-hand side of (4-1) is always nonnegative, we have, for $0 < s < 1$,

$$\begin{aligned}\int_0^\infty\int_{\mathbb{R}^n}&\big(|\nabla_U u(x,\rho)|^2 + \big(\tfrac{1}{2}n + \tfrac{1}{4}\rho^2\big)u(x,\rho)^2\big)\rho^{1-2s}\,d\gamma(x)\,d\rho \\ &\ge -\int_0^\infty\int_{\mathbb{R}^n}\frac{u(x,\rho)^2}{v(x,\rho)}(P_s(\partial_x,\partial_\rho)v(x,\rho))\rho^{1-2s}\,d\gamma(x)\,d\rho \\ &\qquad -\int_{\mathbb{R}^n}\frac{u(x,0)^2}{v(x,0)}\lim_{\rho\to 0}(\rho^{1-2s}\partial_\rho v)(x,\rho)\,d\gamma(x). \qquad (4\text{-}3)\end{aligned}$$

Now we take

$$v(x,\rho) = \frac{4^{-s}}{\Gamma(s)}\rho^{2s}\int_0^\infty k_{t,s}(\rho)e^{-tL}\phi(x)\,dt.$$

Then v solves the extension equation (2-1), i.e., $P_s(\partial_x,\partial_\rho)v = 0$ and $v(x,0) = \phi(x)$. Then from (4-3), we have

$$\begin{aligned}\int_0^\infty\int_{\mathbb{R}^n}&\big(|\nabla_U u(x,\rho)|^2 + \big(\tfrac{1}{2}n + \tfrac{1}{4}\rho^2\big)u(x,\rho)^2\big)\rho^{1-2s}\,d\gamma(x)\,d\rho \\ &\ge -\int_{\mathbb{R}^n}\frac{u(x,0)^2}{v(x,0)}\lim_{\rho\to 0}(\rho^{1-2s}\partial_\rho v)(x,\rho)\,d\gamma(x). \qquad (4\text{-}4)\end{aligned}$$

In view of the above, we need to solve the extension problem for U with a given initial condition ϕ. Since

$$-\lim_{\rho\to 0}\rho^{1-2s}\partial_\rho u(x,\rho) = 2^{1-2s}\frac{\Gamma(1-s)}{\Gamma(s)}U_s\phi,$$

we get the desired inequality. □

Corollary 4.3. *Let $0 < s < 1$ and $f \in L^2(\gamma)$ with $U_s f \in L^2(\gamma)$. Then we have*

$$\langle U_s f, f\rangle_{L^2(\gamma)} \ge \int_{\mathbb{R}^n} f^2(x)\frac{U_s\phi}{\phi}\,d\gamma(x)$$

for any real-valued ϕ in the domain of U_s.

Proof. When u itself solves the extension problem with initial condition f, the proof of Lemma 4.1 shows that the left-hand side of the trace Hardy inequality reduces to $\langle U_s f, f\rangle_{L^2(\gamma)}$. □

Hardy's inequality from trace Hardy. In this subsection we construct a suitable function ϕ so that $(U_s\phi)/\phi$ simplifies. In order to do so, let us quickly recall some basic facts about Laguerre functions. Let $\alpha > -1$ and $k \in \mathbb{N}$. The Laguerre polynomial of degree k and type α, which we denote by $L_k^\alpha(x)$, is a solution of the ordinary differential equation

$$xy''(x) + (\alpha + 1 - x)y'(x) + ky(x) = 0,$$

whose explicit expression is given by

$$L_k^\alpha(x) = \sum_{j=0}^{k} \frac{\Gamma(k+\alpha+1)}{\Gamma(k-j+1)\Gamma(j+\alpha+1)} \frac{(-x)^j}{j!}. \tag{4-5}$$

Recall that the Laguerre functions of type $(n-1)$ are given by

$$\varphi_k^{n-1}(r) = L_k^{n-1}\big(\tfrac{1}{2}r^2\big)e^{-r^2/4}, \quad r \geq 0.$$

For more details about such functions we refer the reader to [Thangavelu 1993, Chapter 1]. Now given $s, \rho > 0$, we consider the function $\phi_{s,\rho}$ which is defined in terms of Laguerre polynomials as follows:

$$\phi_{s,\rho}(x) = \sum_{m=0}^{\infty} C_{2m,\rho}(s) L_m^{n/2-1}(|x|^2) = e^{|x|^2/2} \sum_{m=0}^{\infty} C_{2m,\rho}(s)\varphi_m^{n/2-1}(\sqrt{2}|x|),$$

where the coefficients are given in terms of the L function as

$$C_{k,\rho}(s) = \frac{2\pi}{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)^2} L\big(\rho, \tfrac{1}{4}(2k+n) + \tfrac{1}{2}(1+s), \tfrac{1}{4}(2k+n) + \tfrac{1}{2}(1-s)\big).$$

In the following lemma we show how these functions are related via the fractional power of the operator under study.

Lemma 4.4. *For $-1 < s < 1$, we have*

$$U_s\phi_{-s,\rho} = \frac{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)^2}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)^2} (4\rho)^s \phi_{s,\rho}. \tag{4-6}$$

Proof. Let us take two radial functions g and h on $\mathbb{R}^n$ such that

$$g(x) = \pi^{n/2} e^{|x|^2/2} h(x),$$

where $h \in L^2(\mathbb{R}^n)$. Moreover, we choose h in such a way that the Laguerre coefficients

$$R_m^{n/2-1}(h) = 2\frac{\Gamma(m+1)}{\Gamma(m+n/2)} \int_0^\infty h(r) L_m^{n/2-1}(r^2) e^{-r^2/2} r^{n-1}\, dr$$

are nonzero. By our choice of h and definition of g, it is not hard to see that $Q_k g(x) = e^{|x|^2/2} P_k h(x)$. Also, since h is radial, using a result proved in [Thangavelu 1993, Theorem 3.4.1] we have

$$P_k h(x) = \begin{cases} 0 & \text{if } k = 2m+1, \\ R_m^{n/2-1}(h) L_m^{n/2-1}(|x|^2) e^{-|x|^2/2} & \text{if } k = 2m. \end{cases} \tag{4-7}$$

Now using the definition of the Laguerre function along with the fact that $R_m^{n/2-1}(h) \neq 0$, we see that

$$\begin{aligned}\phi_{s,\rho}(x) &= e^{|x|^2/2} \sum_{m=0}^{\infty} \frac{2\pi}{\Gamma(n/2+1+s)^2} L\big(\rho, \tfrac{1}{4}(4m+n)+\tfrac{1}{2}(1+s), \tfrac{1}{4}(4m+n)+\tfrac{1}{2}(1-s)\big)\varphi_m^{n/2-1}(\sqrt{2}|x|)\\ &= e^{|x|^2/2} \sum_{m=0}^{\infty} C_{2m,\rho}(s)(R_m^{n/2-1}(h))^{-1} R_m^{n/2-1}(h) L_m^{n/2}(|x|^2) e^{|x|^2/2}.\end{aligned}$$

But observation (4-7) and the fact that

$$Q_k g(x) = e^{|x|^2/2} P_k h(x)$$

transform the above equation into

$$\phi_{s,\rho}(x) = \sum_{k=0}^{\infty} C_{k,\rho}(s)(R_{\lfloor k/2\rfloor}^{n/2-1}(h))^{-1} Q_k g(x). \tag{4-8}$$

Hence using the definition of U_s we have

$$U_s\phi_{-s,\rho} = \sum_{k=0}^{\infty} C_{k,\rho}(-s) 2^s \frac{\Gamma\big(\frac{1}{4}(2k+n)+\frac{1}{2}(1+s)\big)}{\Gamma\big(\frac{1}{4}(2k+n)+\frac{1}{2}(1-s)\big)} (R_{\lfloor k/2\rfloor}^{n/2-1}(h))^{-1} Q_k g. \tag{4-9}$$

But in view of the transformation property (2-12), we have

$$\begin{aligned}&C_{k,\rho}(-s)\\ &= \frac{2\pi}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)^2} L\big(\rho, \tfrac{1}{4}(2k+n)+\tfrac{1}{2}(1-s), \tfrac{1}{4}(2k+n)+\tfrac{1}{2}(1+s)\big)\\ &= \frac{2\pi}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)^2} (2\rho)^s \frac{\Gamma\big(\frac{1}{4}(2k+n)+\frac{1}{2}(1-s)\big)}{\Gamma\big(\frac{1}{4}(2k+n)+\frac{1}{2}(1+s)\big)} L\big(\rho, \tfrac{1}{4}(2k+n)+\tfrac{1}{2}(1+s), \tfrac{1}{4}(2k+n)+\tfrac{1}{2}(1-s)\big)\\ &= \frac{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)^2}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)^2} (2\rho)^s \frac{\Gamma\big(\frac{1}{4}(2k+n)+\frac{1}{2}(1-s)\big)}{\Gamma\big(\frac{1}{4}(2k+n)+\frac{1}{2}(1+s)\big)} C_{k,\rho}(s). \end{aligned}\tag{4-10}$$

Hence from (4-9) we obtain

$$U_s\phi_{-s,\rho} = \frac{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)^2}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)^2} (4\rho)^s \phi_{s,\rho}. \qquad \square$$

Now in the rest of the section we will calculate $\phi_{s,\rho}$ almost explicitly in terms of the Macdonald's function K_ν, defined for $z > 0$ by the integral

$$K_\nu(z) := 2^{-\nu-1} z^\nu \int_0^\infty e^{-t-z^2/(4t)} t^{-\nu-1}\, dt.$$

Proposition 4.5. *Let $0 < s < 1$ and $\rho > 0$. Then we have*

$$\phi_{s,\rho}(x) = 2\frac{\sqrt{\pi}\, 2^{-(n/2+1+s)/2}}{\sqrt{2\pi}\,\Gamma\big(\frac{1}{2}(n/2+1+s)\big)} e^{|x|^2/2} (\rho+|x|^2)^{-(n/2+1+s)/2} K_{(n/2+1+s)/2}(\rho+|x|^2). \tag{4-11}$$

Proof. First we note the following formula proved in [Ciaurri et al. 2018, Lemma 3.8]:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\lambda t}((\rho+r^2)^2+t^2)^{(\alpha+2+s)/2}\,dt = |\lambda|^{\alpha+1}\sum_{k=0}^{\infty} c_{k,\rho}^{\lambda}(s)\varphi_k^{\alpha}(\sqrt{(2|\lambda|)}r), \tag{4-12}$$

where the coefficients $c_{k,\rho}^{\lambda}(s)$ are given by

$$c_{k,\rho}^{\lambda}(s) = \frac{2\pi|\lambda|^s}{\Gamma\big(\frac{1}{2}(\alpha+2+s)\big)^2} L\big(\rho|\lambda|, \tfrac{1}{4}(4k+2\alpha+2)+\tfrac{1}{2}(1+s), \tfrac{1}{4}(4k+2\alpha+2)+\tfrac{1}{2}(1-s)\big).$$

This holds for any $\lambda \neq 0$ and $\alpha > -\frac{1}{2}$. In particular, taking $\alpha = \frac{1}{2}n-1$ and $\lambda = 1$ in (4-12), we have

$$\phi_{s,\rho}(x) = e^{|x|^2/2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{it}((\rho+|x|^2)^2+t^2)^{-(n/2+1+s)/2}\,dt. \tag{4-13}$$

The right-hand side of the above equation can be computed in terms of the Macdonald's function K_ν. Now we make use of the formula (see [Prudnikov et al. 1986, p. 390])

$$\int_0^{\infty}\frac{\cos br}{(r^2+z^2)^{\delta}}\,dr = \left(\frac{2z}{b}\right)^{1/2-\delta}\frac{\sqrt{\pi}}{\Gamma(\delta)}K_{1/2-\delta}(bz), \tag{4-14}$$

which is valid for $b > 0$ and $\Re\delta, \Re z > 0$. This gives

$$\begin{aligned}\int_{-\infty}^{\infty} & e^{it}((\rho+|x|^2)^2+t^2)^{-(n/2+1+s)/2}\,dt \\ &= 2\frac{\sqrt{\pi}2^{-(n/2+1+s)/2}}{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)}(\rho+|x|^2)^{-(n/2+1+s)/2}K_{-(n/2+1+s)/2}(\rho+|x|^2).\end{aligned} \tag{4-15}$$

Now using the fact that $K_\nu = K_{-\nu}$, we obtain

$$\phi_{s,\rho}(x) = 2\frac{\sqrt{\pi}2^{-(n/2+1+s)/2}}{\sqrt{2\pi}\,\Gamma\big(\frac{1}{2}(n/2+1+s)\big)}e^{|x|^2/2}(\rho+|x|^2)^{-(n/2+1+s)/2}K_{(n/2+1+s)/2}(\rho+|x|^2), \tag{4-16}$$

proving the proposition. □

We are now ready to prove Theorem 1.2. For the convenience of the reader we state the theorem here as well.

Theorem 4.6. *Let $0 < s < 1$. Assume that $f \in L^2(\gamma)$ such that $U_s f \in L^2(\gamma)$. Then for every $\rho > 0$ we have*

$$\langle U_s f, f\rangle_{L^2(\gamma)} \geq (2\rho)^s\frac{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)}\int_{\mathbb{R}^n}\frac{f(x)^2}{(\rho+|x|^2)^s}w_s(\rho+|x|^2)\,d\gamma(x)$$

for an explicit $w_s(t) \geq 1$. The inequality is sharp, and equality is attained for $f(x) = \phi_{-s,\rho}(x)$.

Proof. Taking $\phi = \phi_{-s,\rho}$ in 4.5, in view of Lemma 4.4 we have

$$\frac{U_s\phi}{\phi} = \frac{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)^2}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)^2}(4\rho)^s\frac{\phi_{s,\rho}}{\phi_{-s,\rho}}.$$

Now we use Proposition 4.5 to simplify the right-hand side of the above equation. Note that

$$\frac{\phi_{s,\rho}}{\phi_{-s,\rho}} = \frac{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)}{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)} 2^{-s}(\rho+|x|^2)^{-s} \frac{K_{(n/2+1+s)/2}(\rho+|x|^2)}{K_{(n/2+1-s)/2}(\rho+|x|^2)}. \tag{4-17}$$

Let

$$w_s(t) := \frac{K_{(n/2+1+s)/2}(t)}{K_{(n/2+1-s)/2}(t)}, \quad t>0.$$

Using the fact that $K_\nu(t)$ is an increasing function of ν for $t>0$, we note that $w_s(t)\geq 1$, for all $t>0$, and

$$\frac{U_s\phi}{\phi} = 2^s\rho^s \frac{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)}{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)} (\rho+|x|^2)^{-s} w_s(\rho+|x|^2).$$

Hence the required inequality follows from Corollary 4.3.

To see that equality holds for $f(x)=\phi_{-s,\rho}(x)$, using Lemma 4.4 we note that

$$\langle U_s\phi_{-s,\rho}, \phi_{-s,\rho}\rangle_{L^2(\gamma)} = \frac{\Gamma\big(\frac{1}{2}(n/2+1+s)\big)^2}{\Gamma\big(\frac{1}{2}(n/2+1-s)\big)^2} (4\rho)^s \int_{\mathbb{R}^n} \phi_{-s,\rho}(x)^2 \frac{\phi_{s,\rho}(x)}{\phi_{-s,\rho}(x)}\, d\gamma(x).$$

We finish by noting that (4-17) allows us to write the above as

$$\langle U_s\phi_{-s,\rho}, \phi_{-s,\rho}\rangle_{L^2(\gamma)} = (2\rho)^s \int_{\mathbb{R}^n} \frac{\phi_{-s,\rho}(x)^2}{(\rho+|x|^2)^s} w_s(\rho+|x|^2)\, d\gamma(x). \qquad \square$$

5. Isometry property for the solution of the extension problem

In this section we prove an isometry property of the solution operator associated to the extension problem for the Ornstein–Uhlenbeck operator under consideration. Such a property has been studied in the context of the extension problem for the Laplacian on $\mathbb{R}^n$ and for the sub-Laplacian on $\mathbb{H}^n$ in [Möllers et al. 2016]. See also the work of Roncal and Thangavelu [2020a], where they proved a similar result in the context of H-type groups.

We consider the Gaussian Sobolev space $\mathcal{H}^s_\gamma(\mathbb{R}^n)$ defined via the relation $f\in\mathcal{H}^s_\gamma(\mathbb{R}^n)$ if and only if $L_{s/2}f\in L^2(\gamma)$, where $L_{s/2}$ is the fractional power under consideration. Instead of $\|L_{s/2}f\|_2$, we use the equivalent norm for this space which is given by

$$\|f\|^2_{(s)} := \langle L_s f, f\rangle_{L^2(\gamma)} = \sum_{\alpha\in\mathbb{N}^n} 2^s \frac{\Gamma\big(\frac{1}{2}(2|\alpha|+n)+\frac{1}{2}(1+s)\big)}{\Gamma\big(\frac{1}{2}(2|\alpha|+n)+\frac{1}{2}(1-s)\big)} |\langle f, H_\alpha\rangle_{L^2(\gamma)}|^2.$$

Recall that the H_α are the normalised Hermite polynomials on $\mathbb{R}^n$ forming an orthonormal basis for $L^2(\gamma)$. As the solution of the extension equation (2-1) is a function of ρ^2, it can be thought of as a function of $(x,y)\in\mathbb{R}^{n+2}$ that is radial in y. Thus it makes sense to define $P_sf(x,y)=u(x,\sqrt{2}|y|)$, where $u(x,\rho)$ is the solution of the extension equation (2-1) given by (2-3). We can now consider $P_sf(x,y)$ as an element of $L^2(\mathbb{R}^{n+2},\gamma)$. For $(\alpha,j)\in\mathbb{N}^n\times\mathbb{N}^2$, we let

$$H_{\alpha,j}(x,y) := H_\alpha(x)H_j(y), \quad (x,y)\in\mathbb{R}^n\times\mathbb{R}^2,$$

where the H_j are two-dimensional Hermite polynomials. Then $P_s f(x, y)$ can be expanded in terms of $H_{\alpha,j}(x, y)$. We will show that P_s takes $\mathcal{H}^s_\gamma(\mathbb{R}^n)$ into $\mathcal{H}^{s+1}(\mathbb{R}^{n+2})$. We equip $\mathcal{H}^{s+1}_\gamma(\mathbb{R}^{n+2})$ with a different but equivalent norm. For $u \in \mathcal{H}^{s+1}_\gamma(\mathbb{R}^{n+2})$, we define

$$\|u\|^2_{(1,s)} = \sum_{(\alpha,j)\in\mathbb{N}^n\times\mathbb{N}^2} 2^{s+1} \frac{\Gamma\big(\frac{1}{2}(2|\alpha|+2|j|+n+1)+\frac{1}{2}(1+(1+s))\big)}{\Gamma\big(\frac{1}{2}(2|\alpha|+2|j|+n+1)+\frac{1}{2}(1-(1+s))\big)} |\langle u^j, H_\alpha\rangle_{L^2(\gamma)}|^2,$$

where for any $j \in \mathbb{N}^2$ we let

$$u^j(x) := \int_{\mathbb{R}^2} u(x, y) H_j(y) e^{-|y|^2/2}\, dy.$$

Equipped with this norm we denote the space $\mathcal{H}^{s+1}_\gamma(\mathbb{R}^{n+2})$ by $\widetilde{\mathcal{H}}^{s+1}_\gamma(\mathbb{R}^{n+2})$.

Theorem 5.1. *For $0 < s < n$, the function $P_s : \mathcal{H}^s_\gamma(\mathbb{R}^n) \to \widetilde{\mathcal{H}}^{s+1}_\gamma(\mathbb{R}^{n+2})$ is a constant multiple of an isometry, i.e., $\|P_s f\|_{(1,s)} = C_{n,s}\|f\|_{(s)}$ for all $f \in \mathcal{H}^s_\gamma(\mathbb{R}^n)$.*

Proof. We have

$$P_s f(x, y) = \sum_{k=0}^{\infty} \frac{2^{-s}}{\Gamma(s)} (\sqrt{2}|y|)^{2s} L\big(\tfrac{1}{2}|y|^2, \tfrac{1}{2}(2k+n)+\tfrac{1}{2}(1+s), \tfrac{1}{2}(2k+n)+\tfrac{1}{2}(1-s)\big) Q_k f.$$

Now from (2-13) we note that

$$\begin{aligned} P_s f(x, y) &= T_{-s,\sqrt{2}|y|}(L_s f)(x) \\ &= \sum_{k=0}^{\infty} \frac{4^s}{\Gamma(-s)} L\big(\tfrac{1}{2}|y|^2, \tfrac{1}{2}(2k+n)+\tfrac{1}{2}(1-s), \tfrac{1}{2}(2k+n)+\tfrac{1}{2}(1+s)\big) \frac{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1+s)\big)}{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1-s)\big)} Q_k f. \end{aligned}$$

Now writing $a := \frac{1}{2}(2k+n)+\frac{1}{2}(1+s)$ and $b := \frac{1}{2}(2k+n)+\frac{1}{2}(1-s)$, we expand $L\big(\frac{1}{2}|y|^2, a, b\big)$ in terms of Hermite polynomials. In order to do this, we use Mehler's formula (see [Urbina-Romero 2019, Chapter 1]) for two-dimensional normalised Hermite polynomials:

$$\sum_{j\in\mathbb{N}^2} H_j(x) H_j(y) r^{|j|} = (1-r^2)^{-1} \exp\bigg(-\frac{r^2(|x|^2+|y|^2)}{1-r^2} - \frac{2r x\cdot y}{1-r^2}\bigg).$$

In view of the definition of the L function, we have

$$L\big(\tfrac{1}{2}|y|^2, a, b\big) = e^{-|y|^2/2} \int_0^\infty e^{-t|y|^2} t^{a-1}(1+t)^{-b}\, dt.$$

Now taking $r^2 = t/(1+t)$ in the above Mehler's formula, we have

$$e^{-t|y|^2} = (1+t)^{-1} \sum_{j\in\mathbb{N}^2} H_j(0) H_j(y) \Big(\frac{t}{1+t}\Big)^{|j|/2},$$

which yields

$$L\big(\tfrac{1}{2}|y|^2, a, b\big) = e^{-|y|^2/2} \sum_{j\in\mathbb{N}^2} H_j(0)H_j(y) \int_0^\infty t^{a+|j|/2-1}(1+t)^{-b-|j|/2-1}\,dt$$
$$= e^{-|y|^2/2} \sum_{j\in\mathbb{N}^2} H_j(0)H_j(y) \frac{\Gamma(a+|j|/2)\Gamma(b-a+1)}{\Gamma(b+|j|/2+1)}.$$

Here the second equality follows from the formula

$$\int_0^\infty (1+t)^{-b} t^{a-1}\,dt = \frac{\Gamma(a)\Gamma(b-a)}{\Gamma(b)}. \tag{5-1}$$

Finally, writing $P_s f(x, y) = v(x, y)$ and using the above observations, we have

$$v(x, y) = c_s e^{-|y|^2/2} \sum_{(\alpha,j)\in\mathbb{N}^n\times\mathbb{N}^2} H_j(0)H_j(y) \frac{\Gamma(a+|j|/2)\Gamma(b-a+1)}{\Gamma(b+|j|/2+1)} \frac{\Gamma\big(\frac{1}{2}(2|\alpha|+n)+\frac{1}{2}(1+s)\big)}{\Gamma\big(\frac{1}{2}(2|\alpha|+n)+\frac{1}{2}(1-s)\big)} \langle f, H_\alpha\rangle_{L^2(\gamma)} H_\alpha(x),$$

where $c_s := 4^s/\Gamma(-s)$. Now note that for any $j \in \mathbb{N}^2$ we obtain

$$v^j(x) = c_s \sum_{\alpha\in\mathbb{N}^n} H_j(0) \frac{\Gamma(a+|j|/2)\Gamma(b-a+1)}{\Gamma(b+|j|/2+1)} \frac{\Gamma\big(\frac{1}{2}(2|\alpha|+n)+\frac{1}{2}(1+s)\big)}{\Gamma\big(\frac{1}{2}(2|\alpha|+n)+\frac{1}{2}(1-s)\big)} \langle f, H_\alpha\rangle_{L^2(\gamma)} H_\alpha(x),$$

which yields

$$\langle v^j, H_\alpha\rangle_{L^2(\gamma)} = c_s H_j(0) \frac{\Gamma(a+|j|/2)\Gamma(b-a+1)}{\Gamma(b+|j|/2+1)} \frac{\Gamma\big(\frac{1}{2}(2|\alpha|+n)+\frac{1}{2}(1+s)\big)}{\Gamma\big(\frac{1}{2}(2|\alpha|+n)+\frac{1}{2}(1-s)\big)} \langle f, H_\alpha\rangle_{L^2(\gamma)}.$$

As shown in [Urbina-Romero 2019], for any $k \in \mathbb{N}$ and for one-dimensional Hermite polynomials we have

$$H_{2k+1}(0) = 0 \quad\text{and}\quad (H_{2k}(0))^2 = \frac{2^{-2k}\Gamma(2k+1)}{\Gamma(k+1)^2}.$$

But making use of the formula $\Gamma(2z) = (2\pi)^{-1/2} 2^{2z-1/2}\Gamma(z)\Gamma\big(z+\frac{1}{2}\big)$, we obtain

$$(H_{2k}(0))^2 = \frac{1}{\sqrt{\pi}} \frac{\Gamma(k+1/2)}{\Gamma(k+1)}.$$

Hence, for $j = (j_1, j_2) \in \mathbb{N}^n$, we have

$$(H_{2j}(0))^2 = \frac{1}{\pi} \frac{\Gamma(j_1+1/2)\Gamma(j_2+1/2)}{\Gamma(j_1+1)\Gamma(j_2+1)}.$$

With these things in hand we proceed to calculate $\|v\|^2_{(1,s)}$, which is given by a constant multiple of

$$\sum_{k=0}^\infty \sum_{j\in\mathbb{N}^2} \left(\frac{\Gamma\big(\frac{1}{2}(2k+2|j|+n+1)+\frac{1}{2}(1+(1+s))\big)}{\Gamma\big(\frac{1}{2}(2k+2|j|+n+1)+\frac{1}{2}(1-(1+s))\big)} \times \left| H_j(0) \frac{\Gamma(a+|j|/2)\Gamma(b-a+1)}{\Gamma(b+|j|/2+1)} \frac{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1+s)\big)}{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1-s)\big)} \right|^2 \|Q_k f\|^2 \right),$$

where $\|Q_k f\|^2 = \sum_{|\alpha|=k} |\langle f, H_\alpha\rangle_{L^2(\gamma)}|^2$. Now we have already noted the fact that $H_{2k+1}(0) = 0$. In what follows both j_1 and j_2 should be even. Using the values of a and b we have

$$\sum_{j=(j_1,j_2)\in\mathbb{N}^2} \frac{\Gamma\big(\frac{1}{2}(2k+2|j|+n+1)+\frac{1}{2}(1+(1+s))\big)}{\Gamma\big(\frac{1}{2}(2k+2|j|+n+1)+\frac{1}{2}(1-(1+s))\big)} \bigg(H_j(0)\frac{\Gamma(a+|j|/2)\Gamma(b-a+1)}{\Gamma(b+|j|/2+1)}\bigg)^2$$
$$= \frac{\Gamma(s+1)^2}{\pi} \sum_{j\in\mathbb{N}^2} \frac{\Gamma\big(\frac{1}{2}(2k+2|j|+n+1)+\frac{1}{2}(1-(1+s))\big)}{\Gamma\big(\frac{1}{2}(2k+2|j|+n+1)+\frac{1}{2}(1+(1+s))\big)} \frac{\Gamma(j_1+1/2)\Gamma(j_2+1/2)}{\Gamma(j_1+1)\Gamma(j_2+1)}.$$

In order to simplify this further we make use of some properties of Hypergeometric functions. We start by recalling that

$$F(\delta,\beta,\eta,z) = \sum_{k=0}^{\infty} \frac{(\delta)_k(\beta)_k}{(\eta)_k k!} z^k = \frac{\Gamma(\eta)}{\Gamma(\delta)\Gamma(\beta)} \sum_{k=0}^{\infty} \frac{\Gamma(\delta+k)\Gamma(\beta+k)}{\Gamma(\eta+k)\Gamma(k+1)} z^k.$$

Here we will be using the following property proved in [Olver and Maximon 2010]:

$$\frac{\Gamma(\eta)\Gamma(\eta-\delta-\beta)}{\Gamma(\eta-\delta)\Gamma(\eta-\beta)} = F(\delta,\beta,\eta,1) = \frac{\Gamma(\eta)}{\Gamma(\delta)\Gamma(\beta)} \sum_{k=0}^{\infty} \frac{\Gamma(\delta+k)\Gamma(\beta+k)}{\Gamma(\eta+k)\Gamma(k+1)}.$$

That is,

$$\sum_{k=0}^{\infty} \frac{\Gamma(\delta+k)\Gamma(\beta+k)}{\Gamma(\eta+k)\Gamma(k+1)} = \frac{\Gamma(\beta)\Gamma(\eta-\delta-\beta)\Gamma(\delta)}{\Gamma(\eta-\delta)\Gamma(\eta-\beta)}, \quad \text{provided } \Re(\eta-\delta-\beta) > 0. \tag{5-2}$$

Taking $\delta = \frac{1}{2}(2k+2j_2+n+1-s)$, $\beta = \frac{1}{2}$ and $\eta = \frac{1}{2}(2k+2j_2+n+3+s)$ in the above formula, we have

$$\sum_{j_1=0}^{\infty} \frac{\Gamma(\delta+j_1)\Gamma(\beta+j_1)}{\Gamma(\eta+j_1)\Gamma(j_1+1)} = \frac{\Gamma(s+1/2)\Gamma(1/2)}{\Gamma(s+1)} \frac{\Gamma\big(\frac{1}{2}(2k+2j_2+n+1-s)\big)}{\Gamma\big(\frac{1}{2}(2k+2j_2+n+2+s)\big)}.$$

This gives

$$\sum_{j\in\mathbb{N}^2} \frac{\Gamma\big(\frac{1}{2}(2k+2|j|+n+1)+\frac{1}{2}(1-(1+s))\big)}{\Gamma\big(\frac{1}{2}(2k+2|j|+n+1)+\frac{1}{2}(1+(1+s))\big)} \frac{\Gamma(j_1+1/2)\Gamma(j_2+1/2)}{\Gamma(j_1+1)\Gamma(j_2+1)}$$
$$= \frac{\Gamma(s+1/2)\Gamma(1/2)}{\Gamma(s+1)} \sum_{j_2=0}^{\infty} \frac{\Gamma\big(\frac{1}{2}(2k+n+1-s)+j_2\big)\Gamma(j_2+1/2)}{\Gamma\big(\frac{1}{2}(2k+n+2+s)+j_2\big)\Gamma(j_2+1)}$$
$$= \frac{\Gamma(s+1/2)\Gamma(1/2)}{\Gamma(s+1)} \frac{\Gamma\big(\frac{1}{2}(2k+n+1-s)\big)}{\Gamma\big(\frac{1}{2}(2k+n+1+s)\big)} \frac{\Gamma(s)\Gamma(1/2)}{\Gamma(s+1/2)}$$
$$= \frac{\Gamma(1/2)^2}{s} \frac{\Gamma\big(\frac{1}{2}(2k+n+1-s)\big)}{\Gamma\big(\frac{1}{2}(2k+n+1+s)\big)}.$$

Therefore, we have

$$\begin{aligned}\|v\|^2_{(1,s)} &= c_s^2\Gamma(s+1)^2\frac{2\Gamma(1/2)^2}{\pi s}\sum_{k=0}^{\infty}2^s\frac{\Gamma\big(\frac{1}{2}(2k+n+1-s)\big)}{\Gamma\big(\frac{1}{2}(2k+n+1+s)\big)}\|Q_kf\|^2\\ &= c_{n,s}\sum_{\alpha\in\mathbb{N}^n}2^s\frac{\Gamma\big(\frac{1}{2}(2|\alpha|+n+1-s)\big)}{\Gamma\big(\frac{1}{2}(2|\alpha|+n+1+s)\big)}|\langle f, H_\alpha\rangle_{L^2(\gamma)}|^2 = c_{n,s}\|f\|^2_{(s)}.\end{aligned}$$ □

6. Hardy–Littlewood–Sobolev inequality for H_s

In this section we are interested in the Hardy–Littlewood–Sobolev inequality for the fractional powers H_s. For the Laplacian on $\mathbb{R}^n$ and the sub-Laplacian on $\mathbb{H}^n$, such inequalities with sharp constants are known. Let us recall the inequality for the sub-Laplacian $\mathcal{L}$ on $\mathbb{H}^n$. Letting $q = 2(n+1)/(n+1-s)$, the Hardy–Littlewood–Sobolev inequality for $\mathcal{L}_s$ (see [Branson et al. 2013; Frank and Lieb 2012]) reads as

$$\frac{\Gamma\big(\frac{1}{2}(1+n+s)\big)^2}{\Gamma\big(\frac{1}{2}(1+n-s)\big)^2}w_{2n+1}^{s/(n+1)}\left(\int_{\mathbb{H}^n}|g(z,w)|^q\,dz\,dw\right)^{2/q} \le \langle\mathcal{L}_sg, g\rangle. \tag{6-1}$$

We first find an integral representation of H_{-s} using the integral representation of fractional powers of the sub-Laplacian, $\mathcal{L}_{-s}$. The integral kernel of $\mathcal{L}_{-s}$ is given by $c_{n,s}|(z,t)|^{-Q+2s}$ as shown in [Roncal and Thangavelu 2016]. Here $|(z,t)| := (|z|^4+t^2)^{1/4}$ denotes the Koranyi norm on the Heisenberg group and $Q = 2n+2$ is its homogeneous dimension. We consider the Schrödinger representation π_λ of $\mathbb{H}^n$ whose action on the representation space $L^2(\mathbb{R}^n)$ is given by

$$\pi_\lambda(z,t)\phi(\xi) = e^{i\lambda t}e^{i\lambda(x\cdot\xi+x\cdot y/2)}\phi(\xi+y).$$

The Fourier transform of a function $f\in L^1(\mathbb{H}^n)$ is the operator-valued function defined on the set of all nonzero real numbers, $\mathbb{R}^*$, given by

$$\hat{f}(\lambda) = \int_{\mathbb{H}^n} f(z,t)\pi_\lambda(z,t)\,dz\,dt.$$

The action of the Fourier transform on a function of the form $\mathcal{L}$ is well known and is given by $\widehat{\mathcal{L}f}(\lambda) = \hat{f}(\lambda)H(\lambda)$, where $H(\lambda)$ is the scaled Hermite operator. In view of this, it can be easily checked that

$$d\pi_\lambda(m(\mathcal{L})) = m(H(\lambda)), \tag{6-2}$$

where $d\pi_\lambda$ stands for the derived representation corresponding to π_λ. We refer the reader to [Thangavelu 1998] for more details in this regard. Recall that the fractional power $\mathcal{L}_{-s}$ is defined as follows (see [Roncal and Thangavelu 2016]):

$$\mathcal{L}_{-s}f(z,t) := (2\pi)^{-n-1}\int_{-\infty}^{\infty}\left(\sum_{k=0}^{\infty}(2|\lambda|)^{-s}\frac{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1-s)\big)}{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1+s)\big)}f^\lambda *_\lambda \varphi_k^\lambda(z)\right)e^{-i\lambda t}|\lambda|^n\,d\lambda.$$

So we have $d\pi_\lambda(\mathcal{L}_{-s}) = H(\lambda)_{-s}$. In particular, for $\lambda = 1$, using spectral decomposition, we have

$$H_{-s} f = \sum_{k=0}^{\infty} 2^{-s} \frac{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1-s)\big)}{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1+s)\big)} P_k f.$$

Now it is not hard to see that

$$H_{-s}(f e^{-|\cdot|^2/2})(x) = e^{-|x|^2/2} \sum_{k=0}^{\infty} 2^{-s} \frac{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1-s)\big)}{\Gamma\big(\frac{1}{2}(2k+n)+\frac{1}{2}(1+s)\big)} Q_k f(x). \tag{6-3}$$

Hence from the definition of L_{-s} we have

$$H_{-s}(f e^{-|\cdot|^2/2})(x) = e^{-|x|^2/2} L_{-s} f(x). \tag{6-4}$$

In this section, we prove an analogue of (6-1) for the operator H_{-s}. We first study $L^p - L^q$ mapping properties of the operator H_{-s}.

In view of relation (6-2) we have

$$H_{-s} f(\xi) = c_{n,s} \int_{\mathbb{H}^n} |(z,t)|^{-Q+2s} \pi_1(z,t) f(\xi)\, dz\, dt.$$

Using the definition of π_1 and writing $z = x+iy$, we obtain

$$\begin{aligned} H_{-s} f(\xi) &= c_{n,s} \int_{\mathbb{H}^n} ((|x|^2+|y|^2)^2+t^2)^{-(n+1-s)/2} e^{it} e^{i(x\cdot\xi + x\cdot y/2)} f(\xi+y)\, dx\, dy\, dt \\ &= c_{n,s} \int_{\mathbb{H}^n} ((|x|^2+|\eta-\xi|^2)^2+t^2)^{-(n+1-s)/2} e^{it} e^{i(x\cdot\xi + x\cdot\eta)/2} f(\eta)\, dx\, d\eta\, dt \\ &= \int_{\mathbb{R}^n} K_H^s(\xi,\eta) f(\eta)\, d\eta, \end{aligned}$$

where the kernel K_H^s is defined by

$$K_H^s(\xi,\eta) = c_{n,s} \int_{\mathbb{R}^n\times\mathbb{R}} ((|x|^2+|\eta-\xi|^2)^2+t^2)^{-(n+1-s)/2} e^{it} e^{i(x\cdot\xi+x\cdot\eta)/2}\, dx\, dt. \tag{6-5}$$

Taking the modulus and then a change of variables leads to

$$|K_H^s(\xi,\eta)| \le c_{n,s} \int_{\mathbb{R}^n} (|x|^2+|\eta-\xi|^2)^{-n+s}\, dx.$$

Now again a change of variable $x \to x|\xi-\eta|$ yields

$$|K_H^s(\xi,\eta)| \le C_{n,s} |\xi-\eta|^{-n+2s}. \tag{6-6}$$

It is a routine matter to check the following $L^p - L^q$-boundedness property; see e.g., [Grafakos 2009, Theorem 6.1.3]. In fact, for $1 < p < q < \infty$ with $1/p - 1/q = 2s/n$, we get

$$\|H_{-s} f\|_{L^q} \le C_{n,s}(p) \|f\|_{L^p}. \tag{6-7}$$

Nevertheless, in the following theorem, we obtain a better estimate for the kernel, improving the $L^p - L^q$ estimates mentioned above.

Theorem 6.1. *For any $1 \le p \le q < \infty$ with $1/p - 1/q \le 1$, there exists a constant $C_{n,s}(p)$ such that for all $f \in L^p(\mathbb{R}^n)$, the inequality $\|H_{-s} f\|_{L^q} \le C_{n,s}(p)\|f\|_{L^p}$ holds.*

Proof. In view of the formula stated in (4-14), from (6-5) we have

$$K_H^s(\xi,\eta) := 2c_{n,s}\frac{\sqrt{\pi}\,2^{-(n/2+1+s)/2}}{\Gamma\big(\frac{1}{2}(n+1-s)\big)}\int_{\mathbb{R}^n}(|x|^2+|\eta-\xi|^2)^{-(n+1-s)/2}K_{-(n+1-s)/2}(|x|^2+|\eta-\xi|^2)e^{ix\cdot(\eta+\xi)/2}\,dx.$$

Now we use the integral representation of K_ν to simplify the above integral giving the kernel as

$$K_\nu(z) = 2^{-\nu-1}z^\nu\int_0^\infty e^{-t-z^2/(4t)}t^{-\nu-1}\,dt. \tag{6-8}$$

A simple change of variables shows that

$$z^\nu K_\nu(z) = 2^{\nu-1}\int_0^\infty e^{-t-z^2/(4t)}t^{\nu-1}\,dt = z^\nu K_{-\nu}(z).$$

Thus

$$\begin{aligned}(|x|^2+|\eta-\xi|^2)^{-(n+1-s)/2}&K_{-(n+1-s)/2}(|x|^2+|\eta-\xi|^2)\\ &= (|x|^2+|\eta-\xi|^2)^{-(n+1-s)/2}K_{(n+1-s)/2}(|x|^2+|\eta-\xi|^2),\end{aligned}$$

leading to the formula

$$(|x|^2+|\eta-\xi|^2)^{-(n+1-s)/2}K_{-(n+1-s)/2}(|x|^2+|\eta-\xi|^2) = 2^{-\nu-1}\int_0^\infty e^{-t-z^2/(4t)}t^{-\nu-1}\,dt,$$

where $\nu = \frac{1}{2}(n+1-s)$ and $z = |x|^2+|\xi-\eta|^2$. Writing $a := \frac{1}{2}(\xi+\eta)$, we estimate the integral

$$\int_{\mathbb{R}^n} e^{-(|x|^2+r^2)^2/(4t)}e^{ix\cdot a}\,dx,$$

where we have let $r = |\xi-\eta|$. First note that

$$\int_{\mathbb{R}^n} e^{-(|x|^2+r^2)^2/(4t)}e^{ix\cdot a}\,dx = e^{-r^4/(4t)}\int_{\mathbb{R}^n} e^{ix\cdot a}e^{-2r^2|x|^2/(4t)}e^{-|x|^4/(4t)}\,dx.$$

Let φ stand for the Fourier transform of the function $e^{-|x|^4/4}$. So the above integral is bounded by

$$e^{-r^4/(4t)}\left(\frac{t}{r^2}\right)^{n/2}t^{n/4}\int_{\mathbb{R}^n}\varphi(t^{1/4}(a-y))e^{-t|y|^2/(2r^2)}\,dy,$$

which is bounded by (after making a change of variables and using $|\varphi(\xi)| \le C$)

$$e^{-r^4/(4t)}t^{n/4},$$

and $K_H^s(\xi,\eta)$ is bounded by

$$\int_0^\infty e^{-t}e^{-r^4/(4t)}t^{-(n+2-2s)/4-1}\,dt = r^{-(n+2-2s)/2}K_{(n+2-2s)/4}(r^2).$$

Finally we have

$$|K_H^s(\xi,\eta)| \le C|\xi-\eta|^{-(n+2-2s)/2}K_{(n+2-2s)/4}(|\xi-\eta|^2) =: G(\xi-\eta). \tag{6-9}$$

Now we see that

$$|H_{-s} f(\xi)| \leq C|f| * G(\xi), \quad \forall \xi \in \mathbb{R}^n. \tag{6-10}$$

Now note that for $r \geq 1$, integrating in polar coordinates, we have

$$\int_{\mathbb{R}^n} G(x)^r \, dx = c_n \int_0^\infty (t^{-(n+2-2s)/2} K_{(n+2-2s)/4}(t^2))^r t^{n-1} \, dt.$$

Using the facts that $K_\nu(z) \sim z^{-1/2} e^{-z}$ for large z and near the origin $z^{-\nu} K_\nu(z)$ is bounded, we conclude that the above integral is finite. Now in view of Young's inequality we have

$$\||f| * G\|_q \leq \|f\|_p \|G\|_r, \quad \text{where } \frac{1}{q} + 1 = \frac{1}{p} + \frac{1}{r}. \tag{6-11}$$

But this is true for any $r \geq 1$. Hence we are done. □

As a corollary to Theorem 6.1 we have the following analogue of (6-1).

Corollary 6.2. *For $q = 2n/(n-s)$, $0 < s < n$, we have the inequality*

$$C_{n,s} \left(\int_{\mathbb{R}^n} |f(x)|^q \, dx \right)^{2/q} \leq \langle H_s f, f \rangle, \tag{6-12}$$

where $C_{n,s}$ is some constant depending only on n and s.

Proof. Replacing s by $\frac{1}{2}s$ and putting $p = 2$ in the above theorem, we have

$$\|H_{-s/2} f\|_q^2 \leq c_{n,s} \|f\|_2^2, \tag{6-13}$$

where $q = 2n/(n-s)$. Now in the above inequality substituting f by $H_{s/2} f$ we have

$$\left(\int_{\mathbb{R}^n} |f(x)|^q \, dx \right)^{2/q} \leq c_{n,s} \langle H_{s/2} f, H_{s/2} f \rangle.$$

But in view of Stirling's formula for the gamma function we know that $H_{s/2}^2$ and H_s differ by a bounded operator on $L^2(\mathbb{R}^n)$. Hence the result follows. □

Corollary 6.3 (Hardy's inequality for H_s). *Let $0 < s < 1$. Assume that $f \in L^2(\mathbb{R}^n)$ such that $H_s f \in L^2(\mathbb{R}^n)$. Then we have*

$$\langle H_s f, f \rangle_{L^2(\mathbb{R}^n)} \geq c_{n,s} \int_{\mathbb{R}^n} \frac{f(x)^2}{(1+|x|^2)^s} \, dx.$$

Proof. Given $f \in L^2(\mathbb{R}^n)$, in view of Holder's inequality we have

$$\int_{\mathbb{R}^n} \frac{f(x)^2}{(1+|x|^2)^s} \, dx \leq A(n,s) \left(\int_{\mathbb{R}^n} |f(x)|^q \, dx \right)^{2/q}, \tag{6-14}$$

where

$$q = \frac{2n}{n-s}, \quad A(n,s) := \left(\int_{\mathbb{R}^n} (1+|x|^2)^{-sq'} \, dx \right)^{1/q'} \quad \text{and} \quad \frac{1}{q'} = 1 - \frac{2n-2s}{2n} = \frac{s}{n}.$$

Hence the result follows from the previous corollary. □

As a consequence of this we have a version of Hardy's inequality for L_s:

Corollary 6.4. *Let* $0 < s < 1$. *Assume that* $f \in L^2(\gamma)$ *such that* $L_s f \in L^2(\gamma)$. *Then we have*

$$\langle L_s f, f\rangle_{L^2(\gamma)} \geq c_{n,s} \int_{\mathbb{R}^n} \frac{f(x)^2}{(1+|x|^2)^s}\, d\gamma(x).$$

Proof. Let $f \in L^2(\gamma)$. It is easy to see that $g(x) := f(x)e^{-|x|^2/2} \in L^2(\mathbb{R}^n)$. By Corollary 6.3 we have

$$\langle H_s g, g\rangle_{L^2(\mathbb{R}^n)} \geq c_{n,s} \int_{\mathbb{R}^n} \frac{g(x)^2}{(1+|x|^2)^s}\, dx.$$

Also from the spectral decomposition we see that

$$H_s g(x) = H_s(f e^{-|\cdot|^2/2})(x) = e^{-|x|^2/2} L_s f(x),$$

which gives $\langle H_s g, g\rangle_{L^2(\mathbb{R}^n)} = \langle L_s f, f\rangle_{L^2(\gamma)}$. Hence the result follows. □

Remark. Frank and Lieb proved in [Frank et al. 2008] that the constant appearing in the left-hand side of the Hardy–Littlewood–Sobolev inequality (6-1) for the sub-Laplacian on the Heisenberg group is sharp. It would be interesting to see the sharp constant in the analogous inequality (6-12), which we have proved for the Hermite operator.

Acknowledgements

The authors are grateful to the anonymous reviewer for the valuable comments and suggestions. Ganguly is supported by an Integrated Ph.D. scholarship from the Indian Institute of Science. Manna is thankful for financial support from DST-INSPIRE [DST/INSPIRE/04/2019/001914] and R(IA)CVR-PDF/2020/224. Thangavelu is supported by a J. C. Bose Fellowship from the Department of Science and Technology, Government of India.

References

[Abramowitz and Stegun 1964] M. Abramowitz and I. A. Stegun (editors), *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Nat. Bureau Standards Appl. Math. Ser. **55**, US Govt. Printing Office, Washington, DC, 1964. Reprinted by Dover, New York, 1974. MR Zbl

[Ahlfors 1953] L. V. Ahlfors, *Complex analysis*: *an introduction to the theory of analytic functions of one complex variable*, McGraw-Hill, New York, 1953. MR Zbl

[Beckner 2012] W. Beckner, "Pitt's inequality and the fractional Laplacian: sharp error estimates", *Forum Math.* **24**:1 (2012), 177–209. MR Zbl

[Boggarapu et al. 2019] P. Boggarapu, L. Roncal, and S. Thangavelu, "On extension problem, trace Hardy and Hardy's inequalities for some fractional Laplacians", *Commun. Pure Appl. Anal.* **18**:5 (2019), 2575–2605. MR Zbl

[Branson et al. 2013] T. P. Branson, L. Fontana, and C. Morpurgo, "Moser–Trudinger and Beckner–Onofri's inequalities on the CR sphere", *Ann. of Math.* (2) **177**:1 (2013), 1–52. MR Zbl

[Caffarelli and Silvestre 2007] L. Caffarelli and L. Silvestre, "An extension problem related to the fractional Laplacian", *Comm. Partial Differential Equations* **32**:7-9 (2007), 1245–1260. MR Zbl

[Ciaurri et al. 2018] Ó. Ciaurri, L. Roncal, and S. Thangavelu, "Hardy-type inequalities for fractional powers of the Dunkl–Hermite operator", *Proc. Edinb. Math. Soc.* (2) **61**:2 (2018), 513–544. MR Zbl

[Cowling and Haagerup 1989] M. Cowling and U. Haagerup, "Completely bounded multipliers of the Fourier algebra of a simple Lie group of real rank one", *Invent. Math.* **96**:3 (1989), 507–549. MR Zbl

[Folland 1989] G. B. Folland, *Harmonic analysis in phase space*, Ann. of Math. Stud. **122**, Princeton Univ. Press, 1989. MR Zbl

[Frank and Lieb 2012] R. L. Frank and E. H. Lieb, "Sharp constants in several inequalities on the Heisenberg group", *Ann. of Math.* (2) **176**:1 (2012), 349–381. MR Zbl

[Frank et al. 2008] R. L. Frank, E. H. Lieb, and R. Seiringer, "Hardy–Lieb–Thirring inequalities for fractional Schrödinger operators", *J. Amer. Math. Soc.* **21**:4 (2008), 925–950. MR Zbl

[Frank et al. 2015] R. L. Frank, M. d. M. González, D. D. Monticelli, and J. Tan, "An extension problem for the CR fractional Laplacian", *Adv. Math.* **270** (2015), 97–137. MR Zbl

[Garofalo and Tralli 2021] N. Garofalo and G. Tralli, "Feeling the heat in a group of Heisenberg type", *Adv. Math.* **381** (2021), art. id. 107635. MR Zbl

[Garofalo and Tralli 2023] N. Garofalo and G. Tralli, "A heat equation approach to intertwining", *J. Anal. Math.* **149**:1 (2023), 113–134. MR

[Grafakos 2009] L. Grafakos, *Modern Fourier analysis*, 2nd ed., Grad. Texts in Math. **250**, Springer, 2009. MR Zbl

[Hardy 1919] G. H. Hardy, "Notes on some points in the integral calculus, LI: On Hilbert's double series theorem and some connected theorems concerning the convergence of infinite series and integrals", *Messenger Math.* **48** (1919), 107–112. Zbl

[Hardy 1920] G. H. Hardy, "Note on a theorem of Hilbert", *Math. Z.* **6**:3-4 (1920), 314–317. MR Zbl

[Kufner et al. 2007] A. Kufner, L. Maligranda, and L.-E. Persson, *The Hardy inequality*: *about its history and some related results*, Vydavatelský Servis, Pilsen, Czech Republic, 2007. MR Zbl

[Möllers et al. 2016] J. Möllers, B. Ørsted, and G. Zhang, "On boundary value problems for some conformally invariant differential operators", *Comm. Partial Differential Equations* **41**:4 (2016), 609–643. MR Zbl

[Olver and Maximon 2010] F. W. J. Olver and L. C. Maximon, "Bessel functions", pp. 215–286 in *NIST handbook of mathematical functions*, edited by F. W. J. Olver et al., US Dept. Commerce, Washington, DC, 2010. MR

[Prudnikov et al. 1986] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and series*, *I*: *Elementary functions*, Gordon & Breach, New York, 1986. MR Zbl

[Roncal and Thangavelu 2016] L. Roncal and S. Thangavelu, "Hardy's inequality for fractional powers of the sublaplacian on the Heisenberg group", *Adv. Math.* **302** (2016), 106–158. MR Zbl

[Roncal and Thangavelu 2020a] L. Roncal and S. Thangavelu, "An extension problem and trace Hardy inequality for the sub-Laplacian on H-type groups", *Int. Math. Res. Not.* **2020**:14 (2020), 4238–4294. MR Zbl

[Roncal and Thangavelu 2020b] L. Roncal and S. Thangavelu, "Holomorphic extensions of eigenfunctions on NA groups", preprint, 2020. arXiv 2005.09894

[Stinga and Torrea 2010] P. R. Stinga and J. L. Torrea, "Extension problem and Harnack's inequality for some fractional operators", *Comm. Partial Differential Equations* **35**:11 (2010), 2092–2122. MR Zbl

[Szegő 1967] G. Szegő, *Orthogonal polynomials*, 3rd ed., Amer. Math. Soc. Colloq. Publ. **23**, Amer. Math. Soc., Providence, RI, 1967. MR

[Thangavelu 1993] S. Thangavelu, *Lectures on Hermite and Laguerre expansions*, Math. Notes **42**, Princeton Univ. Press, 1993. MR Zbl

[Thangavelu 1998] S. Thangavelu, *Harmonic analysis on the Heisenberg group*, Progr. Math. **159**, Birkhäuser, Boston, 1998. MR Zbl

[Thangavelu 2008] S. Thangavelu, "An analogue of Gutzmer's formula for Hermite expansions", *Studia Math.* **185**:3 (2008), 279–290. MR Zbl

[Thangavelu 2010] S. Thangavelu, "On the unreasonable effectiveness of Gutzmer's formula", pp. 199–217 in *Harmonic analysis and partial differential equations* (El Escorial, Spain, 2008), edited by P. Cifuentes et al., Contemp. Math. **505**, Amer. Math. Soc., Providence, RI, 2010. MR Zbl

[Urbina-Romero 2019] W. Urbina-Romero, *Gaussian harmonic analysis*, Springer, 2019. MR Zbl

[Yafaev 1999] D. Yafaev, “Sharp constants in the Hardy–Rellich inequalities”, *J. Funct. Anal.* **168**:1 (1999), 121–144. MR Zbl

Received 17 Dec 2020. Revised 19 Sep 2021. Accepted 28 Oct 2021.

PRITAM GANGULY: pritam1995.pg@gmail.com, pritamg@math.upb.de
Current address: *Institut für Mathematik, Universität Paderborn, Paderborn, Germany*
Department of Mathematics, Indian Institute of Science, Bangalore, India

RAMESH MANNA: rameshmanna@niser.ac.in
School of Mathematical Sciences, National Institute of Science Education and Research, Bhubaneswar, an OCC of Homi Bhabha National Institute, Jatni, 752050, India

SUNDARAM THANGAVELU: veluma@iisc.ac.in
Department of Mathematics, Indian Institute of Science, Bangalore, India

ANALYSIS AND PDE
Vol. 16 (2023), No. 5, pp. 1245–1270
DOI: 10.2140/apde.2023.16.1245

ON THE WELL-POSEDNESS PROBLEM FOR THE DERIVATIVE NONLINEAR SCHRÖDINGER EQUATION

ROWAN KILLIP, MARIA NTEKOUME AND MONICA VIŞAN

We consider the derivative nonlinear Schrödinger equation in one space dimension, posed both on the line and on the circle. This model is known to be completely integrable and L^2-critical with respect to scaling. We first discuss whether ensembles of orbits with L^2-equicontinuous initial data remain equicontinuous under evolution. We prove that this is true under the restriction $M(q) = \int |q|^2 < 4\pi$. We conjecture that this restriction is unnecessary. Further, we prove that the problem is globally well posed for initial data in $H^{1/6}$ under the same restriction on M. Moreover, we show that this restriction would be removed by a successful resolution of our equicontinuity conjecture.

1. Introduction

The derivative nonlinear Schrödinger equation

$$iq_t + q'' + i(|q|^2 q)' = 0 \tag{DNLS}$$

describes the evolution of a complex-valued field q defined either on the line $\mathbb{R}$ or the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. This equation was introduced as an effective model in magnetohydrodynamics; see [Ichikawa and Watanabe 1977; Mio et al. 1976; Mjølhus 1976]. It was soon shown to be completely integrable [Kaup and Newell 1978] and has received enduring attention since that time.

As we shall document more fully below, well-posedness questions for (DNLS), particularly *global* well-posedness, have been particularly stubborn. Local well-posedness is already very challenging: the nonlinearity contains a full derivative, like KdV or mKdV, while the linear part gives only Schrödinger-like smoothing.

The task of converting local into global well-posedness is typically a matter of exploiting conservation laws. As a completely integrable system, (DNLS) has an infinite family of conserved quantities. The first three are as follows:

$$M(q) = \int |q(x)|^2 \, dx, \tag{1-1}$$

$$H(q) = -\frac{1}{2} \int i(q\bar{q}' - \bar{q}q') + |q|^4 \, dx, \tag{1-2}$$

$$H_2(q) = \int |q'|^2 + \frac{3}{4} i |q|^2 (q\bar{q}' - \bar{q}q') + \frac{1}{2}|q|^6 \, dx. \tag{1-3}$$

Killip was supported by NSF grant DMS-1856755 and Vişan by grant DMS-1763074.
MSC2020: 35Q55.
Keywords: derivative nonlinear Schrödinger equation, well-posedness.

The striking fact about (DNLS) is that, with the exception of $M(q)$, none of the Hamiltonians in the hierarchy are coercive. Indeed, algebraic solitons have $M = 4\pi$ but all other Hamiltonians are identically zero. Applying the scaling symmetry

$$q(t,x) \mapsto \sqrt{\lambda}\, q(\lambda^2 t, \lambda x) \tag{1-4}$$

to an algebraic soliton yields a one-parameter family of solutions with identical values for *all* the conserved quantities. However, this family is unbounded in H^s for every $s > 0$.

The quantity $H(q)$ serves as the Hamiltonian for (DNLS) with respect to the Poisson structure

$$\{F, G\} = \int \frac{\delta F}{\delta q}\left(\frac{\delta G}{\delta \bar{q}}\right)' + \frac{\delta F}{\delta \bar{q}}\left(\frac{\delta G}{\delta q}\right)' dx, \tag{1-5}$$

while $M(q)$ generates translations, albeit at speed 2. Although the momentum is given by $\frac{1}{2}M(q)$, our definition of M leads to a more seamless connection to the existing literature.

Given that $M(q)$ is invariant under both (DNLS) and the scaling (1-4), it is natural to ask whether or not (DNLS) is well posed in L^2. This is not known. Indeed, the existing local well-posedness theory requires H^s initial data with $s \geq \frac{1}{2}$. (We will make some further progress on this question in this paper.) It is important to recognize that because $M(q)$ is scaling critical, the mere fact that it forms a coercive conservation law would not suffice to render local well-posedness in L^2 automatically global. One must fear the solution concentrates at one (or more) points in space, a scenario known as type-II blowup. We do not believe this happens.

Conjecture 1.1. *For any $Q \subseteq \mathcal{S}$ that is L^2-bounded and equicontinuous, the totality of states reached by* (DNLS) *orbits originating from Q, that is*

$$Q_* = \{e^{tJ\nabla H} q : q \in Q \text{ and } t \in \mathbb{R}\}, \tag{1-6}$$

is also L^2-equicontinuous.

Here $\mathcal{S}$ denotes Schwartz class in the line case and C^∞ on the torus. In the line case, recent works (discussed below) guarantee that all such initial data lead to global Schwartz solutions. The analogous claim is unknown on the torus, though we believe it to be true. Nevertheless, one can still ask if equicontinuity holds for as long as the orbits do exist. By the arguments presented in this paper, solutions cannot break down without losing equicontinuity. Therefore, a positive resolution of the conjecture for such partial solutions would already guarantee that they are global and so settle the conjecture in its entirety; see Corollary 4.2.

We phrased the conjecture in terms of $\mathcal{S}$ initial data because it is a class that is dense in all relevant spaces. It also serves to emphasize that the central question to be addressed is not inherently tied to low regularity.

Equicontinuity in L^2 is most easily understood via Fourier transformation: it means that $|\hat{q}|^2$ forms a tight family of measures. Notice that, in view of the uncertainty principle, concentration on the physical side must be accompanied by a loss of tightness on the Fourier side.

In setting this conjecture, we have in mind four principal reasons: (1) It is challenging, yet recent developments give us hope for a successful resolution. (2) It encapsulates a single essential obstacle, namely, understanding conservation laws for (DNLS). (3) A proof of this conjecture would have significant consequences for the well-posedness problem. Indeed, such equicontinuity results form an essential part of a recent program developed in [Bringmann et al. 2021; Harrop-Griffiths et al. 2020; Killip and Vişan 2019] that has proved successful in obtaining optimal well-posedness results for completely integrable PDE. (4) We are able to verify that it is true in the regime $M(q) < 4\pi$; see Theorem 1.3 below.

Given the nature of completely integrable systems, it is natural to imagine that an equicontinuity conjecture of the same form holds for all other PDE in the (DNLS) hierarchy. Indeed, we truly believe that this is so and will shortly formulate just such a conjecture. However, the particular claim that we believe will be of greatest use in understanding the hierarchy is best expressed through the perturbation determinant. Let us turn our attention now to presenting this object, beginning with the requisite background.

The Lax pair introduced by Kaup and Newell [1978] for (DNLS) employs

$$L_{\mathrm{KN}} = \begin{bmatrix} -i\lambda^2 - \partial & \lambda q \\ -\lambda\bar{q} & i\lambda^2 - \partial \end{bmatrix}.$$

For what follows, it will be convenient to make some cosmetic changes to this choice. Specifically, we set $\lambda = e^{i\pi/4}\sqrt{\kappa}$ with $\kappa \geq 1$ and replace $e^{i\pi/4}q \mapsto q$. This yields

$$L(\kappa) := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \kappa - \partial & \sqrt{\kappa}\, q \\ i\sqrt{\kappa}\,\bar{q} & \kappa + \partial \end{bmatrix} \quad \text{and, for } q \equiv 0, \quad L_0(\kappa) := \begin{bmatrix} \kappa - \partial & 0 \\ 0 & -(\kappa + \partial) \end{bmatrix}.$$

These modifications maintain the crucial property that for smooth functions,

$$q(t) \text{ solves (DNLS)} \quad \Longleftrightarrow \quad \frac{d}{dt} L(t;\kappa) = [P(t;\kappa),\ L(t;\kappa)],$$

where

$$P(\kappa) = \begin{bmatrix} 2i\kappa^2 - \kappa|q|^2 & 2i\kappa^{3/2}q - \kappa^{1/2}|q|^2 q + i\kappa^{1/2}q' \\ 2\kappa^{3/2}\bar{q} + i\kappa^{1/2}|q|^2\bar{q} - \kappa^{1/2}\bar{q}' & -2i\kappa^2 + \kappa|q|^2 \end{bmatrix}.$$

This guarantees that the Lax operators L at different times are conjugate, at least formally. This in turn suggests that the perturbation determinant $\det[L_0^{-1}(\kappa)L(\kappa)]$ should be well defined and conserved by the flow.

To make this precise, it is convenient for us to mimic the analysis of the AKNS-ZS system employed in [Killip et al. 2018]: Let us first define $(\kappa \pm \partial)^{-1/2}$ as the Fourier multipliers $(\kappa \pm i\xi)^{-1/2}$, where the complex square root is determined by $\sqrt{\kappa} > 0$ and continuity. We then define

$$\Lambda(q) := (\kappa - \partial)^{-1/2} q (\kappa + \partial)^{-1/2} \quad \text{and} \quad \Gamma(q) := (\kappa + \partial)^{-1/2} \bar{q} (\kappa - \partial)^{-1/2}, \tag{1-7}$$

which are Hilbert–Schmidt operators for $q \in L^2$; see Lemma 2.1. Thus

$$a(\kappa; q) = \det[1 - i\kappa\Lambda\Gamma] \tag{1-8}$$

is well defined for $q \in L^2$ (and extends holomorphically to all $\operatorname{Re}\kappa > 0$); moreover, for $q \in \mathcal{S}$ it agrees with the formal notion of the perturbation determinant.

While $a(\kappa)$ does encode all the Hamiltonians of the (DNLS) hierarchy, this is more easily seen through its logarithm,

$$\alpha(\kappa; q) := -\log[a(\kappa; q)] = \sum_{\ell \geq 1} \frac{1}{\ell} \operatorname{tr}\{(i\kappa\Lambda\Gamma)^{\ell}\}, \tag{1-9}$$

which serves as a generating function for these conservation laws. Due to the possibility of $a(\kappa)$ vanishing, $\alpha(\kappa)$ may not be defined for all $\kappa \geq 1$. Nevertheless, the series in (1-9) does converge for fixed $q \in L^2$ and κ sufficiently large; see Proposition 2.6.

We have not yet addressed the conservation of $a(\kappa; q)$ under the (DNLS) flow. In the line case, this could be effected by demonstrating that $a(\kappa; q)$ coincides with the reciprocal of the transmission coefficient and then appealing to the inverse scattering theory. However, two direct proofs have appeared recently in the literature: Klaus and Schippa [2022] argued by differentiating the series (following a model introduced in [Killip et al. 2018]), while Tang and Xu [2021] developed a microscopic representation of this conservation law (in the style of [Harrop-Griffiths et al. 2020]). While these papers impose a small $M(q)$ requirement, this is solely to guarantee the convergence of the series (1-9). This issue is remedied by our Proposition 2.6.

To state the grand version of Conjecture 1.1, covering a wide range of commuting flows, let us first introduce a replacement for the set Q_* defined in (1-6). Given $q \in \mathcal{S}$, we first define

$$C_q = \{\tilde{q} \in \mathcal{S} : a(\kappa; \tilde{q}) = a(\kappa; q) \text{ for all } \kappa > 0\} \tag{1-10}$$

and write C_q^0 for the connected component (in the L^2 topology) of C_q containing q. Finally, given a set $Q \subseteq \mathcal{S}$, we define

$$Q_{**} = \bigcup_{q \in Q} C_q^0. \tag{1-11}$$

Conjecture 1.2. *If $Q \subseteq \mathcal{S}$ is L^2-bounded and equicontinuous, then so too is the set Q_{**} defined in* (1-11).

We have several motivations in choosing connected components when defining Q_{**}. This formulation of the conjecture retains a vestige of the behavior of orbits, while emphasizing that this is a question about conservation laws and is ultimately independent of the well-posedness of any flow. Note also that while the zero solution and the family of algebraic solitons all share $a(\kappa) \equiv 1$, they are not in the same connected component under the (DNLS) hierarchy.

Our most compelling evidence in favor of these two conjectures is that both hold in the regime where $M(q) < 4\pi$.

Theorem 1.3. *Let $Q \subseteq \mathcal{S}$ be an L^2-equicontinuous set satisfying*

$$\sup\{\|q\|_{L^2}^2 : q \in Q\} < 4\pi. \tag{1-12}$$

Then the set Q_{**} defined in (1-11) *is L^2-bounded and equicontinuous.*

The significance of 4π is this: It is the value of M at which the polynomial conservation laws lose their efficacy. It is also the value of M for the algebraic soliton, which is maximal among all solitary wave solutions. Unlike mass-critical NLS, (DNLS) admits solitons of arbitrarily small L^2-norm, and consequently, there is no notion of a scattering threshold.

The proof of Theorem 1.3, which will be given in Section 3, is both short and simple. Indeed, the hypothesis (1-12) even allows us to forgo the restriction to connected components.

It has been observed before that $\operatorname{tr}(i\kappa\Lambda\Gamma)$ may be used to understand how the L^2-norm of q is distributed across frequencies (compare Lemma 2.2). The key observation that allows us to reach all the way to 4π (as opposed to mere smallness, compare [Klaus and Schippa 2022; Tang and Xu 2021]) is the manner in which we handle the remainder, specifically, the observation that the remainder may be summed in κ for any $q \in L^2$; see (3-8).

While the 4π restriction is crucial to our proof of Theorem 1.3, it does not play any role in our subsequent analysis of the consequences of such equicontinuity. For this reason, we introduce a general threshold M_*.

Definition 1.4. Let M_* denote the maximal constant such that for any L^2-equicontinuous set $Q \subseteq \mathcal{S}$ satisfying

$$\sup\{\|q\|_{L^2}^2 : q \in Q\} < M_*, \tag{1-13}$$

the set defined in (1-11) is L^2-equicontinuous.

Evidently, Theorem 1.3 shows that $M_* \geq 4\pi$ and we conjecture that $M_* = \infty$. Our primary contribution to the well-posedness problem is low-regularity well-posedness below the M_* threshold.

Theorem 1.5. *Fix $\frac{1}{6} \leq s < \frac{1}{2}$. The* (DNLS) *evolution is globally well posed, both on the line and on the circle, in the space*

$$B^s_{M_*} = \{q \in H^s : \|q\|_{L^2}^2 < M_*\} \tag{1-14}$$

endowed with the H^s topology.

A natural prerequisite for proving this theorem is a priori H^s bounds. In Section 4, we show how such bounds follow from L^2-equicontinuity; see Theorem 4.3.

To prove Theorem 1.5 we employ the method of commuting flows introduced in [Killip and Vişan 2019]. In that paper, the method was used to prove well-posedness of the Korteweg–de Vries equation. It has also been adapted and extended to treat the well-posedness problem for other completely integrable PDE [Bringmann et al. 2021; Harrop-Griffiths et al. 2020], to prove symplectic non-squeezing [Ntekoume 2022], and to construct dynamics for KdV in thermal equilibrium [Killip et al. 2020].

In contrast to those papers, we do not employ a change of unknown; this simplifies some of the analysis. On the other hand, new difficulties attend the construction of regularized flows: Because they are rooted in $\alpha(\kappa; q)$, the regularized Hamiltonians $H_\kappa(q)$ cannot be defined throughout $B^0_{M_*}$ for any single value of κ. Instead, we need to use an exhaustion by equicontinuous subsets. Ultimately, these problems originate in the L^2-criticality of the problem. Nevertheless, we will be able to prove that the regularized flows admit a satisfactory notion of well-posedness all the way down to L^2! The $s \geq \frac{1}{6}$ restriction arises later when we show that the regularized flows converge to the full (DNLS) evolution.

At this moment we do not know whether $s = \frac{1}{6}$ is sharp in either geometry or indeed, whether the threshold regularity will differ between the line and the circle. Moreover, we do not know of any results (in either geometry) that would preclude well-posedness all the way down to the scaling critical space L^2. On the other hand, the self-similar solutions constructed in [Fujiwara et al. 2020] (see also [Kitaev 1985]) show that smooth solutions can break-down in a dramatic way if one permits mere weak-L^2 decay at spatial infinity.

The restriction $s < \frac{1}{2}$ in Theorem 1.5 does not represent a meaningful breakdown of our methods. However, treating larger values would require additional arguments. This seems unwarranted given that a great deal is already known about H^s-solutions for $s \geq \frac{1}{2}$, as we shall now discuss.

Local well-posedness in H^s for $s > \frac{3}{2}$ was proved by Tsutsumi and Fukuda [1980; 1981]. This was extended to $s \geq \frac{1}{2}$ by Takaoka [1999] for (DNLS) posed on the line and by Herr [2006] for the periodic problem. The endpoint $s = \frac{1}{2}$ is significant: for lesser s, the data-to-solution map can no longer be uniformly continuous on bounded sets; see [Biagioni and Linares 2001; Takaoka 1999].

Global well-posedness in $H^1(\mathbb{R})$ for initial data satisfying $M(q) < 2\pi$ was obtained by Hayashi and Ozawa [1992]. This result was extended first to $s > \frac{2}{3}$ and then to $s > \frac{1}{2}$ by Colliander, Keel, Staffilani, Takaoka, and Tao [Colliander et al. 2001; 2002], under the same L^2 restriction. See [Miao et al. 2011] for a refinement of these arguments to handle the endpoint case $s = \frac{1}{2}$, as well as [Takaoka 2001] for earlier efforts in this direction.

Hayashi and Ozawa [1992] also proved that solutions with initial data in $\mathcal{S}$ remain in $\mathcal{S}$ for as long as they remain bounded in H^1.

Wu [2015] proved global well-posedness in $H^1(\mathbb{R})$ for initial data satisfying $M(q) < 4\pi$; see also his earlier work [Wu 2013] which first overcame the 2π barrier. An alternate variational proof was given in [Fukaya et al. 2017], which also constructed global solutions for highly modulated initial data of arbitrary L^2 size. The result in [Wu 2015] was extended to the periodic setting in [Mosincat and Oh 2015]. Finally, the argument in [Colliander et al. 2002] was further advanced in [Guo and Wu 2017; Mosincat 2017] to treat the endpoint case $s = \frac{1}{2}$ and $M(q) < 4\pi$; see also [Win 2010] for earlier work in the periodic setting.

We note that the results of this paper provide an alternate proof of the main results in [Mosincat and Oh 2015; Wu 2015]; see Corollary 4.2. In particular, Proposition 4.1 shows that H^1 bounds follow from Theorem 1.3.

The well-posedness of (DNLS) has also been investigated in Fourier–Lebesgue spaces; [Deng et al. 2021; Grünrock 2005; Grünrock and Herr 2008]. This allowed the authors to obtain a uniformly continuous data-to-solution map in spaces that are closer to the critical scaling; recall that this property breaks down in H^s spaces when $s < \frac{1}{2}$. An almost sure global well-posedness result for randomized initial data was proved in [Nahmod et al. 2012].

As a completely integrable PDE, (DNLS) is also amenable to inverse scattering techniques. Building on the pioneering work of Liu [2017], global well-posedness and asymptotic analysis of soliton-free solutions in $H^{2,2}(\mathbb{R}) = \{f \in H^2(\mathbb{R}) : x^2 f \in L^2(\mathbb{R})\}$ were addressed in [Liu et al. 2016; 2018].

Global well-posedness for *all* $H^{2,2}(\mathbb{R})$ initial data was proved by Jenkins, Liu, Perry, and Sulem in [Jenkins et al. 2020b]. This work builds on the authors' prior successes in [Jenkins et al. 2018b]. These authors also proved a soliton resolution result [Jenkins et al. 2018a] for generic data in $H^{2,2}(\mathbb{R})$. See also their excellent review article [Jenkins et al. 2020a].

The inverse scattering approach was also applied by Pelinovsky and Shimabukuro [2018] to prove global well-posedness in $H^{1,1}(\mathbb{R}) \cap H^2(\mathbb{R})$ for soliton-free solutions and then in joint work with Saalmann [Pelinovsky et al. 2017] for data giving rise to finitely many solitons; see also [Saalmann 2017].

Recently there has been a surge of activity on the well-posedness problem for (DNLS). We first note the paper [Klaus and Schippa 2022], which showed a priori H^s bounds, $0 < s < \frac{1}{2}$, for solutions with $M(q)$ small. The smallness assumption allows them to guarantee that the series (1-9) converges rapidly for κ large, and so the series can be conflated with its first term. The paper [Tang and Xu 2021] presents a microscopic representation of the conservation of $\alpha(\kappa; q)$. In [Bahouri and Perelman 2022], the authors achieve the major breakthrough of proving that for *every* initial datum in $H^{1/2}(\mathbb{R})$, the orbit remains bounded in the same space (irrespective of the size of $M(q)$). For the periodic (DNLS), the paper [Isom et al. 2020] shows that for $s \geq 1$ and $M(q)$ small, the $H^s(\mathbb{T})$-norm of solutions grows at most polynomially in time.

While these exciting results appeared too recently to affect what we do in this paper, their novelty and insightfulness give us every hope that the conjectures presented herein may soon be resolved.

2. Preliminaries

Our conventions for the Fourier transform are

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\xi x} f(x)\, dx, \quad \text{so} \quad f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i\xi x} \hat{f}(\xi)\, d\xi$$

for functions on the line, and

$$\hat{f}(\xi) = \int_0^1 e^{-i\xi x} f(x)\, dx, \quad \text{so} \quad f(x) = \sum_{\xi \in 2\pi\mathbb{Z}} \hat{f}(\xi) e^{i\xi x}$$

for functions on the torus $\mathbb{T}$. These definitions of the Fourier transform are unitary on L^2 and yield the Plancherel identities

$$\|f\|_{L^2(\mathbb{R})} = \|\hat{f}\|_{L^2(\mathbb{R})} \quad \text{and} \quad \|f\|_{L^2(\mathbb{T})} = \sum_{\xi \in 2\pi\mathbb{Z}} |\hat{f}(\xi)|^2,$$

as well as the following convolution identity on $\mathbb{R}$:

$$\widehat{fg} = \frac{1}{\sqrt{2\pi}} \hat{f} * \hat{g}.$$

We use the standard Littlewood–Paley decomposition of a function,

$$q = \sum_{N \in 2^{\mathbb{N}}} q_N,$$

based on a smooth partition of unity on the Fourier side. Here q_1 denotes the projection onto frequencies $|\xi| \leq 1$; for $N \geq 2$, frequencies $|\xi| \sim N$ are contained in q_N.

The fact that the operators Λ and Γ defined in (1-7) are Hilbert–Schmidt was noticed already in [Killip et al. 2018, Lemma 4.1]:

Lemma 2.1. *For $q \in L^2$ and $\kappa > 0$, we have*

$$\|\Lambda\|^2_{\mathfrak{I}_2(\mathbb{R})} = \|\Gamma\|^2_{\mathfrak{I}_2(\mathbb{R})} \approx \int_{\mathbb{R}} \log\Bigl(4+\frac{\xi^2}{\kappa^2}\Bigr)\frac{|\hat q(\xi)|^2}{\sqrt{4\kappa^2+\xi^2}}\,d\xi \lesssim \kappa^{-1}\|q\|^2_{L^2}, \tag{2-1}$$

$$\|\Lambda\|^2_{\mathfrak{I}_2(\mathbb{T})} = \|\Gamma\|^2_{\mathfrak{I}_2(\mathbb{T})} \approx \sum_{\xi\in 2\pi\mathbb{Z}} \log\Bigl(4+\frac{\xi^2}{\kappa^2}\Bigr)\frac{|\hat q(\xi)|^2}{\sqrt{4\kappa^2+\xi^2}} \lesssim \kappa^{-1}\|q\|^2_{L^2}. \tag{2-2}$$

Proof. The estimate (2-1) follows from the computation

$$\|\Lambda\|^2_{\mathfrak{I}_2(\mathbb{R})} = \frac{1}{2\pi}\int_{\mathbb{R}} |\hat q(\xi)|^2 \int_{\mathbb{R}} \frac{1}{\sqrt{\kappa^2+\eta^2}\sqrt{\kappa^2+(\eta+\xi)^2}}\,d\eta\,d\xi \approx \int_{\mathbb{R}} \log\Bigl(4+\frac{\xi^2}{\kappa^2}\Bigr)\frac{|\hat q(\xi)|^2}{\sqrt{4\kappa^2+\xi^2}}\,d\xi.$$

To compute the above integral in η, one treats separately the regions $|\eta| \le 2|\xi|$ and $|\eta| > 2|\xi|$; the logarithm term arises only when considering the first region.

On the torus, similar arguments yield

$$\|\Lambda\|^2_{\mathfrak{I}_2(\mathbb{T})} = \sum_{\xi\in 2\pi\mathbb{Z}} |\hat q(\xi)|^2 \sum_{\eta\in 2\pi\mathbb{Z}} \frac{1}{\sqrt{\kappa^2+\eta^2}\sqrt{\kappa^2+(\eta+\xi)^2}} \approx \sum_{\xi\in 2\pi\mathbb{Z}} \log\Bigl(4+\frac{\xi^2}{\kappa^2}\Bigr)\frac{|\hat q(\xi)|^2}{\sqrt{4\kappa^2+\xi^2}},$$

which settles (2-2). □

These Hilbert–Schmidt bounds ensure that $i\kappa\Lambda\Gamma$ is trace class and thus that the determinant in (1-8) is well defined. The trace of this operator will also be important and is easily evaluated.

Lemma 2.2. *Let $q \in L^2$ and $\kappa > 0$. Then*

$$\operatorname{tr}(i\kappa\Lambda\Gamma) = \int \frac{i\kappa|\hat q(\xi)|^2}{2\kappa - i\xi}\,d\xi \qquad \text{on } \mathbb{R}, \tag{2-3}$$

$$\operatorname{tr}(i\kappa\Lambda\Gamma) = \frac{1+e^{-\kappa}}{1-e^{-\kappa}} \sum_{\xi\in 2\pi\mathbb{Z}} \frac{i\kappa|\hat q(\xi)|^2}{2\kappa - i\xi} \qquad \text{on } \mathbb{T}. \tag{2-4}$$

Proof. To prove (2-3), we simply compute the trace on the Fourier side:

$$\operatorname{tr}(i\kappa\Lambda\Gamma) = \frac{i\kappa}{2\pi}\iint \frac{|\hat q(\xi)|^2}{(\eta - i\kappa)(\eta+\xi+i\kappa)}\,d\eta\,d\xi = \int \frac{i\kappa|\hat q(\xi)|^2}{2\kappa - i\xi}\,d\xi.$$

In the circle setting, we use the partial fraction decomposition of the cotangent:

$$\sum_{\eta\in 2\pi\mathbb{Z}} \Bigl(\frac{1}{\kappa+i\eta} + \frac{1}{\kappa - i\eta}\Bigr) = \coth\Bigl(\frac{\kappa}{2}\Bigr) = \frac{1+e^{-\kappa}}{1-e^{-\kappa}}. \tag{2-5}$$

In this way, we find

$$\operatorname{tr}(i\kappa\Lambda\Gamma) = i\kappa \sum_{\xi\in 2\pi\mathbb{Z}} |\hat q(\xi)|^2 \frac{1}{\xi+2i\kappa} \sum_{\eta\in 2\pi\mathbb{Z}} \Bigl(\frac{1}{\eta - i\kappa} - \frac{1}{\eta+\xi+i\kappa}\Bigr) = \sum_{\xi\in 2\pi\mathbb{Z}} \frac{i\kappa|\hat q(\xi)|^2}{2\kappa - i\xi}\,\frac{1+e^{-\kappa}}{1-e^{-\kappa}}.$$

Notice that the sum over η simplifies to (2-5) because $\xi \in 2\pi\mathbb{Z}$. □

In Section 5, it will be convenient to express the next term in the series (1-9) as a paraproduct. This is the role of the next lemma.

Lemma 2.3. *Let $q \in L^2$ and $\kappa > 0$. Then*

$$\operatorname{tr}([\Lambda\Gamma]^2) = \int_{\mathbb{R}} \Big(\frac{1}{2\kappa+\partial}\bar q\Big)^2 (4\kappa-\partial)\Big(\frac{1}{2\kappa-\partial}q\Big)^2\,dx \qquad \textit{on } \mathbb{R}, \tag{2-6}$$

$$\operatorname{tr}([\Lambda\Gamma]^2) = \frac{1+e^{-\kappa}}{1-e^{-\kappa}} \int_{\mathbb{T}} \Big(\frac{1}{2\kappa+\partial}\bar q\Big)^2 (4\kappa-\partial)\Big(\frac{1}{2\kappa-\partial}q\Big)^2\,dx \quad \textit{on } \mathbb{T}. \tag{2-7}$$

Proof. The method is exactly that of the previous lemma, only the details change. In the line case, we have a more complicated (but still elementary) contour integral. In the circle case, one must verify that

$$\sum_{\xi\in 2\pi\mathbb{Z}} \frac{1}{(\kappa+i\xi)(\kappa-i[\xi+\eta_1])(\kappa+i[\xi+\eta_1+\eta_2])(\kappa-i[\xi+\eta_1+\eta_2+\eta_3])}$$
$$= \frac{1+e^{-\kappa}}{1-e^{-\kappa}}\cdot \frac{4\kappa-i(\eta_1+\eta_3)}{(2\kappa-i\eta_1)(2\kappa+i\eta_2)(2\kappa-i\eta_3)(2\kappa+i\eta_4)}.$$

This follows from (2-5) via a careful partial fraction decomposition. □

Our next lemma records operator estimates for frequency localized potentials.

Lemma 2.4 (operator estimates). *Fix $q \in L^2$, $N \in 2^{\mathbb{N}}$, and $\kappa \ge 1$, and write $\Lambda_N = \Lambda(q_N)$ and $\Gamma_N = \Gamma(q_N)$. Then*

$$\|\Lambda_N\|_{\mathfrak{I}_2} = \|\Gamma_N\|_{\mathfrak{I}_2} \approx \sqrt{\frac{1}{\kappa+N}\log\Big(4+\frac{N^2}{\kappa^2}\Big)}\,\|q_N\|_{L^2}, \tag{2-8}$$

$$\|\Lambda_N\|_{\mathrm{op}} = \|\Gamma_N\|_{\mathrm{op}} \lesssim \min\Big\{\frac{\sqrt{N}}{\kappa}, \sqrt{\frac{1}{\kappa+N}\log\Big(4+\frac{N^2}{\kappa^2}\Big)}\Big\}\|q_N\|_{L^2}, \tag{2-9}$$

$$\sum_{N\le N_0} \|\Lambda_N\|_{\mathrm{op}} \lesssim \kappa^{-1}\min\{\sqrt{N_0}, \sqrt{\kappa}\}\|q\|_{L^2}. \tag{2-10}$$

Proof. The claim (2-8) follows immediately from Lemma 2.1.

Using the Bernstein inequality, we estimate

$$\|\Lambda_N\|_{\mathrm{op}} \le \|(\kappa-\partial)^{-1/2}\|_{\mathrm{op}}\|q_N\|_{\mathrm{op}}\|(\kappa+\partial)^{-1/2}\|_{\mathrm{op}} \le \frac{1}{\kappa}\|q_N\|_{L^\infty} \lesssim \frac{\sqrt{N}}{\kappa}\|q_N\|_{L^2}.$$

Combining this with (2-8) yields (2-9).

The case $N_0 \le \kappa$ of (2-10) is clear. If $N_0 > \kappa$, an application of (2-9) yields

$$\sum_{N\le N_0}\|\Lambda_N\|_{\mathrm{op}} \lesssim \sum_{N\le\kappa}\frac{\sqrt{N}}{\kappa}\|q\|_{L^2} + \sum_{\kappa<N\le N_0}\sqrt{\frac{1}{N}\log\Big(4+\frac{N^2}{\kappa^2}\Big)}\|q\|_{L^2} \lesssim \frac{\sqrt{\kappa}}{\kappa}\|q\|_{L^2},$$

as desired. □

Lemma 2.5. *For all $\kappa \geq 1$, we have*

$$\|(\kappa+\partial)^{-1} f(\kappa-\partial)^{-1}\|_{\mathfrak{I}_2} \lesssim \kappa^{-1/2}\|f\|_{H^{-1}}, \tag{2-11}$$

$$\|q(\kappa+\partial)^{-3/4}\|_{\mathfrak{I}_2} \lesssim \kappa^{-1/4}\|q\|_{L^2}, \tag{2-12}$$

$$\|(\kappa-\partial)^{-1/4} q(\kappa+\partial)^{-1/4}\|_{\mathrm{op}} \lesssim \|q\|_{L^2}. \tag{2-13}$$

Proof. We first turn to (2-11). We will only consider here the line setting; in the periodic case, one can apply a similar argument to the one in the proof of Lemma 2.1. A straightforward computation yields

$$\|(\kappa+\partial)^{-1} f(\kappa-\partial)^{-1}\|_{\mathfrak{I}_2}^2 = \frac{1}{2\pi}\iint \frac{|\hat f(\xi)|^2}{[\kappa^2+(\xi+\eta)^2](\kappa^2+\eta^2)}\,d\eta\,d\xi.$$

Considering separately the regions $|\eta| \leq 2|\xi|$ and $|\eta| > 2|\xi|$ when integrating in η, we find

$$\|(\kappa+\partial)^{-1} f(\kappa-\partial)^{-1}\|_{\mathfrak{I}_2}^2 \lesssim \int \frac{|\hat f(\xi)|^2}{\kappa(\kappa^2+\xi^2)}\,d\xi \lesssim \kappa^{-1}\|f\|_{H^{-1}}^2.$$

By direct computation (compare [Simon 2005, Theorem 4.1]), we have

$$\|q(\kappa+\partial)^{-3/4}\|_{\mathfrak{I}_2} \lesssim \|q\|_{L^2}\|(\kappa+i\xi)^{-3/4}\|_{L^2_\xi} \lesssim \kappa^{-1/4}\|q\|_{L^2},$$

which settles (2-12).

Similarly, by Cwikel's theorem (see [Cwikel 1977] or [Simon 2005, Theorem 4.2]), we find that

$$\begin{aligned}\|(\kappa-\partial)^{-1/4} q(\kappa+\partial)^{-1/4}\|_{\mathrm{op}} &\leq \|(\kappa-\partial)^{-1/4}\sqrt{|q|}\|_{\mathrm{op}} \left\| \frac{q}{\sqrt{|q|}}(\kappa+\partial)^{-1/4}\right\|_{\mathrm{op}} \\ &\lesssim \|(\kappa\pm i\xi)^{-1/4}\|_{L^4_{\mathrm{weak}}}^2 \|\sqrt{|q|}\|_{L^4}^2 \lesssim \|q\|_{L^2}.\end{aligned}$$

□

Proposition 2.6. *Let Q be a bounded and equicontinuous subset of L^2. Then*

$$\lim_{\kappa\to\infty} \sup_{q\in Q} \sqrt{\kappa}\,\|\Lambda(q)\|_{\mathrm{op}} = 0. \tag{2-14}$$

Moreover, there exists $\kappa_0 \geq 1$ such that the series (1-9) *converges uniformly for $\kappa \geq \kappa_0$ and $q \in Q$.*

Proof. Fix $\varepsilon > 0$ and let $\eta > 0$ be a small parameter to be chosen later. Using (2-10) and Lemma 2.1, we get

$$\sqrt{\kappa}\,\|\Lambda(q)\|_{\mathrm{op}} \lesssim \sqrt{\kappa}\,\|\Lambda(q_{>\eta\kappa})\|_{\mathrm{op}} + \sqrt{\kappa}\sum_{N\leq\eta\kappa}\|\Lambda_N(q)\|_{\mathrm{op}} \lesssim \|q_{>\eta\kappa}\|_{L^2} + \sqrt{\eta}\,\|q\|_{L^2}.$$

Choosing η small enough depending on the L^2 bound of Q, and then κ sufficiently large depending on η and the equicontinuity property of Q, we may ensure that

$$\sqrt{\kappa}\,\|\Lambda(q)\|_{\mathrm{op}} < \varepsilon \quad \text{for all } q \in Q,$$

which yields (2-14).

To continue, we choose κ_0 sufficiently large such that for any $\kappa \geq \kappa_0$ we have $\sqrt{\kappa}\,\|\Lambda(q)\|_{\mathrm{op}} \leq \frac{1}{2}$ uniformly for $q \in Q$. Lemma 2.1 then yields

$$\|(i\kappa\Lambda\Gamma)^{\ell+1}\|_{\mathfrak{I}_1} \leq \kappa^{\ell+1}\|\Lambda\|_{\mathfrak{I}_2}^2\|\Lambda\|_{\mathrm{op}}^{2\ell} \lesssim 2^{-\ell}\|q\|_{L^2}^2, \tag{2-15}$$

uniformly for $\kappa \geq \kappa_0$ and $q \in Q$, which ensures convergence of the series (1-9). □

As discussed in the introduction, this convergence result allows the arguments of [Klaus and Schippa 2022; Tang and Xu 2021] to be extended beyond the regime of small L^2-norm and so show that $\alpha(\kappa; q)$ is conserved under the (DNLS) flow, for κ sufficiently large. This conservation is inherited by $a(\kappa; q)$ for *all* $\operatorname{Re}\kappa > 0$ because this is a holomorphic function in this region.

3. Equicontinuity in L^2

The goal of this section is to prove Theorem 1.3. We begin with a convenient notion of the momentum at high frequencies in each geometry:

$$\beta_{\mathbb{R}}^{[2]}(\kappa; q) := \int_{\mathbb{R}} \frac{\xi^2|\hat{q}(\xi)|^2}{4\kappa^2+\xi^2}\,d\xi \quad \text{and} \quad \beta_{\mathbb{T}}^{[2]}(\kappa; q) := \sum_{\xi\in 2\pi\mathbb{Z}} \frac{\xi^2|\hat{q}(\xi)|^2}{4\kappa^2+\xi^2}. \tag{3-1}$$

The curious notation is explained by the fact that these expressions coincide with the quadratic (in q) parts of the quantities in (4-3). For our immediate purposes, however, the following relation with the formulas of Lemma 2.2 is more important:

$$\begin{aligned} \operatorname{Im}\operatorname{tr}(i\kappa\Lambda\Gamma) &= \frac{1}{2}[M(q) - \beta_{\mathbb{R}}^{[2]}(\kappa; q)] && \text{on } \mathbb{R}, \\ \operatorname{Im}\operatorname{tr}(i\kappa\Lambda\Gamma) &= \frac{1}{2}\frac{1+e^{-\kappa}}{1-e^{-\kappa}}[M(q) - \beta_{\mathbb{T}}^{[2]}(\kappa; q)] && \text{on } \mathbb{T}. \end{aligned} \tag{3-2}$$

Given an infinite subset $\mathcal{K} \subseteq 2^{\mathbb{N}}$, we then define a norm via

$$\|q\|_{\mathcal{K}}^2 := \|q\|_{L^2}^2 + \sum_{\kappa\in\mathcal{K}} \beta^{[2]}(\kappa; q). \tag{3-3}$$

This in turn leads to a very convenient formulation of equicontinuity.

Lemma 3.1. *A set $Q \subseteq L^2$ is bounded and equicontinuous if and only if there exists an infinite set $\mathcal{K} \subseteq 2^{\mathbb{N}}$ such that* $\sup_{q\in Q}\|q\|_{\mathcal{K}} < \infty$.

Proof. This is immediately evident from the observation that

$$\|q\|_{\mathcal{K}}^2 \approx \|q\|_{L^2}^2 + \sum_{\kappa\in\mathcal{K}} \|q_{>\kappa}\|_{L^2}^2 \approx \|q\|_{L^2}^2 + \sum_{N\in 2^{\mathbb{N}}} \#\{\kappa\in\mathcal{K} : \kappa < N\}\|q_N\|_{L^2}^2. \qquad \square$$

Before beginning the proof of Theorem 1.3, we need two further preliminaries. The first will allow us to pass from the determinant to the exponentiated trace, and the second to take logarithms.

Lemma 3.2. *Let $A \in \mathfrak{I}_1$. Then*

$$|\det(1+A) - \exp\{\operatorname{tr}(A)\}| \leq \tfrac{1}{2}\|A\|_{\mathfrak{I}_2}^2 \exp\{\|A\|_{\mathfrak{I}_1}\}. \tag{3-4}$$

Proof. Let λ_i enumerate the nonzero eigenvalues of A repeated according to algebraic multiplicity. By relating eigenvalues and singular values, Weyl proved that

$$\sum |\lambda_i| \leq \|A\|_{\mathfrak{I}_1} \quad \text{and} \quad \sum |\lambda_i|^2 \leq \|A\|_{\mathfrak{I}_2}^2.$$

Now let us compare

$$\det(1+A) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\substack{i_1,\dots,i_n \\ \text{distinct}}} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_n}, \quad \text{and} \quad \exp\{\operatorname{tr}(A)\} = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{i_1,\dots,i_n} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_n}.$$

Evidently, the difference contains only sums over n-tuples $(i_1, \dots, i_n)$ that contain at least one pair of identical indices. Thus,

$$\text{LHS of (3-4)} \leq \sum_{n=2}^{\infty} \frac{1}{n!}\binom{n}{2}\Big[\sum_j |\lambda_j|^2\Big]\Big[\sum_i |\lambda_i|\Big]^{n-2} \leq \frac{1}{2}\|A\|_{\mathfrak{I}_2}^2 \sum_{n=2}^{\infty} \frac{1}{(n-2)!}\|A\|_{\mathfrak{I}_1}^{n-2},$$

and so (3-4) follows. □

Lemma 3.3. *Given $C > 0$ and $0 < \varepsilon < \pi$, let*

$$\mathcal{R} = \{z : |\operatorname{Re} z| \leq C \text{ and } 0 < \operatorname{Im} z < 2\pi - \varepsilon\}. \tag{3-5}$$

Then

$$|\operatorname{Im}(z-w)| \leq \frac{\pi e^C}{\sin(\varepsilon/2)}|e^w - e^z| \quad \text{uniformly for } z, w \in \mathcal{R}. \tag{3-6}$$

Proof. This reduces to elementary trigonometry once one realizes that the worst-case scenario is $\operatorname{Re} z = \operatorname{Re} w = -C$. □

We are now ready for the climax of the section.

Proof of Theorem 1.3. Let us begin right away with the key computation. Given any $q \in L^2$, we may apply (2-8), (2-10), and (in the final step) Cauchy–Schwarz to deduce that

$$\begin{aligned}
\sum_{\kappa\in 2^{\mathbb{N}}} \|i\kappa\Lambda(q)\Gamma(q)\|_{\mathfrak{I}_2}^2 &\lesssim \sum_{\kappa\in 2^{\mathbb{N}}} \kappa^2 \sum_{N_1\sim N_2\geq N_3,N_4} \|\Lambda_{N_1}(q)\|_{\mathfrak{I}_2}\|\Lambda_{N_2}(q)\|_{\mathfrak{I}_2}\|\Lambda_{N_3}(q)\|_{\mathrm{op}}\|\Lambda_{N_4}(q)\|_{\mathrm{op}} \\
&\lesssim M(q)\sum_{\kappa\in 2^{\mathbb{N}}}\sum_{N_1\sim N_2} \frac{1}{N_2+\kappa}\log\Big(4+\frac{N_2^2}{\kappa^2}\Big)\|q_{N_1}\|_{L^2}\|q_{N_2}\|_{L^2}\min\{N_2,\kappa\} \\
&\lesssim M(q)\sum_{N_1\sim N_2}\|q_{N_1}\|_{L^2}\|q_{N_2}\|_{L^2}\Big(\sum_{\kappa\leq N_2}\frac{\kappa}{N_2}\log\Big(4+\frac{N_2^2}{\kappa^2}\Big)+\sum_{\kappa>N_2}\frac{N_2}{\kappa}\Big) \\
&\lesssim M(q)^2.
\end{aligned} \tag{3-7}$$

Combining this with Lemmas 2.1 and 3.2, we find

$$\sum_{\kappa\in 2^{\mathbb{N}}} \big|a(\kappa;q) - \exp\{-\operatorname{tr}[i\kappa\Lambda(q)\Gamma(q)]\}\big| \leq CM(q)^2 e^{CM(q)} \tag{3-8}$$

for some absolute C.

As we did not explicitly require that $M(\tilde{q}) = M(q)$ for $\tilde{q} \in C_q^0$, let us pause to see that this follows from the equality $a(\kappa; \tilde{q}) \equiv a(\kappa; q)$. From (3-4) and Lemma 2.2 we see that for $\kappa \to \infty$,

$$\begin{aligned} 0 = |a(\kappa; \tilde{q}) - a(\kappa; q)| &= \big|\exp\{-\operatorname{tr}[i\kappa\Lambda(\tilde{q})\Gamma(\tilde{q})] - \exp\{-\operatorname{tr}[i\kappa\Lambda(q)\Gamma(q)]\}\big| + o(1) \\ &= \Big|\exp\Big\{-\frac{i}{2}M(\tilde{q})\Big\} - \exp\Big\{-\frac{i}{2}M(q)\Big\}\Big| + o(1). \end{aligned}$$

Thus $M(\tilde{q})$ is preserved modulo $4\pi\mathbb{Z}$. As $\tilde{q}$ belongs to the same connected component as q, we must have that $M(\tilde{q}) = M(q)$. For later use, we note the consequence

$$\sup_{q \in Q_{**}} M(q) = \sup_{q \in Q} M(q). \tag{3-9}$$

While this argument did not require the hypothesis (1-12), we will need it to unwrap this phase ambiguity when we address equicontinuity. This is our next topic.

Given an equicontinuous set Q satisfying (1-12), choose $\varepsilon > 0$ and an infinite subset $\mathcal{K} \subseteq 2^{\mathbb{N}}$ such that

$$\sup_{q \in Q, \kappa \in \mathcal{K}} \frac{1+e^{-\kappa}}{1-e^{-\kappa}} M(q) \le 4\pi - 2\varepsilon \quad \text{and} \quad \sup_{q \in Q} \|q\|_{\mathcal{K}} < \infty.$$

Proceeding very much as we did above, we see that

$$\sum_{\kappa \in \mathcal{K}} \big|\exp\{-\operatorname{tr}[i\kappa\Lambda(\tilde{q})\Gamma(\tilde{q})] - \exp\{-\operatorname{tr}[i\kappa\Lambda(q)\Gamma(q)]\}\big| \le 2CM(q)^2 e^{CM(q)}$$

for any $\tilde{q} \in C_q^0$. Combining this with (3-2) and Lemma 3.3, we deduce that

$$\sum_{\kappa \in \mathcal{K}} |\beta^{[2]}(\kappa; \tilde{q}) - \beta^{[2]}(\kappa; q)| \lesssim_\varepsilon 1.$$

This in turn guarantees that

$$\sup\{\|\tilde{q}\|_{\mathcal{K}}^2 : \tilde{q} \in Q_{**}\} \le \sup\{\|q\|_{\mathcal{K}}^2 : q \in Q\} + O_\varepsilon(1) < \infty,$$

from which equicontinuity follows via Lemma 3.1. □

4. Conservation laws and equicontinuity

The primary goal of this section is to prove H^s bounds for (DNLS) solutions, for $0 < s < \frac{1}{2}$, as a prerequisite for proving Theorem 1.5. In addition, we will prove equicontinuity in these spaces, which is also needed to prove that theorem.

Before turning to that subject, we pause to show how L^2-equicontinuity can be used to restore coercivity to the traditional polynomial conservation laws. As a representative example, we show how $H_2(q)$ can be used to control the H^1-norm.

Proposition 4.1. *Let $Q \subseteq H^1$ be L^2-bounded and equicontinuous. Then*

$$\|q\|_{H^1}^2 \lesssim H_2(q) + M(q)^3, \tag{4-1}$$

uniformly for all $q \in Q$.

Proof. Splitting into low and high frequency parts and estimating using the Bernstein and Gagliardo–Nirenberg inequalities, respectively, we obtain

$$\|q\|_{L^6}^6 \lesssim \|q_{\leq N}\|_{L^6}^6 + \|q_{>N}\|_{L^6}^6 \lesssim N^2\|q\|_{L^2}^6 + \|q_{>N}\|_{L^2}^4\|q'\|_{L^2}^2.$$

This allows us to control the quartic term in H_2, and hence the H^1-norm, as follows:

$$\begin{aligned}\|q'\|_{L^2}^2 &\leq H_2(q) + \frac{3}{2}\int |q(x)|^3|q'(x)|\,dx\\ &\leq H_2(q) + \varepsilon\|q'\|_{L^2}^2 + \frac{9}{16\varepsilon}\|q\|_{L^6}^6\\ &\leq H_2(q) + \Big(\varepsilon + C\frac{9}{16\varepsilon}\|q_{>N}\|_{L^2}^4\Big)\|q'\|_{L^2}^2 + C\frac{9}{16\varepsilon}N^2M(q)^3,\end{aligned}$$

for any $\varepsilon > 0$. The claim (4-1) now follows by choosing ε small and then N large, exploiting the equicontinuity of Q. □

Proposition 4.1 allows us to extend local H^1 solutions globally in time, provided we remain below the M_* bound introduced in Definition 1.4.

Corollary 4.2. *The* (DNLS) *evolution is globally well posed*, *both on the line and on the circle*, *in the space*

$$B^1_{M_*} = \{q \in H^1 : \|q\|_{L^2}^2 < M_*\} \tag{4-2}$$

endowed with the H^1 topology. Moreover, *initial data in $\mathcal{S}$ leads to solutions that belong to $\mathcal{S}$ at all times.*

Proof. In the line case, this result can be deduced from [Bahouri and Perelman 2022]; indeed, the restriction $M(q) < M_*$ is not needed in this case. Below we give an alternate argument that works also in the periodic setting.

As discussed in the introduction, local well-posedness in H^1 was proved already in [Takaoka 1999; Herr 2006]. Thus, given initial data $q(0) \in B^1_{M_*} \cap \mathcal{S}$, there is a corresponding maximal lifespan solution $q \in C_t([0, T); H^1)$ to (DNLS). Moreover, [Hayashi and Ozawa 1992] shows that $q(t) \in \mathcal{S}$ for all $t \in [0, T)$. Combining [Klaus and Schippa 2022; Tang and Xu 2021] with Proposition 2.6 yields that $a(\kappa; q(t)) = a(\kappa, q(0))$ for all $t \in [0, T)$ and $\kappa > 0$. By the definition of M_* and Proposition 4.1, the solution q satisfies a priori H^1 bounds on $[0, T)$, which in turn guarantees that $T = \infty$.

Finally, global well-posedness in $B^1_{M_*}$ follows from local well-posedness and the density of $\mathcal{S}$ in H^1. □

Let us now turn to low-regularity questions. Bounded sets in H^s, with $s > 0$, are automatically bounded and equicontinuous in L^2. As we shall work only below the M_* threshold in this section, such L^2-equicontinuity is retained globally in time. Our goal is to propagate H^s bounds. The key to doing this is a certain renormalization of $\alpha(\kappa; q)$ that we introduce now:

$$\begin{aligned}\beta_{\mathbb{R}}(\kappa; q) &:= \|q\|_{L^2}^2 - 2\operatorname{Im}\alpha(\kappa; q) &&\text{on } \mathbb{R},\\ \beta_{\mathbb{T}}(\kappa; q) &:= \|q\|_{L^2}^2 - \frac{1-e^{-\kappa}}{1+e^{-\kappa}}\,2\operatorname{Im}\alpha(\kappa; q) &&\text{on } \mathbb{T}.\end{aligned} \tag{4-3}$$

Proposition 2.6 guarantees that these quantities are well defined for κ sufficiently large across our whole family of orbits.

The quadratic (in q) parts of these expressions were presented already in (3-1). As we saw there, these provide a sense of the L^2-norm of the high-frequency part of q. To address higher regularity, for $0 < s < \frac{1}{2}$ we consider the quantity

$$\beta_s(\kappa; q) := \int_\kappa^\infty \beta(\varkappa; q)\varkappa^{2s}\,\frac{d\varkappa}{\varkappa}.$$

The quadratic term in this expression is given by

$$\beta_s^{[2]}(\kappa; q) = \int_\kappa^\infty \beta^{[2]}(\varkappa; q)\varkappa^{2s}\,\frac{d\varkappa}{\varkappa} = \int_\kappa^\infty \Big\langle \frac{-\partial^2}{4\varkappa^2 - \partial^2} q, q\Big\rangle \varkappa^{2s}\,\frac{d\varkappa}{\varkappa} \approx_s \Big\langle \frac{-\partial^2}{(\kappa^2 - \partial^2)^{1-s}} q, q\Big\rangle.$$

From this we see that for any $0 < \eta < 1$,

$$\|q_{>\kappa}\|_{H^s}^2 \lesssim \beta_s^{[2]}(\kappa; q) \lesssim \eta^{2(1-s)}\|q\|_{H^s}^2 + \|q_{>\eta\kappa}\|_{H^s}^2, \tag{4-4}$$

and so $\beta_s^{[2]}(\kappa; q)$ captures the H^s-norm of the high-frequency part of q. Indeed, a bounded set $Q \subseteq H^s$ is equicontinuous in H^s if and only if $\beta_s^{[2]}(\kappa; q) \to 0$ uniformly on Q as $\kappa \to \infty$.

Theorem 4.3. *Fix $0 < s < \frac{1}{2}$ and let $Q \subseteq \mathcal{S}$ be H^s-bounded and satisfy* (1-13). *Then*, *recalling the notation Q_{**} from* (1-11), *we have*

$$\sup_{q \in Q_{**}} \|q\|_{H^s} \lesssim C\big(\sup_{q \in Q}\|q\|_{L^2}^2, \sup_{q\in Q}\|q\|_{H^s}^2\big). \tag{4-5}$$

Moreover, *if Q is H^s-equicontinuous*, *then so is Q_{**}.*

Proof. As Q is H^s-bounded, it is automatically L^2-bounded and equicontinuous. By (3-9), Q_{**} inherits L^2-boundedness from Q. As Q satisfies (1-13), we deduce that Q_{**} is also L^2-equicontinuous. By Proposition 2.6, we may choose $\kappa_0 \geq 1$ such that

$$\sqrt{\kappa}\,\|\Lambda(q)\|_{\mathrm{op}} \leq \tfrac{1}{2} \quad \text{uniformly for } q \in Q_{**} \text{ and } \kappa \geq \kappa_0. \tag{4-6}$$

As shown there, this ensures that $\alpha(\kappa; q)$ and so also $\beta(\kappa; q)$ are well defined for all $q \in Q_{**}$ and $\kappa \geq \kappa_0$.

Arguing as in (2-15), we also see that (4-6) implies

$$\begin{aligned} |\beta_s(\kappa; q) - \beta_s^{[2]}(\kappa; q)| &\lesssim \int_\kappa^\infty \varkappa^{2s+2}\|\Lambda(q)\Gamma(q)\|_{\mathfrak{I}_2}^2\,\frac{d\varkappa}{\varkappa} \\ &\lesssim \int_\kappa^\infty \varkappa^{2s+2} \sum_{N_1 \sim N_2 \geq N_3 \geq N_4} \|\Lambda_{N_1}\|_{\mathfrak{I}_2}\|\Lambda_{N_2}\|_{\mathfrak{I}_2}\|\Lambda_{N_3}\|_{\mathrm{op}}\|\Lambda_{N_4}\|_{\mathrm{op}}\,\frac{d\varkappa}{\varkappa} \end{aligned} \tag{4-7}$$

uniformly for $q \in Q_{**}$ and $\kappa \geq \kappa_0$. To continue from here, we decompose the full sum into the subregions S_j defined by

$$S_1 = \{N_2 \leq \kappa\}, \qquad \begin{aligned} S_2 &= \{\kappa < N_2 \leq \varkappa \text{ and } N_3 \leq \eta\kappa\}, & S_4 &= \{N_2 > \varkappa \text{ and } N_3 \leq \eta\kappa\}, \\ S_3 &= \{\kappa < N_2 \leq \varkappa \text{ and } N_3 > \eta\kappa\}, & S_5 &= \{N_2 > \varkappa \text{ and } N_3 > \eta\kappa\}, \end{aligned}$$

where $\eta \in (0, 1)$ is a small parameter to be chosen later. We will estimate separately each of the contributions

$$I_j(\kappa; q) := \int_\kappa^\infty \varkappa^{2s+2} \sum_{S_j} \|\Lambda_{N_1}\|_{\mathfrak{I}_2} \|\Lambda_{N_2}\|_{\mathfrak{I}_2} \|\Lambda_{N_3}\|_{\mathrm{op}} \|\Lambda_{N_4}\|_{\mathrm{op}} \,\frac{d\varkappa}{\varkappa}.$$

Applying (2-8) and (2-10) from Lemma 2.4, we have

$$\begin{aligned} I_1 &\lesssim \int_\kappa^\infty \varkappa^{2s+2} \sum_{N_1 \sim N_2 \le \kappa} \frac{N_2}{\varkappa^3} \|q_{N_1}\|_{L^2} \|q_{N_2}\|_{L^2} \|q\|_{L^2}^2 \,\frac{d\varkappa}{\varkappa} \\ &\lesssim \kappa^{2s-1} \|q\|_{L^2}^2 \sum_{N_1 \sim N_2 \le \kappa} N_2 \|q_{N_1}\|_{L^2} \|q_{N_2}\|_{L^2} \lesssim \kappa^{2s} \|q\|_{L^2}^4. \end{aligned}$$

Proceeding analogously and using (4-4), we find

$$\begin{aligned} I_2 &\lesssim \sum_{N_1 \sim N_2 > \kappa} \int_{N_2}^\infty \eta\kappa \varkappa^{2s-1} N_2^{-2s} \|q_{N_1}\|_{H^s} \|q_{N_2}\|_{H^s} \|q\|_{L^2}^2 \,\frac{d\varkappa}{\varkappa} \\ &\lesssim \sum_{N_1 \sim N_2 > \kappa} \eta\kappa N_2^{-1} \|q_{N_1}\|_{H^s} \|q_{N_2}\|_{H^s} \|q\|_{L^2}^2 \lesssim \eta \|q\|_{L^2}^2 \beta_s^{[2]}(\kappa; q), \\ I_3 &\lesssim \sum_{N_1 \sim N_2 > \kappa} \int_{N_2}^\infty \varkappa^{2s-1} N_2^{1-2s} \|q_{N_1}\|_{H^s} \|q_{N_2}\|_{H^s} \|q\|_{L^2} \|q_{>\eta\kappa}\|_{L^2} \,\frac{d\varkappa}{\varkappa} \\ &\lesssim \sum_{N_1 \sim N_2 > \kappa} \|q_{N_1}\|_{H^s} \|q_{N_2}\|_{H^s} \|q\|_{L^2} \|q_{>\eta\kappa}\|_{L^2} \lesssim \|q\|_{L^2} \|q_{>\eta\kappa}\|_{L^2} \beta_s^{[2]}(\kappa; q), \\ I_4 &\lesssim \sum_{N_1 \sim N_2 > \kappa} \int_\kappa^{N_2} \eta\kappa \varkappa^{2s} \log\Bigl(4 + \frac{N_2^2}{\varkappa^2}\Bigr) N_2^{-1-2s} \|q_{N_1}\|_{H^s} \|q_{N_2}\|_{H^s} \|q\|_{L^2}^2 \,\frac{d\varkappa}{\varkappa} \\ &\lesssim \sum_{N_1 \sim N_2 > \kappa} \eta\kappa N_2^{-1} \|q_{N_1}\|_{H^s} \|q_{N_2}\|_{H^s} \|q\|_{L^2}^2 \lesssim \eta \|q\|_{L^2}^2 \beta_s^{[2]}(\kappa; q), \end{aligned}$$

and finally,

$$\begin{aligned} I_5 &\lesssim \sum_{N_1 \sim N_2 > \kappa} \int_\kappa^{N_2} \varkappa^{2s} \log\Bigl(4 + \frac{N_2^2}{\varkappa^2}\Bigr) N_2^{-2s} \|q_{N_1}\|_{H^s} \|q_{N_2}\|_{H^s} \|q\|_{L^2} \|q_{>\eta\kappa}\|_{L^2} \,\frac{d\varkappa}{\varkappa} \\ &\lesssim \sum_{N_1 \sim N_2 > \kappa} \|q_{N_1}\|_{H^s} \|q_{N_2}\|_{H^s} \|q\|_{L^2} \|q_{>\eta\kappa}\|_{L^2} \lesssim \|q\|_{L^2} \|q_{>\eta\kappa}\|_{L^2} \beta_s^{[2]}(\kappa; q). \end{aligned}$$

Collecting all our estimates, we conclude that

$$|\beta_s(\kappa; q) - \beta_s^{[2]}(\kappa; q)| \lesssim \kappa^{2s} \|q\|_{L^2}^4 + (\eta \|q\|_{L^2}^2 + \|q\|_{L^2} \|q_{>\eta\kappa}\|_{L^2}) \beta_s^{[2]}(\kappa; q)$$

uniformly on Q_{**}. As Q_{**} is L^2-bounded and equicontinuous, we may choose η small and then $\kappa_1 \ge \kappa_0$ large to deduce that

$$\sup_{q \in Q_{**}} \beta_s^{[2]}(\kappa; q) \lesssim \sup_{q \in Q} \beta_s^{[2]}(\kappa; q) + \kappa^{2s} \sup_{q \in Q} \|q\|_{L^2}^4, \quad \text{for all } \kappa \ge \kappa_1. \tag{4-8}$$

The claim (4-5) now follows from (4-4) by choosing $\kappa = \kappa_1$.

It remains to prove that H^s-equicontinuity for Q is inherited by Q_{**}. This requires a different estimate for I_1. Using (2-8) and (2-9), we obtain

$$I_1 \lesssim \sum_{N_4\le\cdots\le N_1\le\kappa} \sqrt{N_3N_4}\,\|q_{N_1}\|_{L^2}\|q_{N_2}\|_{L^2}\|q_{N_3}\|_{L^2}\|q_{N_4}\|_{L^2}\int_\kappa^\infty \varkappa^{2s-1}\,\frac{d\varkappa}{\varkappa} \lesssim \kappa^{2s-4\sigma}\|q\|_{H^s}^4,$$

where $\sigma=\min\{s,\frac14\}$. Now that we know (4-5), we may employ it here to deduce the following analogue of (4-8):

$$\sup_{q\in Q_{**}} \beta_s^{[2]}(\kappa;q) \lesssim \sup_{q\in Q}\beta_s^{[2]}(\kappa;q)+\kappa^{2s-4\sigma}C\Big(\sup_{q\in Q}\|q\|_{L^2}^2,\sup_{q\in Q}\|q\|_{H^s}^2\Big)^4 \tag{4-9}$$

uniformly for $\kappa\ge\kappa_1$. As $4\sigma>2s$, equicontinuity follows by sending $\kappa\to\infty$. □

5. Global well-posedness in H^s for $s \ge \frac16$

In order to treat the line and circle simultaneously, it is convenient to introduce

$$A(\kappa;q)=\alpha(\kappa;q)\quad\text{on }\mathbb{R}\qquad\text{and}\qquad A(\kappa;q)=\frac{1-e^{-\kappa}}{1+e^{-\kappa}}\alpha(\kappa;q)\quad\text{on }\mathbb{T}. \tag{5-1}$$

This leads to parallel leading asymptotic expansions:

$$A(\kappa;q)=\frac{i}{2}M(q)+\frac{1}{4\kappa}H(q)+O\Big(\frac{1}{\kappa^2}\Big),$$

as follows from Lemmas 2.2 and 2.3. This expansion is important; it guides our choice of regularized Hamiltonian flows. We choose

$$H_\kappa(q):=4\kappa\,\mathrm{Re}\,A(\kappa;q),$$

since, formally at least, $H(q)=H_\kappa(q)+O(\kappa^{-1})$, which suggests that the flow generated by $H_\kappa(q)$ approximates the (DNLS) flow as the parameter κ diverges to infinity.

The flow generated by $H_\kappa(q)$ with respect to the Poisson structure (1-5) is

$$\frac{d}{dt}q=\Big(\frac{\delta H_\kappa}{\delta\bar q}\Big)'=2\kappa\Big(\frac{\delta A(\kappa;q)}{\delta\bar q}+\overline{\frac{\delta A(\kappa;q)}{\delta q}}\Big)',\quad\text{since }\frac{\delta\bar A}{\delta\bar q}=\overline{\frac{\delta A}{\delta q}}. \tag{H_κ}$$

Our first task in this section is to prove that the H_κ flow is well posed on L^2-equicontinuous sets of Schwartz initial data satisfying (1-13), provided κ is chosen sufficiently large depending on the equicontinuous family; see Proposition 5.3. Moreover, we will show that the corresponding solutions belong to $\mathcal{S}$ for all times.

In Lemma 5.2, the H_κ flow will be shown to conserve $M(q)$ and $\alpha(\varkappa;q)$; thus, it satisfies both the H^s-bounds and the H^s-equicontinuity guaranteed by Theorem 4.3. Together with Proposition 5.3, this immediately yields well-posedness of the H_κ flow on H^s for all $0\le s<\frac12$ under the restriction (1-13); see Corollary 5.4.

To prove that the (DNLS) flow is well posed in H^s for $\frac16\le s<\frac12$, it then suffices to prove that this is well approximated by (H_κ) flows as $\kappa\to\infty$. An important ingredient in our argument is the commutativity of the H_κ and (DNLS) flows, at least on $\mathcal{S}$. This follows from Lemma 5.2 and the well-posedness of these flows on $\mathcal{S}$ by mimicking the arguments in [Arnold 1989, §39]. In view of this commutativity, proving convergence of the (H_κ) flows to the (DNLS) flow amounts to showing that the flow generated by the difference of the Hamiltonians $H(q)-H_\kappa(q)$ converges to the identity as $\kappa\to\infty$. This final stage of the proof will be carried out in Theorem 5.5.

In order to make sense of (H_κ), we must prove that $\alpha(\kappa;q)$ is in fact differentiable. To solve (H_κ) locally in time, we further need to show that this functional derivative is itself a Lipschitz function of q. These goals require us to define $\alpha(\kappa;q)$ on open sets in L^2, rather than merely equicontinuous sets. The next result addresses these issues.

Here and below we write Q_ε to denote the ε neighborhood of Q in the L^2-metric.

Lemma 5.1. *Let Q be a bounded and equicontinuous subset of L^2. Then there exist $\varepsilon>0$ and $\kappa_0\ge 1$ such that for all $\kappa\ge\kappa_0$, we have that $\alpha(\kappa;q)$ is a real-analytic function of $q\in Q_\varepsilon$. Moreover, we have the bounds*

$$\Big\|\frac{\delta\alpha(\kappa;q)}{\delta q}\Big\|_{H^1}+\Big\|\frac{\delta\alpha(\kappa;q)}{\delta\bar q}\Big\|_{H^1}\lesssim\kappa\|q\|_{L^2},\tag{5-2}$$

$$\Big\|\frac{\delta\alpha(\kappa;q)}{\delta q}-\frac{\delta\alpha(\kappa;\tilde q)}{\delta q}\Big\|_{H^1}+\Big\|\frac{\delta\alpha(\kappa;q)}{\delta\bar q}-\frac{\delta\alpha(\kappa;\tilde q)}{\delta\bar q}\Big\|_{H^1}\lesssim\kappa\|q-\tilde q\|_{L^2},\tag{5-3}$$

where the implicit constants depend only on Q. Additionally, for every $\kappa\ge\kappa_0$ and $q\in Q_\varepsilon$, there exists $\gamma(\kappa;q)\in H^1$ such that

$$\Big(\frac{\delta\alpha(\kappa;q)}{\delta\bar q}\Big)'=2\kappa\frac{\delta\alpha(\kappa;q)}{\delta\bar q}-i\kappa q[\gamma(\kappa;q)+1],\tag{5-4}$$

$$\Big(\frac{\delta\alpha(\kappa;q)}{\delta q}\Big)'=-2\kappa\frac{\delta\alpha(\kappa;q)}{\delta q}+i\kappa\bar q[\gamma(\kappa;q)+1],\tag{5-5}$$

$$\gamma(\kappa;q)'=2\bar q\frac{\delta\alpha(\kappa;q)}{\delta\bar q}-2q\frac{\delta\alpha(\kappa;q)}{\delta q}.\tag{5-6}$$

Lastly, for each integer $m\ge 0$ we have

$$\Big\|\Big(\frac{\delta\alpha(\kappa;q)}{\delta\bar q}\Big)'\Big\|_{H^m}\lesssim_m\kappa\|q\|_{H^m},\tag{5-7}$$

$$\Big\|\langle x\rangle^{2m}\Big(\frac{\delta\alpha(\kappa;q)}{\delta\bar q}\Big)'\Big\|_{L^2}\lesssim_m\kappa\|\langle x\rangle^{2m}q\|_{L^2},\tag{5-8}$$

uniformly for $q\in Q_\varepsilon$ and $\kappa\ge\kappa_0$.

Proof. Proposition 2.6 shows that given $\delta\in(0,1]$, there exists $\kappa_0\ge1$ such that

$$\sup_{q\in Q}\sqrt{\kappa}\,\|\Lambda(q)\|_{\mathrm{op}}\le\frac{\delta}{4}\quad\text{uniformly for }\kappa\ge\kappa_0.$$

As $\Lambda(q)$ is linear in q, Lemma 2.1 allows us to deduce

$$\sup_{q\in Q_\varepsilon}\sqrt{\kappa}\,\|\Lambda(q)\|_{\mathrm{op}}\le\frac{\delta}{2}\quad\text{uniformly for }\kappa\ge\kappa_0,\tag{5-9}$$

provided ε is chosen sufficiently small (depending on δ).

Now we must explain how to choose δ. In view of (2-15), $\delta \le 1$ guarantees that the series (1-9) converges on Q_ε. We place an additional requirement to aid in the proofs of (5-2) and (5-3). From Lemma 2.5 we find that

$$\|(\kappa+\partial)^{-1/4} q (\kappa-\partial)^{-1/4}\|_{\mathrm{op}} \cdot \|(\kappa-\partial)^{-3/4} q (\kappa+\partial)^{-3/4}\|_{\mathrm{op}} \lesssim \|q\|_{L^2} \cdot \kappa^{-1/2} \|\Lambda(q)\|_{\mathrm{op}}.$$

Thus, we may choose δ even smaller if necessary to ensure also that

$$\kappa \|(\kappa+\partial)^{-1/4} q (\kappa-\partial)^{-1/4}\|_{\mathrm{op}} \cdot \|(\kappa-\partial)^{-3/4} q (\kappa+\partial)^{-3/4}\|_{\mathrm{op}} \le \tfrac{1}{2} \tag{5-10}$$

uniformly for $q \in Q_\varepsilon$ and $\kappa \ge \kappa_0$.

Turning now to (5-2), we argue by duality. For $f \in H^{-1}$, we have

$$\left\langle f, \frac{\delta\alpha(\kappa;q)}{\delta\bar q}\right\rangle = \sum_{\ell\ge 0} (i\kappa)^{\ell+1} \operatorname{tr}\{[(\kappa-\partial)^{-1} q (\kappa+\partial)^{-1}\bar q]^\ell (\kappa-\partial)^{-1} q (\kappa+\partial)^{-1} \bar f\}.$$

The $\ell = 0$ term is readily computed exactly via Lemma 2.2. For example,

$$i\kappa \operatorname{tr}\{(\kappa-\partial)^{-1} q (\kappa+\partial)^{-1}\bar f\} = i\kappa \left\langle \frac{1}{2\kappa+\partial} f, q\right\rangle \quad \text{in the line case.}$$

In either geometry, this is easily seen to satisfy the desired bound.

For $\ell \ge 1$, we employ (5-10) and Lemma 2.5 to estimate

$$\begin{aligned}
&\left|(i\kappa)^{\ell+1} \operatorname{tr}\{[(\kappa-\partial)^{-1} q (\kappa+\partial)^{-1}\bar q]^\ell (\kappa-\partial)^{-1} q (\kappa+\partial)^{-1}\bar f\}\right| \\
&\quad \lesssim \kappa^{\ell+1} \|(\kappa+\partial)^{-1} \bar f (\kappa-\partial)^{-1}\|_{\mathfrak{I}_2} \|q(\kappa+\partial)^{-3/4}\|_{\mathfrak{I}_2}^2 \|(\kappa+\partial)^{-1/4} q (\kappa-\partial)^{-1/4}\|_{\mathrm{op}}^\ell \\
&\qquad\qquad \times \|(\kappa-\partial)^{-3/4} q (\kappa+\partial)^{-3/4}\|_{\mathrm{op}}^{\ell-1} \\
&\quad \lesssim 2^{-\ell} \kappa \|f\|_{H^{-1}} \|q\|_{L^2}^3,
\end{aligned}$$

with an implicit constant independent of ℓ. This proves that the estimate (5-2) holds for the $\bar q$ derivative; the bound on the q derivative follows in a parallel fashion.

The proof of (5-3) proceeds analogously, noting that one can always exhibit the difference $q - \tilde q$ in place of a q.

We define $\gamma(\kappa;q)$ via the associated linear functional

$$\begin{aligned}
\langle f, \gamma(\kappa;q)\rangle = &\sum_{\ell\ge 1} (i\kappa)^\ell \operatorname{tr}\{[(\kappa-\partial)^{-1} q (\kappa+\partial)^{-1}\bar q]^\ell (\kappa-\partial)^{-1}\bar f\} \\
&+ \sum_{\ell\ge 1} (i\kappa)^\ell \operatorname{tr}\{[(\kappa+\partial)^{-1}\bar q (\kappa-\partial)^{-1} q]^\ell (\kappa+\partial)^{-1}\bar f\},
\end{aligned} \tag{5-11}$$

and will prove $\gamma \in H^1$ by showing that this functional is bounded for $f \in H^{-1}$.

Regarding the $\ell = 1$ terms, Lemma 2.5 and direct computation show that

$$\left|\kappa \operatorname{tr}\{(\kappa\mp\partial)^{-1} q (\kappa\pm\partial)^{-1}\bar q (\kappa\mp\partial)^{-1}\bar f\}\right| \lesssim \sqrt{\kappa}\, \|q(\kappa\pm\partial)^{-1}\bar q\|_{\mathfrak{I}_2} \|f\|_{H^{-1}} \lesssim \sqrt{\kappa}\, \|q\|_{L^2}^2 \|f\|_{H^{-1}}.$$

For $\ell \geq 2$, we employ Lemma 2.5 and (5-10) as follows:

$$\begin{aligned}
&\left|\kappa^{\ell}\operatorname{tr}\{[(\kappa\mp\partial)^{-1}q(\kappa\pm\partial)^{-1}\bar q]^{\ell}(\kappa\mp\partial)^{-1}f\}\right| \\
&\quad\lesssim \kappa^{\ell-1/2}\|f\|_{H^{-1}}\|q(\kappa+\partial)^{-3/4}\|_{\mathfrak{I}_2}^2\|(\kappa-\partial)^{-1/2}\|_{\mathrm{op}}\|(\kappa+\partial)^{-1/4}q(\kappa-\partial)^{-1/4}\|_{\mathrm{op}}^{\ell} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\times\|(\kappa-\partial)^{-3/4}q(\kappa+\partial)^{-3/4}\|_{\mathrm{op}}^{\ell-2} \\
&\quad\lesssim 2^{-\ell}\sqrt{\kappa}\,\|q\|_{L^2}^4\|f\|_{H^{-1}},
\end{aligned}$$

where the implicit constant is independent of ℓ. Thus $\gamma \in H^1$ and

$$\|\gamma(\kappa;q)\|_{H^1} \lesssim \sqrt{\kappa}\|q\|_{L^2}^2.$$

The proofs of (5-4) and (5-5) follow parallel arguments. In the former case, we pair $\delta\alpha(\kappa;q)/\delta\bar q$ with f', which we then rewrite as a trace. The result then follows by noting the operator identity $f' = -(\kappa-\partial)f - f(\kappa+\partial) + 2\kappa f$ and simplifying.

The proof of (5-6) follows the same style: one pairs $\gamma(\kappa;q)$ with f' and employs the operator identity $f' = [\kappa+\partial, f] = -[\kappa-\partial, f]$.

The proof of (5-7) mimics closely that of (5-2), once one understands how to move the derivatives from the test function f to copies of q. Introducing the notation $f_h(x) = f(x-h)$, we observe that by the translation invariance of the trace,

$$\begin{aligned}
\left\langle f^{(m)}, \left(\frac{\delta\alpha(\kappa;q)}{\delta\bar q}\right)'\right\rangle &= -\frac{\partial^m}{\partial h^m}\bigg|_{h=0}\left\langle f_h', \frac{\delta\alpha(\kappa;q)}{\delta\bar q}\right\rangle = -\frac{\partial^m}{\partial h^m}\bigg|_{h=0}\left\langle f', \frac{\delta}{\delta\bar q}\alpha(\kappa;q_{-h})\right\rangle \\
&= -\frac{\partial^m}{\partial h^m}\bigg|_{h=0}\sum_{\ell\geq 0}(i\kappa)^{\ell+1}\operatorname{tr}\{[(\kappa-\partial)^{-1}q_{-h}(\kappa+\partial)^{-1}\bar q_{-h}]^{\ell}(\kappa-\partial)^{-1}q_{-h}(\kappa+\partial)^{-1}\bar f'\}.
\end{aligned}$$

Next, we apply the estimates used to prove (5-2) together with the elementary inequality

$$\|q^{(n)}\|_{L^2} \lesssim \|q\|_{L^2}^{1-n/m}\|q\|_{H^m}^{n/m} \quad \text{for all } 0 < n \leq m.$$

This yields the estimate (5-7). Note that summability in ℓ is guaranteed by (5-10), just as before.

Lastly, we turn to (5-8). The argument is very similar; the key ingredient is to move the polynomial weight $\langle x\rangle^{2m}$ from the test function f to a copy of q. This is achieved via the identity

$$q(\kappa+\partial)^{-1}P\bar f = \sum_{n\geq 0}(-1)^n[P^{(n)}q](\kappa+\partial)^{-n-1}\bar f,$$

valid for any polynomial $P(x)$, which follows easily by induction using

$$[(\kappa+\partial)^{-1}, P(x)] = -(\kappa+\partial)^{-1}P'(x)(\kappa+\partial)^{-1}. \qquad \square$$

Lemma 5.2. *Let $Q \subseteq \mathcal{S}$ be L^2-bounded and equicontinuous, and let ε and κ_0 be as in Lemma 5.1. Then for all $\kappa,\ \varkappa \geq \kappa_0$,*

$$\{H, \alpha(\kappa)\} = 0, \quad \{M, \alpha(\kappa)\} = 0, \quad \text{and} \quad \{\alpha(\varkappa), \alpha(\kappa)\} = 0$$

on Q_ε. Consequently, $A(\kappa)$, $A(\varkappa)$, M, H, and H_κ all Poisson commute on Q_ε.

Proof. As discussed in the introduction, the commutativity of $\alpha(\kappa)$ with the Hamiltonian H was proved in [Klaus and Schippa 2022; Tang and Xu 2021] whenever the series defining $\alpha(\kappa)$ can be guaranteed to converge. Such convergence is guaranteed by Lemma 5.1.

Recalling (1-5) and employing (5-4), (5-5), and (5-6), we find

$$\{M, \alpha(\kappa)\} = -2\kappa \int \gamma' \, dx = 0.$$

Notice that (5-6) guarantees $\gamma' \in L^1$.

If $\kappa = \varkappa$ the third equality is clear. When $\kappa \neq \varkappa$, we may proceed to compute the Poisson bracket by applying (5-4) and (5-5) directly to the derivatives of $\alpha(\kappa)$ or by employing integration by parts and then the corresponding formulae for the partial derivatives of $\alpha(\varkappa)$. Comparing the two approaches yields

$$\varkappa\{\alpha(\kappa), \alpha(\varkappa)\} = \kappa\{\alpha(\kappa), \alpha(\varkappa)\}, \quad \text{and so } \{\alpha(\varkappa), \alpha(\kappa)\} = 0. \qquad \square$$

Proposition 5.3. *For each L^2-equicontinuous set $Q \subseteq \mathcal{S}$ satisfying* (1-13), *there exists $\kappa_0 \geq 1$ sufficiently large such that for all $\kappa \geq \kappa_0$, the (H_κ) flow is globally well posed for initial data in Q. Moreover, the solutions remain in $\mathcal{S}$ for all time. Lastly, the set*

$$Q_* := \{e^{tJ\nabla H_\kappa} q : q \in Q,\ t \in \mathbb{R},\ \textit{and } \kappa \geq \kappa_0\}$$

is bounded and equicontinuous in L^2.

Proof. Recall the set Q_{**} introduced in (1-11). By (3-9), the hypothesis (1-13), and the definition of M_*, this set is bounded and equicontinuous in L^2. We fix $\varepsilon > 0$ and $\kappa_0 \geq 1$ as the values obtained by applying Lemma 5.1 to the set Q_{**}.

Next we construct a local solution for initial data $q(0) \in Q_{**}$. For $\kappa \geq \kappa_0$, Lemma 5.1 ensures that one can run the usual contraction mapping argument for the integral equation

$$q(t) = q(0) + \int_0^t 2\kappa \left(\frac{\delta A(\kappa; q(s))}{\delta \bar{q}} + \overline{\frac{\delta A(\kappa; q(s))}{\delta q}} \right)' ds$$

to find a unique solution $q \in C([0, T]; L^2)$, provided T is chosen sufficiently small. In fact, T is chosen so small that $q(t)$ and indeed all Picard iterates remain in the ε-neighborhood of Q_{**}.

Combining the estimates (5-7) and (5-8) with the Gronwall inequality shows that $q(t) \in \mathcal{S}$ for all $t \in [0, T]$. This in turn allows us to apply Lemma 5.2 to conclude that $\alpha(\varkappa; q(t))$ and hence $a(\varkappa; q(t))$ are conserved. Taken together, these observations guarantee that $q([0, T]) \subseteq Q_{**}$ and so the local solutions may be concatenated to yield a global solution lying wholly within Q_{**}. Finally, as Q_* is a subset of Q_{**}, it is L^2-bounded and equicontinuous. $\square$

Combining Proposition 5.3 with Theorem 4.3 immediately yields well-posedness of the (H_κ) flow in the following sense:

Corollary 5.4. *Fix $0 < s < \frac{1}{2}$ and let $Q \subseteq \mathcal{S}$ be H^s-bounded and satisfy* (1-13). *Then there exists $\kappa_0 \geq 1$ such that for all $\kappa \geq \kappa_0$ the (H_κ) flow is globally well posed for initial data in Q. Moreover,*

$$Q_* := \{e^{tJ\nabla H_\kappa} q : q \in Q,\ t \in \mathbb{R},\ \textit{and } \kappa \geq \kappa_0\} \subseteq \mathcal{S} \quad \textit{is } H^s\textit{-bounded.}$$

If Q is H^s-equicontinuous, then so is Q_*.

In order to complete the proof of Theorem 1.5, we must prove that H^s-Cauchy sequences of initial data $q_n(0) \in \mathcal{S}$ satisfying (1-13) lead to Cauchy sequences of solutions to (DNLS). As mentioned above, this will be accomplished by showing that the flow

$$\frac{d}{dt}q = \left[iq' - |q|^2 q - 2\kappa\left(\frac{\delta A(\kappa; q)}{\delta \bar{q}} + \overline{\frac{\delta A(\kappa; q)}{\delta q}}\right)\right]', \tag{H_κ^{diff}}$$

generated by $H(q) - H_\kappa(q)$, converges to the identity as $\kappa \to \infty$. Due to commutativity of the flows, $\mathcal{S}$-valued solutions to (H_κ^{diff}) can be built via

$$e^{tJ\nabla H_\kappa^{\text{diff}}} q = e^{tJ\nabla H} e^{-tJ\nabla H_\kappa} q$$

using Corollaries 4.2 and 5.4. In view of Lemma 5.2, these solutions conserve M and $\alpha(\varkappa)$.

The proof of our final theorem makes a fitting end for this paper by highlighting the power of equicontinuity. It is also here that we will finally see the origin of the restriction $s \geq \frac{1}{6}$. It is needed to make sense of the nonlinearity in (H_κ^{diff}) pointwise in time.

Theorem 5.5. *Fix $\frac{1}{6} \leq s < \frac{1}{2}$ and $T > 0$. Given a sequence $q_n(0) \in \mathcal{S}$ of initial data that converges in H^s and satisfies* (1-13), *let $q_n(t)$ denote the corresponding solutions to* (DNLS). *Then $q_n(t)$ converges in H^s, uniformly for $|t| \leq T$.*

Proof. By hypothesis, the set $Q = \{q_n(0) : n \in \mathbb{N}\}$ is bounded and equicontinuous in the H^s-metric. Let $\kappa_0 \geq 1$ be as given by Corollary 5.4. Then for $\kappa \geq \kappa_0$, the (H_κ) flow is well posed for initial data in Q, and the set

$$Q_* := \{e^{tJ\nabla H_\kappa} q_n(0) : n \in \mathbb{N},\ t \in \mathbb{R},\ \text{and } \kappa \geq \kappa_0\} \subseteq \mathcal{S}$$

is bounded and equicontinuous in H^s.

The commutativity of the (H_κ) and the (DNLS) flows allows us to rewrite our sequence of solutions as

$$q_n(t) = e^{tJ\nabla H_\kappa^{\text{diff}}} e^{tJ\nabla H_\kappa} q_n(0).$$

Moreover, by Theorem 4.3, the set

$$\{e^{tJ\nabla H_\kappa^{\text{diff}}} q : q \in Q_*,\ t \in \mathbb{R},\ \text{and } \kappa \geq \kappa_0\} \subseteq Q_{**}$$

is bounded and equicontinuous in H^s.

We will show that $q_n(t)$ forms a Cauchy sequence in H^s, uniformly for $|t| \leq T$. By the definition of Q_*, we estimate

$$\sup_{|t|\leq T} \|q_n(t) - q_m(t)\|_{H^s} \leq 2 \sup_{q\in Q_*} \sup_{|t|\leq T} \|e^{tJ\nabla H_\kappa^{\text{diff}}} q - q\|_{H^s} + \sup_{|t|\leq T} \|e^{tJ\nabla H_\kappa} q_n(0) - e^{tJ\nabla H_\kappa} q_m(0)\|_{H^s}$$

for all $\kappa \geq \kappa_0$. For any such fixed κ, the well-posedness of the (H_κ) flow ensures that the last term of the right-hand side converges to 0 as $n,\ m \to \infty$. Thus, it suffices to prove that the difference flow converges to the identity uniformly on Q_*:

$$\lim_{\kappa\to\infty} \sup_{q\in Q_*} \sup_{|t|\leq T} \|e^{tJ\nabla H_\kappa^{\text{diff}}} q - q\|_{H^s} = 0.$$

In fact, as Q_{**} is H^s-equicontinuous, it suffices to show that

$$\lim_{\kappa\to\infty} \sup_{q\in Q_*} \sup_{|t|\le T} \|e^{tJ\nabla H_\kappa^{\mathrm{diff}}} q - q\|_{H^{-4}} = 0. \tag{5-12}$$

By the fundamental theorem of calculus and (H_κ^{diff}), proving (5-12) reduces to showing that

$$\lim_{\kappa\to\infty} \sup_{q\in Q_{**}} \|F\|_{H^{-3}} = 0, \quad \text{where} \quad F := iq' - |q|^2 q - 2\kappa\left(\frac{\delta A(\kappa;q)}{\delta \bar q} + \overline{\frac{\delta A(\kappa;q)}{\delta q}}\right). \tag{5-13}$$

A straightforward computation shows that $F^{[1]}$, the linear (in q) term in F, is given by $-i\partial^3/(4\kappa^2-\partial^2)q$. This clearly converges to zero in H^{-3} as $\kappa\to\infty$, uniformly on Q_{**}, or indeed, on any L^2-bounded set.

We turn now to the contribution of $F^{[3]}$, the term in F that is cubic in q. Employing Lemma 2.3, we find the cubic terms

$$\left(\frac{\delta A(\kappa;q)}{\delta \bar q}\right)^{[3]} = -\frac{\kappa^2}{2\kappa-\partial}\left\{\left(\frac{1}{2\kappa+\partial}\bar q\right)(4\kappa-\partial)\left(\frac{1}{2\kappa-\partial}q\right)^2\right\} = -\frac{2\kappa^2}{2\kappa-\partial}\left\{q\left(\frac{1}{2\kappa-\partial}q\right)\left(\frac{1}{2\kappa+\partial}\bar q\right)\right\},$$

$$\left(\overline{\frac{\delta A(\kappa;q)}{\delta q}}\right)^{[3]} = -\frac{\kappa^2}{2\kappa+\partial}\left\{\left(\frac{1}{2\kappa-\partial}\bar q\right)(4\kappa+\partial)\left(\frac{1}{2\kappa+\partial}q\right)^2\right\} = -\frac{2\kappa^2}{2\kappa+\partial}\left\{q\left(\frac{1}{2\kappa+\partial}q\right)\left(\frac{1}{2\kappa-\partial}\bar q\right)\right\}.$$

This allows us to compute the full cubic term as follows:

$$\begin{aligned}
F^{[3]} &= 2\kappa^2\frac{\partial}{2\kappa-\partial}\left[q\left(\frac{1}{2\kappa-\partial}q\right)\left(\frac{1}{2\kappa+\partial}\bar q\right)\right] - 2\kappa^2\frac{\partial}{2\kappa+\partial}\left[q\left(\frac{1}{2\kappa+\partial}q\right)\left(\frac{1}{2\kappa-\partial}\bar q\right)\right] \\
&\qquad + 2\kappa^2 q\left(\frac{1}{2\kappa-\partial}q\right)\left(\frac{1}{2\kappa+\partial}\bar q\right) + 2\kappa^2 q\left(\frac{1}{2\kappa+\partial}q\right)\left(\frac{1}{2\kappa-\partial}\bar q\right) - q^2\bar q \\
&= \frac{2\partial}{2\kappa-\partial}\left[q\left(\frac{\kappa}{2\kappa-\partial}q\right)\left(\frac{\kappa}{2\kappa+\partial}\bar q\right)\right] - \frac{2\partial}{2\kappa+\partial}\left[q\left(\frac{\kappa}{2\kappa+\partial}q\right)\left(\frac{\kappa}{2\kappa-\partial}\bar q\right)\right] + q^2\left(\frac{\partial^2}{4\kappa^2-\partial^2}\bar q\right) \\
&\qquad + q\bar q\left(\frac{\partial^2}{4\kappa^2-\partial^2}q\right) - \frac{1}{2}q\left(\frac{\partial}{2\kappa-\partial}q\right)\left(\frac{\partial}{2\kappa+\partial}\bar q\right) - \frac{1}{2}q\left(\frac{\partial}{2\kappa+\partial}q\right)\left(\frac{\partial}{2\kappa-\partial}\bar q\right).
\end{aligned}$$

To estimate its contribution, we pair with $f\in H^3$ and apply Hölder's inequality. Boundedness is easily deduced from

$$\|f\|_{L^\infty} + \kappa\left\|\frac{\partial}{2\kappa\pm\partial}f\right\|_{L^\infty} \lesssim \|f\|_{H^3}, \tag{5-14}$$

$$\|q\|_{L^3} + \left\|\frac{\partial}{2\kappa\pm\partial}q\right\|_{L^3} + \left\|\frac{\partial^2}{4\kappa^2-\partial^2}q\right\|_{L^3} \lesssim \|q\|_{H^s}. \tag{5-15}$$

Evidently (5-15) requires $s\ge\frac{1}{6}$. The gain of a power of κ in (5-14) guarantees that the contribution of the first two terms in $F^{[3]}$ decays to zero as $\kappa\to\infty$. For the remaining terms, we use H^s-equicontinuity to obtain decay: as $s\ge\frac{1}{6}$, we have that

$$\lim_{\kappa\to\infty}\sup_{q\in Q_{**}}\left\|\frac{\partial}{2\kappa\pm\partial}q\right\|_{L^3} \lesssim \lim_{\kappa\to\infty}\sup_{q\in Q_{**}}\left\|\frac{\partial}{2\kappa\pm\partial}q\right\|_{H^s} = 0.$$

Finally, we turn our attention to the remaining terms (quintic and higher) in the series expansion of F. By Lemma 2.1, (1-9), and the embedding $H^3 \hookrightarrow L^\infty$,

$$\left|\int f F^{[\geq 5]}\,dx\right| \lesssim \sum_{\ell\geq 2} \kappa^{\ell+2} \|\Lambda(q)\|_{\mathfrak{I}_2}^2 \|\Lambda(q)\|_{\mathrm{op}}^{2\ell-1} \|\Lambda(f)\|_{\mathrm{op}} \lesssim \|q\|_{L^2}^2 \|f\|_{H^3} \sum_{\ell\geq 2} \kappa^{\ell} \|\Lambda(q)\|_{\mathrm{op}}^{2\ell-1}.$$

The convergence we require does not follow from Proposition 2.6; we would lose by a factor of $\sqrt{\kappa}$. However arguing in the same fashion, we find

$$\begin{aligned}\|\Lambda(q)\|_{\mathrm{op}} \lesssim \|\Lambda(q_{\leq\eta\kappa})\|_{\mathrm{op}} + \|\Lambda(q_{>\eta\kappa})\|_{\mathrm{op}} &\lesssim \kappa^{-1}\|q_{\leq\eta\kappa}\|_{L^\infty} + \kappa^{-1/2}\|q_{>\eta\kappa}\|_{L^2} \\ &\lesssim \kappa^{-1/2-s}(\eta^{1/2-s}\|q\|_{H^s} + \eta^{-s}\|q_{>\eta\kappa}\|_{H^s}),\end{aligned}$$

for any $\eta > 0$. When $s > \frac{1}{6}$, we may simply take $\eta = 1$ to deduce that

$$\lim_{\kappa\to\infty} \sup_{q\in Q_{**}} \|F^{[5]}\|_{H^{-3}} = 0.$$

For the endpoint case $s = \frac{1}{6}$, this follows from the H^s-equicontinuity of Q_{**} by choosing η small and then κ large. This completes the proof of the theorem. $\square$

References

[Arnold 1989] V. I. Arnold, *Mathematical methods of classical mechanics*, 2nd ed., Grad. Texts in Math. **60**, Springer, 1989. MR Zbl

[Bahouri and Perelman 2022] H. Bahouri and G. Perelman, "Global well-posedness for the derivative nonlinear Schrödinger equation", *Invent. Math.* **229**:2 (2022), 639–688. MR Zbl

[Biagioni and Linares 2001] H. A. Biagioni and F. Linares, "Ill-posedness for the derivative Schrödinger and generalized Benjamin–Ono equations", *Trans. Amer. Math. Soc.* **353**:9 (2001), 3649–3659. MR Zbl

[Bringmann et al. 2021] B. Bringmann, R. Killip, and M. Vişan, "Global well-posedness for the fifth-order KdV equation in $H^{-1}(\mathbb{R})$", *Ann. PDE* **7**:2 (2021), art. id. 21. MR Zbl

[Colliander et al. 2001] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "Global well-posedness for Schrödinger equations with derivative", *SIAM J. Math. Anal.* **33**:3 (2001), 649–669. MR Zbl

[Colliander et al. 2002] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "A refined global well-posedness result for Schrödinger equations with derivative", *SIAM J. Math. Anal.* **34**:1 (2002), 64–86. MR Zbl

[Cwikel 1977] M. Cwikel, "Weak type estimates for singular values and the number of bound states of Schrödinger operators", *Ann. of Math.* (2) **106**:1 (1977), 93–100. MR Zbl

[Deng et al. 2021] Y. Deng, A. R. Nahmod, and H. Yue, "Optimal local well-posedness for the periodic derivative nonlinear Schrödinger equation", *Comm. Math. Phys.* **384**:2 (2021), 1061–1107. MR Zbl

[Fujiwara et al. 2020] K. Fujiwara, V. Georgiev, and T. Ozawa, "Self-similar solutions to the derivative nonlinear Schrödinger equation", *J. Differential Equations* **268**:12 (2020), 7940–7961. MR Zbl

[Fukaya et al. 2017] N. Fukaya, M. Hayashi, and T. Inui, "A sufficient condition for global existence of solutions to a generalized derivative nonlinear Schrödinger equation", *Anal. PDE* **10**:5 (2017), 1149–1167. MR Zbl

[Grünrock 2005] A. Grünrock, "Bi- and trilinear Schrödinger estimates in one space dimension with applications to cubic NLS and DNLS", *Int. Math. Res. Not.* **2005**:41 (2005), 2525–2558. MR Zbl

[Grünrock and Herr 2008] A. Grünrock and S. Herr, "Low regularity local well-posedness of the derivative nonlinear Schrödinger equation with periodic initial data", *SIAM J. Math. Anal.* **39**:6 (2008), 1890–1920. MR Zbl

[Guo and Wu 2017] Z. Guo and Y. Wu, "Global well-posedness for the derivative nonlinear Schrödinger equation in $H^{1/2}(\mathbb{R})$", *Discrete Contin. Dyn. Syst.* **37**:1 (2017), 257–264. MR Zbl

[Harrop-Griffiths et al. 2020] B. Harrop-Griffiths, R. Killip, and M. Vişan, "Sharp well-posedness for the cubic NLS and mKdV in $H^s(\mathbb{R})$", preprint, 2020. arXiv 2003.05011

[Hayashi and Ozawa 1992] N. Hayashi and T. Ozawa, "On the derivative nonlinear Schrödinger equation", *Phys. D* **55**:1-2 (1992), 14–36. MR Zbl

[Herr 2006] S. Herr, "On the Cauchy problem for the derivative nonlinear Schrödinger equation with periodic boundary condition", *Int. Math. Res. Not.* **2006** (2006), art. id. 96763. MR Zbl

[Ichikawa and Watanabe 1977] Y. H. Ichikawa and S. Watanabe, "Solitons, envelope solitons in collisionless plasmas", *J. Phys. Colloques* **38**:C6 (1977), 15–26.

[Isom et al. 2020] B. Isom, D. Mantzavinos, and A. Stefanov, "Growth bound and nonlinear smoothing for the periodic derivative nonlinear Schrödinger equation", preprint, 2020. arXiv 2012.09933

[Jenkins et al. 2018a] R. Jenkins, J. Liu, P. Perry, and C. Sulem, "Soliton resolution for the derivative nonlinear Schrödinger equation", *Comm. Math. Phys.* **363**:3 (2018), 1003–1049. MR Zbl

[Jenkins et al. 2018b] R. Jenkins, J. Liu, P. A. Perry, and C. Sulem, "Global well-posedness for the derivative non-linear Schrödinger equation", *Comm. Partial Differential Equations* **43**:8 (2018), 1151–1195. MR Zbl

[Jenkins et al. 2020a] R. Jenkins, J. Liu, P. Perry, and C. Sulem, "The derivative nonlinear Schrödinger equation: global well-posedness and soliton resolution", *Quart. Appl. Math.* **78**:1 (2020), 33–73. MR Zbl

[Jenkins et al. 2020b] R. Jenkins, J. Liu, P. Perry, and C. Sulem, "Global existence for the derivative nonlinear Schrödinger equation with arbitrary spectral singularities", *Anal. PDE* **13**:5 (2020), 1539–1578. MR Zbl

[Kaup and Newell 1978] D. J. Kaup and A. C. Newell, "An exact solution for a derivative nonlinear Schrödinger equation", *J. Math. Phys.* **19**:4 (1978), 798–801. MR Zbl

[Killip and Vişan 2019] R. Killip and M. Vişan, "KdV is well-posed in H^{-1}", *Ann. of Math.* (2) **190**:1 (2019), 249–305. MR Zbl

[Killip et al. 2018] R. Killip, M. Vişan, and X. Zhang, "Low regularity conservation laws for integrable PDE", *Geom. Funct. Anal.* **28**:4 (2018), 1062–1090. MR Zbl

[Killip et al. 2020] R. Killip, J. Murphy, and M. Vişan, "Invariance of white noise for KdV on the line", *Invent. Math.* **222**:1 (2020), 203–282. MR Zbl

[Kitaev 1985] A. V. Kitaev, "Self-similar solutions of a modified nonlinear Schrödinger equation", *Teoret. Mat. Fiz.* **64**:3 (1985), 347–369. In Russian; translated in *Theoret. and Math. Phys.* **64**:3 (1986), 878–894. MR Zbl

[Klaus and Schippa 2022] F. Klaus and R. Schippa, "A priori estimates for the derivative nonlinear Schrödinger equation", *Funkcial. Ekvac.* **65**:3 (2022), 329–346. MR Zbl

[Liu 2017] J. Liu, *Global well-posedness for the derivative nonlinear Schrodinger equation through inverse scattering*, Ph.D. thesis, University of Kentucky, 2017, available at https://uknowledge.uky.edu/math_etds/50/.

[Liu et al. 2016] J. Liu, P. A. Perry, and C. Sulem, "Global existence for the derivative nonlinear Schrödinger equation by the method of inverse scattering", *Comm. Partial Differential Equations* **41**:11 (2016), 1692–1760. MR Zbl

[Liu et al. 2018] J. Liu, P. A. Perry, and C. Sulem, "Long-time behavior of solutions to the derivative nonlinear Schrödinger equation for soliton-free initial data", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **35**:1 (2018), 217–265. MR Zbl

[Miao et al. 2011] C. Miao, Y. Wu, and G. Xu, "Global well-posedness for Schrödinger equation with derivative in $H^{1/2}(\mathbb{R})$", *J. Differential Equations* **251**:8 (2011), 2164–2195. MR Zbl

[Mio et al. 1976] K. Mio, T. Ogino, K. Minami, and S. Takeda, "Modulational instability and envelope-solitons for nonlinear Alfvén waves propagating along the magnetic field in plasmas", *J. Phys. Soc. Japan* **41**:2 (1976), 667–673.

[Mjølhus 1976] E. Mjølhus, "On the modulational instability of hydromagnetic waves parallel to the magnetic field", *J. Plasma Phys.* **16**:3 (1976), 321–334.

[Mosincat 2017] R. Mosincat, "Global well-posedness of the derivative nonlinear Schrödinger equation with periodic boundary condition in $H^{1/2}$", *J. Differential Equations* **263**:8 (2017), 4658–4722. MR Zbl

[Mosincat and Oh 2015] R. Mosincat and T. Oh, “A remark on global well-posedness of the derivative nonlinear Schrödinger equation on the circle”, *C. R. Math. Acad. Sci. Paris* **353**:9 (2015), 837–841. MR Zbl

[Nahmod et al. 2012] A. R. Nahmod, T. Oh, L. Rey-Bellet, and G. Staffilani, “Invariant weighted Wiener measures and almost sure global well-posedness for the periodic derivative NLS”, *J. Eur. Math. Soc.* **14**:4 (2012), 1275–1330. MR Zbl

[Ntekoume 2022] M. Ntekoume, “Symplectic nonsqueezing for the KdV flow on the line”, *Pure Appl. Anal.* **4**:3 (2022), 401–448. MR Zbl

[Pelinovsky and Shimabukuro 2018] D. E. Pelinovsky and Y. Shimabukuro, “Existence of global solutions to the derivative NLS equation with the inverse scattering transform method”, *Int. Math. Res. Not.* **2018**:18 (2018), 5663–5728. MR Zbl

[Pelinovsky et al. 2017] D. E. Pelinovsky, A. Saalmann, and Y. Shimabukuro, “The derivative NLS equation: global existence with solitons”, *Dyn. Partial Differ. Equ.* **14**:3 (2017), 271–294. MR Zbl

[Saalmann 2017] A. Saalmann, “Global existence for the derivative NLS equation in the presence of solitons”, preprint, 2017. Zbl arXiv 1704.00071

[Simon 2005] B. Simon, *Trace ideals and their applications*, 2nd ed., Math. Surv. Monogr. **120**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Takaoka 1999] H. Takaoka, “Well-posedness for the one-dimensional nonlinear Schrödinger equation with the derivative nonlinearity”, *Adv. Differential Equations* **4**:4 (1999), 561–580. MR Zbl

[Takaoka 2001] H. Takaoka, “Global well-posedness for Schrödinger equations with derivative in a nonlinear term and data in low-order Sobolev spaces”, *Electron. J. Differential Equations* **2001** (2001), art. id. 42. MR Zbl

[Tang and Xu 2021] X. Tang and G. Xu, “Microscopic conservation laws for the derivative nonlinear Schrödinger equation”, *Lett. Math. Phys.* **111**:6 (2021), art. id. 138. MR Zbl

[Tsutsumi and Fukuda 1980] M. Tsutsumi and I. Fukuda, “On solutions of the derivative nonlinear Schrödinger equation: existence and uniqueness theorem”, *Funkcial. Ekvac.* **23**:3 (1980), 259–277. MR Zbl

[Tsutsumi and Fukuda 1981] M. Tsutsumi and I. Fukuda, “On solutions of the derivative nonlinear Schrödinger equation, II”, *Funkcial. Ekvac.* **24**:1 (1981), 85–94. MR Zbl

[Win 2010] Y. Y. S. Win, “Global well-posedness of the derivative nonlinear Schrödinger equations on T”, *Funkcial. Ekvac.* **53**:1 (2010), 51–88. MR Zbl

[Wu 2013] Y. Wu, “Global well-posedness for the nonlinear Schrödinger equation with derivative in energy space”, *Anal. PDE* **6**:8 (2013), 1989–2002. MR Zbl

[Wu 2015] Y. Wu, “Global well-posedness on the derivative nonlinear Schrödinger equation”, *Anal. PDE* **8**:5 (2015), 1101–1112. MR Zbl

Received 29 Jan 2021. Revised 28 Jun 2021. Accepted 6 Oct 2021.

ROWAN KILLIP: `killip@math.ucla.edu`
Department of Mathematics, University of California, Los Angeles, Los Angeles, CA, United States

MARIA NTEKOUME: `maria.ntekoume@rice.edu`
Department of Mathematics, Rice University, Houston, TX, United States

MONICA VIŞAN: `visan@math.ucla.edu`
Math Sciences, University of California, Los Angeles, Los Angeles, CA, United States

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 16 (2023), No. 5, pp. 1271–1288
DOI: 10.2140/apde.2023.16.1271

EXPONENTIAL INTEGRABILITY IN GAUSS SPACE

PAATA IVANISVILI AND RYAN RUSSELL

Talagrand showed that finiteness of $\mathbb{E}\, e^{|\nabla f(X)|^2/2}$ implies finiteness of $\mathbb{E}\, e^{f(X)-\mathbb{E} f(X)}$, where X is the standard Gaussian vector in $\mathbb{R}^n$ and f is a smooth function. However, in this paper we show that finiteness of $\mathbb{E}\, e^{|\nabla f|^2/2}(1+|\nabla f|)^{-1}$ implies finiteness of $\mathbb{E}\, e^{f(X)-\mathbb{E} f(X)}$, and we also obtain quantitative bounds

$$\log \mathbb{E}\, e^{f-\mathbb{E} f} \leq 10\, \mathbb{E}\, e^{|\nabla f|^2/2}(1+|\nabla f|)^{-1}.$$

Moreover, the extra factor $(1+|\nabla f|)^{-1}$ is the best possible in the sense that there is a smooth f with $\mathbb{E}\, e^{f-\mathbb{E} f} = \infty$ but $\mathbb{E}\, e^{|\nabla f|^2/2}(1+|\nabla f|)^{-c} < \infty$ for all $c > 1$. As an application we show corresponding dual inequalities for the discrete time dyadic martingales and their quadratic variations.

1. Introduction

Bobkov and Götze [1999] showed that for a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ with $\mathbb{E} f(X) = 0$ we have

$$\mathbb{E} e^{f(X)} \leq (\mathbb{E} e^{\alpha |\nabla f(X)|^2})^{1/(2\alpha-1)} \quad \text{for any } \alpha > \tfrac{1}{2}, \tag{1-1}$$

for the class of random vectors X in $\mathbb{R}^n$ satisfying the log-Sobolev inequality *with constant* 1. In particular, the estimate (1-1) holds true when $X \sim \mathcal{N}(0, I_{n\times n})$ is the standard Gaussian vector in $\mathbb{R}^n$ and $I_{n\times n}$ is the identity matrix. The inequality implies the measure concentration inequality $\mathbb{P}(f(X) > \lambda) \leq e^{-\lambda^2/2}$ for all $\lambda \geq 0$ provided that $|\nabla f| \leq 1$ and $\mathbb{E} f(X) = 0$. In [Bobkov and Götze 1999] it was asked what happens in the endpoint case when $\alpha = \frac{1}{2}$, i.e., does finiteness of $\mathbb{E} e^{|\nabla f(X)|^2/2}$ imply finiteness of $\mathbb{E} e^{f(X)}$ even for $n = 1$ and $X \sim \mathcal{N}(0, 1)$?

From the aforementioned paper, it is not hard to see that the Bobkov–Götze exponential inequality (1-1) is optimal in terms of the powers, i.e., one cannot replace $1/(2\alpha - 1)$ with $1/(c\alpha - 1)$ for some $c < 2$, and one cannot replace $e^{\alpha|\nabla f|^2}$ with $e^{c\alpha|\nabla f|^2}$ for some $c < 1$. Notice that the finiteness of $\mathbb{E} e^{\beta|\nabla f(X)|^2}$ for some $\beta \in \left(0, \frac{1}{2}\right)$ does not imply finiteness of $\mathbb{E} e^{f(X)}$; for instance, consider $X \sim \mathcal{N}(0, 1)$ and $f(x) = \frac{1}{2}(x^2 - 1)$. Therefore, perhaps

$$\mathbb{E} e^{f(X)} < h(\mathbb{E} e^{|\nabla f(X)|^2/2})$$

is the *best possible* inequality one may seek for some $h : [1, \infty) \to [0, \infty)$.

According to a discussion on page 8 in [Bobkov and Götze 1999], Talagrand showed that even though (1-1) *fails* at the endpoint exponent $\alpha = \frac{1}{2}$, surprisingly, the finiteness of $\mathbb{E} e^{|\nabla f(X)|^2/2}$ still implies finiteness of $\mathbb{E} e^{f}$ for $X \sim \mathcal{N}(0, I_{n\times n})$. We are not aware of Talagrand's proof as it was never published; we do not know if he solved the problem only for $n = 1$ or for all $n \geq 1$.

MSC2020: 26D10, 35E10, 42B35.
Keywords: exponential integrability, heat flow, Gauss space, measure concentration.

In this paper we show that the finiteness of $\mathbb{E}e^{|\nabla f|/2}(1+|\nabla f(X)|)^{-1}$ implies the finiteness of $\mathbb{E}e^{f(X)}$ for all $n \geq 1$, and the extra factor $(1+|\nabla f|)^{-1}$ is the best possible in the sense that it cannot be replaced by $(1+|\nabla f|)^{-c}$ for some $c > 1$. Moreover, we provide quantitative bounds.

Theorem 1.1. *For any $n \geq 1$, we have*

$$\log \mathbb{E}e^{f(X)-\mathbb{E}f} \leq 10\mathbb{E}e^{|\nabla f(X)|^2/2}(1+|\nabla f(X)|)^{-1} \tag{1-2}$$

for all $f \in C_0^\infty(\mathbb{R}^n)$, where $X \sim \mathcal{N}(0, I_{n\times n})$.

To see the sharpness of the factor $(1+|\nabla f|)^{-1}$ in (1-2), let $n=1$, and let $f(x)=\frac{1}{2}x^2$. Then $\mathbb{E}e^{f(X)}=\infty$. On the other hand,

$$\mathbb{E}e^{|\nabla f(X)|^2/2}(1+|\nabla f(X)|)^{-c} = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \frac{dx}{(1+|x|)^c} < \infty$$

for all $c > 1$. It remains to multiply f by a smooth cut-off function $\mathbb{1}_{|x|\leq R}$ and take the limit $R \to \infty$.

Using standard mass transportation arguments the exponential integrability (1-2) may be extended to random vectors X having *uniformly* log-concave densities.

Corollary 1.2. *Let X be an arbitrary random vector in $\mathbb{R}^n$ with density $e^{-u(x)}\,dx$ such that* Hess $u \geq RI_{n\times n}$ *for some $R > 0$. Then*

$$\log \mathbb{E}e^{f(X)-\mathbb{E}f} \leq 10\mathbb{E}e^{|\nabla f(X)|^2/(2R)}(1+R^{-1/2}|\nabla f(X)|)^{-1} \tag{1-3}$$

for all $f \in C_0^\infty(\mathbb{R}^n)$.

Exponential integrability has been studied for other random vectors X as well. Let us briefly record some known results where we assume f to be real-valued with $\mathbb{E}f(Y)=0$. In all examples Y is uniformly distributed on the set where it is given.

$$\log \mathbb{E}e^{f(Y)} \leq \mathbb{E}\tfrac{1}{4}|\nabla_{\mathbb{S}^2} f(Y)|^2, \qquad Y \in \mathbb{S}^2 = \{\|x\|=1,\ x \in \mathbb{R}^3\}, \tag{1-4}$$

$$\log \mathbb{E}e^{f(Y)} \leq 1+\mathbb{E}\tfrac{1}{16}|\nabla f(Y)|^2, \quad Y \in \mathbb{D} = \{\|x\|\leq 1,\ x \in \mathbb{R}^2\}, \tag{1-5}$$

$$\log \mathbb{E}e^{f(Y)} \leq \log \mathbb{E}e^{D(f)^2(Y)}, \qquad Y \in \{-1,1\}^n, \tag{1-6}$$

$$\log \mathbb{E}e^{f(Y)} \leq \log \mathbb{E}e^{4|\nabla f|^2(Y)}, \qquad Y \in [-1,1]^n, \qquad \text{(only for convex } f), \tag{1-7}$$

where in (1-6) by the symbol $D(f)^2$ we denote the *discrete gradient*; see [Bobkov and Götze 1999]. The estimate (1-4), also known as the Mozer–Trudinger inequality (with the best constants due to Onofri), has been critical for geometric applications [Moser 1971; Onofri 1982]. A slightly weaker version of (1-6), namely,

$$\mathbb{E}e^{f(Y)} \leq \mathbb{E}e^{\pi^2 D(f)^2(Y)/8},$$

was obtained by Efraim and Lust-Piquard [2008].

The proof of the main theorem follows from *heat flow* arguments. We construct a certain increasing quantity $A(s)$ with respect to a parameter $s \in [0,1]$. We will see that

$$\log \mathbb{E}e^{f(X)} = A(0) \leq A(1) \leq \mathbb{E}f(X) + 10\mathbb{E}e^{|\nabla f(X)|^2/2}(1+|\nabla f(X)|)^{-1}.$$

To describe the expression for $A(t)$, let $\Phi(t) = \mathbb{P}(X_1 \leq t)$ be the Gaussian cumulative distribution function, and set $k(x) = -\log(\Phi'(t)/\Phi(t))$. Our main object will be a certain function $F : [0, \infty) \to [0, \infty)$ defined as

$$F(x) = \int_0^x e^{k((k')^{-1}(t))}\, dt \quad \text{for all } x \in [0, \infty), \tag{1-8}$$

where $(k')^{-1}$ is the inverse function to k' (it will be explained in the next section why F is well defined). For $g : \mathbb{R}^n \to (0, \infty)$, we consider its heat flow $U_s g(y) := \mathbb{E} g(y + \sqrt{s} X)$, where $s \in [0, 1]$. Then

$$A(s) := U_s\left[\log U_{1-s} g + F\left(\frac{\sqrt{s}|\nabla U_{1-s} g|}{U_{1-s} g}\right)\right](0)$$

will have the desired properties: $A'(s) \geq 0$, $A(0) = \log \mathbb{E} g$, and $A(1) = \mathbb{E} \log g + \mathbb{E} F(|\nabla g|/g)$. The argument gives the inequality

$$\log \mathbb{E} g - \mathbb{E} \log g \leq \mathbb{E} F\left(\frac{|\nabla g|}{g}\right). \tag{1-9}$$

If we set $g(x) = e^{f(x)}$ with $f : \mathbb{R}^n \to \mathbb{R}$ and use the chain rule $|\nabla g|/g = |\nabla f|$, we obtain

$$\log \mathbb{E} e^{f - \mathbb{E} f} \leq \mathbb{E} F(|\nabla f|). \tag{1-10}$$

The last step is to show the pointwise estimate $F(s) \leq 10 e^{s^2/2}(1+s)^{-1}$ for all $s \geq 0$. We remark that the obtained inequality (1-10) is stronger than (1-2) and it should be considered as a corollary of (1-10); however, due to a complicated expression for F we decided to state the main result in the form of (1-2).

The computation of $A'(s)$ is technical and is done in Section 2C, where we also explain how the expression $A(t)$ was "discovered". We should note that the main reason that makes $A' \geq 0$ is the fact that $k'/k'' > 0$ and the inequality[1]

$$1 - k'' - k' e^k \geq 0,$$

which for $k = -\log(\Phi'(t)/\Phi(t))$ serendipitously turns out to be an equality.

Sections 2A and 2B are technical and can be skipped when reading the paper for the first time. In these sections we show that $F \in C^2([0, \infty))$ is an increasing convex function with values $F(0) = F'(0) = 0$ and $F''(0) = 1$. Furthermore, the *modified* hessian matrix of

$$M(x, y) := \log x + \int_0^{y/x} e^{k((k')^{-1}(t))}\, dt \tag{1-11}$$

is positive semidefinite:

$$\begin{pmatrix} M_{xx} + \frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy} \end{pmatrix} \geq 0 \quad \text{for all } (x, y) \in (0, \infty) \times [0, \infty). \tag{1-12}$$

In Section 2C we demonstrate that the condition (1-12) implies the inequality

$$M(\mathbb{E} g(X), 0) \leq \mathbb{E} M(g(X), |\nabla g(X)|) \tag{1-13}$$

for all smooth bounded $g : \mathbb{R}^n \to (0, \infty)$. At the end of Section 2C, we deduce Theorem 1.1 and Corollary 1.2 from (1-13).

[1] It is an equality for $k = -\log(\Phi'(t)/\Phi(t))$ yet an inequality would be sufficient for our purposes.

As an application, in Section 3 we show that the dual inequality to (1-9), in the sense of duality described in Section 3.2 of [Ivanisvili et al. 2018], corresponds to the following theorem.

Theorem 1.3. *For any positive martingale* $\{\xi_n\}_{n\geq 0}$ *on a probability space* $([0,1], \mathcal{B}, dx)$ *adapted to a discrete time dyadic filtration* $([0,1), \varnothing) = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots$ *such that* $\xi_N = \xi_{N+1} = \cdots = \xi_\infty > 0$ *for a sufficiently large* N, *we have*

$$\log \mathbb{E}\xi_\infty - \mathbb{E}\log \xi_\infty \leq \mathbb{E}G\biggl(\frac{\xi_\infty}{[\xi_\infty]^{1/2}}\biggr), \tag{1-14}$$

where $[\xi_\infty] = \sum_{k\geq 0}(\xi_{k+1}-\xi_k)^2$ *is the quadratic variation, and* $G(t) := \int_t^\infty \int_s^\infty r^{-2}e^{(s^2-r^2)/2}\,dr\,ds$.

In Lemma 3.2 we obtain the two-sided estimate

$$\tfrac{1}{3}\log(1+t^{-2}) \leq G(t) \leq \log(1+t^{-2}) \quad \text{for all } t \geq 0.$$

In particular, (1-14) implies that

$$\log \mathbb{E}\xi_\infty - \mathbb{E}\log \xi_\infty \leq \mathbb{E}\log\biggl(1+\frac{[\xi_\infty]}{\xi_\infty^2}\biggr). \tag{1-15}$$

Estimate (1-15) shows how well $\log \xi_\infty$ is concentrated around $\log \mathbb{E}\xi_\infty$ provided that one can control the quadratic variation of ξ_∞. Theorem 1.3 posits a duality approach developed in [Ivanisvili et al. 2018]. This may be considered as complementary to the e-entropy bound

$$\mathbb{E}e^{\xi_\infty - \mathbb{E}\xi_\infty} \leq \frac{e^{-\varepsilon}}{1-\varepsilon}$$

which holds for all discrete time simple martingales ξ_n (not necessarily positive and dyadic) provided that $[\xi_\infty] \leq \varepsilon^2$; see Corollary 1.12 in [Stolyarov et al. 2022].

The proof of (1-14) uses the special function

$$N(p,t) := \log(p) + \int_{p/\sqrt{t}}^\infty \int_s^\infty r^{-2}e^{(s^2-r^2)/2}\,dr\,ds$$

which we find by dualizing $M(x,y) = \log x + F(y/x)$. We deduce that N is *heat convex*, i.e.,

$$2N(p,t) \leq N(p+a, t+a^2) + N(p-a, t+a^2) \tag{1-16}$$

for all reals $p,\ a,\ t$ such that $p \pm a \geq 0$ and $t \geq 0$. Finally, after iterating (1-16), we recover (1-14).

2. Proofs of Theorem 1.1 and Corollary 1.2

2A. ***Step 1*: *an implicit function F and its properties.*** Let

$$k(x) := -\log(\log\Phi(x))' = \tfrac{1}{2}x^2 + \log\biggl(\int_{-\infty}^x e^{-s^2/2}\,ds\biggr) \quad \text{for all } x \in \mathbb{R}.$$

Define a real-valued function F as

$$F(k'(t)) = \int_{-\infty}^t k''(s)e^{k(s)}\,ds \quad \text{for all } t \in \mathbb{R}. \tag{2-1}$$

The following lemma, in particular, shows that F is well defined.

Lemma 2.1. *We have*

(1) $k'(-\infty)=0,\ k'(x)\sim x$ *as* $x\to\infty$, *and* $k''>0$ (*and hence* $k'>0$);

(2) $F:[0,\infty)\to[0,\infty),\ F(0)=F'(0)=0,\ F''(0)=1,\ F'(k')=e^k$, *and* $F''(k')=(k'/k'')e^k$.

Proof. Let us investigate the asymptotic behavior of k and its derivatives at $x=-\infty$. Let $x<0$, and for $m\geq 0$ define

$$I_m := e^{x^2/2}\int_{-\infty}^{x} e^{-s^2/2}s^{-m}\,ds.$$

Integration by parts reveals $I_m=-x^{-(m+1)}-(m+1)I_{m+2}$. By iterating we obtain

$$\begin{aligned} e^{x^2/2}\int_{-\infty}^{x} e^{-s^2/2}\,ds = I_0 &= -x^{-1}+x^{-3}-3\cdot x^{-5}+3\cdot 5\cdot x^{-7}+3\cdot 5\cdot 7\cdot I_8\\ &= -x^{-1}+x^{-3}-3x^{-5}+O(|x|^{-7}) \qquad\text{as } x\to-\infty, \end{aligned}$$

because $|I_8|\leq\int_{-\infty}^{x}s^{-8}\,ds\leq|x|^{-7}$. Thus, as $x\to-\infty$ we have

$$\begin{gathered} e^{k(x)}=I_0=-x^{-1}+x^{-3}-3x^{-5}+O(|x|^{-7}), \qquad e^{-k(x)}=-x-x^{-1}+2x^{-3}+O(|x|^{-5}),\\ k'(x)=x+e^{-k(x)}=-x^{-1}+2x^{-3}+O(|x|^{-5}), \qquad k''(x)=1-k'(x)e^{-k(x)}=x^{-2}+O(|x|^{-4}),\\ k''(x)e^{k(x)}=-x^{-3}+O(|x|^{-5}), \qquad\text{and}\qquad \frac{k'(x)e^{k(x)}}{k''(x)}=1+O(|x|^{-2}). \end{gathered}$$

The claim

$$k'(x)=x+\frac{1}{e^{x^2/2}\int_{-\infty}^{x}e^{-s^2/2}\,ds}\sim x \quad\text{as } x\to\infty$$

is trivial. Next, we show that $k''>0$. By elementary calculus we have

$$k''=1-\frac{xe^{x^2/2}\int_{-\infty}^{x}e^{-s^2/2}\,ds+1}{\left(e^{x^2/2}\int_{-\infty}^{x}e^{-s^2/2}\,ds\right)^2}=\frac{e^{x^2}}{e^{2k(x)}}\left[\left(\int_{-\infty}^{x}e^{-s^2/2}\,ds\right)^2-xe^{-x^2/2}\int_{-\infty}^{x}e^{-s^2/2}\,ds-e^{-x^2}\right].$$

If we let $h(x):=e^{-x^2/2}$ and $H(x):=\int_{-\infty}^{x}e^{-t^2/2}\,dt$, then it suffices to show

$$u(x):=H^2-xhH-h^2>0.$$

Clearly $H'=h$ and $h'=-xh$. Next

$$\begin{aligned} u'&=2Hh-hH+x^2hH-xh^2+2xh^2=Hh+x^2hH+xh^2\\ &=(H+x^2H+xh)h=\left(H+h\frac{x}{1+x^2}\right)(1+x^2)h. \end{aligned}$$

Let $v(x)=H(x)+h(x)x/(1+x^2)$. Then, we have

$$v'(x)=h-h\frac{x^2}{1+x^2}+h\frac{1-x^2}{(1+x^2)^2}=\left(\frac{1}{1+x^2}+\frac{1-x^2}{(1+x^2)^2}\right)h=\frac{2}{(1+x^2)^2}h>0.$$

Since $v(-\infty)=0$ and $v'>0$, we obtain $v(x)>0$ for all $x\in\mathbb{R}$. In particular $u'>0$, and taking into account that $u(-\infty)=0$, we conclude $u(x)>0$ for all $x\in\mathbb{R}$.

To verify the second part of the lemma notice that $F(0)=0$ by considering the limit as $t\to-\infty$ in (2-1). Taking the derivative in t of (2-1) and dividing both sides by $k''>0$ we obtain $F'(k')=e^k$. Considering the limit as $t\to-\infty$ we realize $F'(0)=0$. Taking the second derivative gives $F''(k')=(k'/k'')e^k$. We have $(k'/k'')e^k=1+O(x^{-2})\to 1$ as $x\to-\infty$. Hence $F''(0)=1$, proving the lemma. □

It follows that $k'>0$ and $k':\mathbb{R}\to[0,\infty)$. Thus, we may consider the inverse map $t\mapsto k'(t)$ denoted by $(k')^{-1}:[0,\infty)\to\mathbb{R}$. After a suitable change of variables in (2-1), we write

$$\begin{aligned} F(x) &= \int_{-\infty}^{(k')^{-1}(x)} k''(s)e^{k(s)}\,ds \\ &= \int_0^x e^{k((k')^{-1}(u))}\,du, \qquad (\text{by } s=(k')^{-1}(u)), \end{aligned}$$

which coincides with the expression announced in (1-8).

Lemma 2.2. *We have* $F(x)\le 10e^{x^2/2}(1+x)^{-1}$ *for all* $x\ge 0$.

Proof. Notice that $k'(u)=u+e^{-k(u)}\ge u$ (for all $u\in\mathbb{R}$) and $k''>0$. Therefore, $u\ge(k')^{-1}(u)$ for $u\ge 0$, so the inequality $k(u)\ge k((k')^{-1}(u))$ follows from the fact that $k'>0$. Thus

$$F(x)\le\int_0^x e^{k(u)}\,du=\int_0^x e^{u^2/2}\int_{-\infty}^u e^{-s^2/2}\,ds\,du\le\sqrt{2\pi}\int_0^x e^{u^2/2}\,du.$$

Next, we claim the simple chain of inequalities

$$\int_0^x e^{u^2/2}\,du \overset{\text{(A)}}{\le} \frac{2x}{1+x^2}e^{x^2/2} \overset{\text{(B)}}{\le} \frac{3}{1+x}e^{x^2/2}.$$

Indeed, inequality (A) follows from the fact that it is true at $x=0$ and

$$\frac{d}{dx}\left(\frac{2x}{1+x^2}e^{x^2/2}-\int_0^x e^{u^2/2}du\right)=e^{x^2/2}\left(1-\frac{4x^2}{(1+x^2)^2}\right)\ge e^{x^2/2}\left(1-\frac{4x^2}{(2x)^2}\right)=0.$$

In contrast, inequality (B) is immediate. Therefore, we conclude that

$$F(x)\le 3\sqrt{2\pi}e^{x^2/2}(1+x)^{-1}\le 10e^{x^2/2}(1+x)^{-1}\quad\text{for all } x\ge 0.$$ □

2B. ***Step 2*: *Monge–Ampère type PDE.*** Define

$$M(x,y)=\log x+F(y/x)\quad\text{for all } (x,y)\in(0,\infty)\times[0,\infty). \tag{2-2}$$

Clearly $M\in C^2$ and $M_y(x,0)=0$, where $M_x=\partial M/\partial x$ and $M_y=\partial M/\partial y$. Next, let us consider the matrix

$$A(x,y):=\begin{pmatrix} M_{xx}+\frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy}\end{pmatrix}. \tag{2-3}$$

We claim the following:

Lemma 2.3. *For each* $(x,y)\in(0,\infty)\times[0,\infty)$, *the matrix* $A(x,y)$ *is positive semidefinite with* $\det(A)=0$.

Proof. Let us calculate the partial derivatives of M. Let $t := yx^{-1}$. We have

$$M_x = x^{-1} - yx^{-2}F'(yx^{-1}) = x^{-1}(1 - tF'(t)), \quad M_y = x^{-1}F'(t),$$
$$M_{xx} = -x^{-2} + 2yx^{-3}F'(yx^{-1}) + (yx^{-2})^2F''(yx^{-1}) = x^{-2}(-1 + 2tF'(t) + t^2F''(t)),$$
$$M_{yx} = -x^{-2}(F'(t) + tF''(t)), \quad M_{yy} = x^{-2}F''(t) \overset{\text{Lemma 2.1}}{>} 0.$$

To see that $A(x, y)$ is positive semidefinite, it suffices (due to the inequality $M_{yy} > 0$) to check that $\det(A) = 0$. We have

$$\det(A) = M_{xx}M_{yy} - M_{xy}^2 + \frac{M_yM_{yy}}{y} = x^{-4}\left[(-1 + 2tF' + t^2F'')F'' - (F' + tF'')^2 + \frac{F'F''}{t}\right]$$
$$= x^{-4}\left[-F'' - (F')^2 + \frac{F'F''}{t}\right].$$

Next, for $t = k'$ we have $F'(k') = e^k$ and $F''(k') = k'e^k/k''$ by Lemma 2.1. Therefore

$$-F'' - (F')^2 + \frac{F'F''}{t} = -\frac{k'e^k}{k''} - e^{2k} + \frac{e^{2k}}{k''} = \frac{e^{2k}}{k''}(1 - k'' - k'e^{-k}) = 0,$$

as $k'(x) = x + e^{-k(x)}$ (and hence $k'' = 1 - k'e^{-k}$). □

2C. *Step 3*: *the heat flow argument.* First we would like to give an explanation for how the flow is constructed. For simplicity consider $n = 1$. If we succeed in proving the inequality

$$M(\mathbb{E}g(\xi), 0) \le \mathbb{E}M(g(\xi), |g'(\xi)|), \quad g : \mathbb{R} \to (0, \infty), \tag{2-4}$$

where $\xi \sim \mathcal{N}(0, 1)$ and $M(x, y) = \log x + F(y/x)$, then we obtain

$$\log \mathbb{E}g + F(0) \le \mathbb{E}\log g + \mathbb{E}F(|g'|/g),$$

which for $g = e^f$ coincides with (1-10). So the goal is to prove (2-4). We consider a discrete approximation of ξ, namely, let

$$\vec{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_m),$$

where the ε_j are i.i.d. symmetric Bernoulli ± 1 random variables. By the central limit theorem,

$$\frac{\varepsilon_1 + \dots + \varepsilon_m}{\sqrt{m}} \xrightarrow{d} \xi \quad \text{as } m \to \infty.$$

We hope to prove the *hypercube analog* of (2-4), i.e.,

$$M(\mathbb{E}\tilde{g}(\vec{\varepsilon}), 0) \le \mathbb{E}M(\tilde{g}(\vec{\varepsilon}), |D\tilde{g}(\vec{\varepsilon})|), \quad \tilde{g}(\vec{\varepsilon}) = g\left(\frac{\varepsilon_1 + \dots + \varepsilon_m}{\sqrt{m}}\right), \tag{2-5}$$

for all $m \ge 1$, where the *discrete* gradient $|D\tilde{g}(\vec{\varepsilon})| := \sqrt{\sum_{j=1}^m |D_j\tilde{g}(\vec{\varepsilon})|^2}$ is defined as follows:

$$D_j\tilde{g}(\varepsilon_1, \dots, \varepsilon_m) = \frac{\tilde{g}(\varepsilon_1, \dots, \varepsilon_j, \dots, \varepsilon_m) - \tilde{g}(\varepsilon_1, \dots, -\varepsilon_j, \dots, \varepsilon_m)}{2} \quad \text{for } j = 1, \dots, m.$$

One sees that as $m \to \infty$ we have

$$D_j \tilde{g}(\vec{\varepsilon}) = g'\left(\frac{\varepsilon_1 + \cdots + \varepsilon_m}{\sqrt{m}}\right)\frac{\varepsilon_j}{\sqrt{m}} + O\left(\frac{1}{m}\right)$$

and

$$|D\tilde{g}(\vec{\varepsilon})| = \sqrt{\left[g'\left(\frac{\varepsilon_1 + \cdots + \varepsilon_m}{\sqrt{m}}\right)\right]^2 + O\left(\frac{1}{\sqrt{m}}\right)},$$

at least for bounded smooth functions g with uniformly bounded derivatives. Thus taking the limit $m \to \infty$ we observe that the right-hand side of (2-5) converges to the right-hand side of (2-4); in particular, (2-5) implies (2-4).

Next, we take this one step further and consider the inequality (2-5) for all $\tilde{g} : \{-1, 1\}^m \to \mathbb{R}$ instead of the specific functions defined in (2-5); in doing so we are ever so slightly enlarging the class of test functions to include those that are not invariant with respect to permutations of $(\varepsilon_1, \dots, \varepsilon_n)$. To prove that

$$M(\mathbb{E}h, 0) \le \mathbb{E}M(h, |Dh|) \qquad \text{for all } h : \{-1, 1\}^m \to (0, \infty) \quad \text{and all } m \ge 1, \tag{2-6}$$

one trivial argument would be to invoke the product structure of $\{-1, 1\}^m$. For example, if we manage to show an *intermediate* “4-point” inequality

$$M(\mathbb{E}_{\varepsilon_1}h, |D\mathbb{E}_{\varepsilon_1}h|) \le \mathbb{E}_{\varepsilon_1}M(h, |Dh|), \tag{2-7}$$

where $\mathbb{E}_{\varepsilon_1}$ averages only with respect to ε_1, then by iterating (2-7) we deduce the inequality

$$M(\mathbb{E}h, 0) = M(\mathbb{E}_{\varepsilon_m} \cdots \mathbb{E}_{\varepsilon_1}h, |D\mathbb{E}_{\varepsilon_m} \cdots \mathbb{E}_{\varepsilon_1}h|) \le \mathbb{E}_{\varepsilon_1} \cdots \mathbb{E}_{\varepsilon_m}M(h, |Dh|) = \mathbb{E}M(h, |Dh|).$$

Upon closer inspection, we see that (2-7) follows[2] from the 4-point inequality

$$2M(x, y) \le M\big(x + a, \sqrt{a^2 + (y + b)^2}\big) + M\big(x - a, \sqrt{a^2 + (y - b)^2}\big) \tag{2-8}$$

for all real numbers $x,\ y,\ a,\ b$ such that $x \pm a > 0$. To prove (2-8) for one specific M seems to be a possible task; however, if we take into account that M is defined by (2-2) which involves an implicitly defined F, the 4-point inequality (2-8) becomes complicated (see [Ivanisvili and Volberg 2020], where one such inequality was proved for $M(x, y) = -\Re(x + iy)^{3/2}$ by tedious computations involving high degree polynomials with integer coefficients).

Expanding (2-8) at the point $(a, b) = (0, 0)$ via Taylor series, one easily obtains a necessary assumption: the infinitesimal form of (2-8), i.e.,

$$\begin{pmatrix} M_{xx} + \frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy} \end{pmatrix} \ge 0. \tag{2-9}$$

Of course, the infinitesimal condition (2-9) does not necessarily imply its global two-point inequality (2-8) (and in particular (2-6)). Also, it may seem implausible to believe that the positive semidefiniteness of (2-9) implies the inequality (2-4) in Gauss space. Surprisingly this last guess turns out to be correct, and

[2] In fact they are equivalent provided that $y \mapsto M(x, y)$ is nondecreasing.

perhaps the reason lies in the fact that one only needs to verify (2-5) as $m \to \infty$ (and only for symmetric functions $\tilde{g}$). Let us "take the limit" and see how the heat flow arises.

Let $\mathbb{E}_{m-k}$ be the average with respect to the variables $\varepsilon_1, \dots, \varepsilon_{m-k}$, and let $\mathbb{E}^k$ be the average with respect to the remaining variables $\varepsilon_{m-k+1}, \dots, \varepsilon_m$. Then the 4-point inequality (2-7) implies that

$$k \mapsto \mathbb{E}^k M(\mathbb{E}_{m-k}\tilde{g}, |D\mathbb{E}_{m-k}\tilde{g}|) \quad \text{is nondecreasing on } [0, m].$$

The expression $\mathbb{E}^k M(\mathbb{E}_{m-k}\tilde{g}, |D\mathbb{E}_{m-k}\tilde{g}|)$ we rewrite as $\mathbb{E}^k M(A, B)$, where

$$A = \mathbb{E}_{m-k} g\biggl(\frac{\sum_{j=1}^{k} \varepsilon_j}{\sqrt{k}}\sqrt{\frac{k}{m}} + \frac{\sum_{j=k+1}^{m} \varepsilon_j}{\sqrt{m-k}}\sqrt{1-\frac{k}{m}}\biggr),$$
$$B = \sqrt{\frac{k}{m}\biggl[\mathbb{E}_{m-k} g'\biggl(\frac{\sum_{j=1}^{k} \varepsilon_j}{\sqrt{k}}\sqrt{\frac{k}{m}} + \frac{\sum_{j=k+1}^{m} \varepsilon_j}{\sqrt{m-k}}\sqrt{1-\frac{k}{m}}\biggr)\biggr]^2 + O\Bigl(\frac{k}{m^{3/2}}\Bigr)}.$$

Taking $k, m \to \infty$ so that $\frac{k}{m} \to s \in [0, 1]$, one can conclude that

$$s \mapsto \mathbb{E}_X M(\mathbb{E}_Y g(X\sqrt{s} + Y\sqrt{1-s}), \sqrt{s}|\mathbb{E}_Y g'(X\sqrt{s} + Y\sqrt{1-s})|) \quad \text{is nondecreasing on } [0, 1],$$

where $X, Y \in \mathcal{N}(0, 1)$ are independent and $\mathbb{E}_X$ takes the expectation with respect to the random variable X. In other words, if we let $U_s g(y) = \mathbb{E} g(y + \sqrt{s}X)$ to be a heat flow defined as

$$\frac{\partial}{\partial s} U_s g = \frac{1}{2}\frac{\partial^2}{\partial x^2} U_s g, \qquad U_0 g = g,$$

then

$$s \mapsto U_s M(U_{1-s} g, \sqrt{s}|U_{1-s} g'|)(0) \quad \text{is nondecreasing on } [0, 1]. \tag{2-10}$$

Luckily we may ignore all the steps by starting from the map (2-10) and taking its derivative in s to divine when it has nonnegative sign. Slightly abusing the notations, denote $D = \partial/\partial x$, and, for simplicity, let us work with the map $s \mapsto U_s M(U_{1-s} g, \sqrt{s} U_{1-s} g')$, where we omit the absolute value in the second argument of M. Let $b = U_{1-s} g$. Clearly $db/ds = -\frac{1}{2} D^2 b$. We have

$$\begin{aligned}
\frac{d}{ds} U_s M(b, \sqrt{s} Db) &= \frac{1}{2} D^2 U_s M(b, \sqrt{s} Db) + U_s\biggl(-\frac{1}{2} D^2 b M_x + \Bigl(\frac{1}{2\sqrt{s}} Db - \frac{\sqrt{s}}{2} D^3 b\Bigr) M_y\biggr) \\
&= \frac{U_s}{2}\biggl(D(M_x Db + M_y \sqrt{s} D^2 b) - M_x D^2 b + \frac{M_y}{\sqrt{s}} Db - M_y \sqrt{s} D^3 b\biggr) \\
&= \frac{U_s}{2}\biggl(M_{xx}(Db)^2 + 2M_{xy}\sqrt{s}\, Db D^2 b + M_{yy} s (D^2 b)^2 + \frac{M_y}{\sqrt{s}} Db\biggr).
\end{aligned}$$

Notice that

$$\begin{aligned}
&M_{xx}(Db)^2 + 2M_{xy}\sqrt{s}\, Db D^2 b + M_{yy} s (D^2 b)^2 + \frac{M_y}{\sqrt{s}} Db \\
&\qquad = \begin{pmatrix} Db & \sqrt{s} D^2 b \end{pmatrix} \begin{pmatrix} M_{xx} + \frac{M_y}{\sqrt{s} Db} & M_{xy} \\ M_{xy} & M_{yy} \end{pmatrix} \begin{pmatrix} Db \\ \sqrt{s} D^2 b \end{pmatrix} \ge 0.
\end{aligned}$$

It remains to extend the argument to higher dimensions and put the absolute value back into the second argument of M.

Theorem 2.4. *Let $M : (0,\infty)\times[0,\infty)\to\mathbb{R}$ be such that $M\in C^2$ with $M_y(x,0)=0$ and*

$$\begin{pmatrix} M_{xx}+\frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy}\end{pmatrix}\geq 0. \tag{2-11}$$

Then the map

$$s\mapsto U_s M(U_{1-s}g, \sqrt{s}|\nabla U_{1-s}g|) \quad \text{is nondecreasing on } [0,1] \tag{2-12}$$

for all smooth bounded functions $g:\mathbb{R}^n\to(0,\infty)$ with uniformly bounded first and second derivatives.

Proof. Let $M(x,y)=B(x,y^2)$. Let B_1 and B_2 be partial derivatives of B. Positive semidefiniteness of the matrix (2-11) in terms of B converts to

$$\begin{pmatrix} B_{11}(x,y^2)+2B_2(x,y^2) & 2yB_{12}(x,y^2) \\ 2yB_{12}(x,y^2) & 2B_2(x,y^2)+4y^2B_{22}(x,y^2)\end{pmatrix}\geq 0 \tag{2-13}$$

for all $x>0$ and all $y\geq 0$ (in fact this holds for all $y\in\mathbb{R}$). Next, let $G=U_{1-s}g$. Clearly $dG/ds=-\frac{1}{2}\Delta G$. We have

$$\frac{d}{ds}U_sB(U_{1-s}g, s|U_{1-s}\nabla g|^2)=\tfrac{1}{2}U_s[\Delta B(G,s|\nabla G|^2)-B_1\Delta G+2B_2|\nabla G|^2-2B_2 s\nabla G\cdot\nabla\Delta G].$$

Next, let $D_j=\partial/\partial x_j$. Then

$$\begin{aligned}
D_jB(G,s|\nabla G|^2)&=B_1D_jG+B_2sD_j|\nabla G|^2,\\
D_j^2B(G,s|\nabla G|^2)&=B_{11}(D_jG)^2+2B_{12}D_jGsD_j|\nabla G|^2+B_{22}s^2(D_j|\nabla G|^2)^2+B_1D_j^2G+B_2sD_j^2|\nabla G|^2,\\
\Delta B(G,s|\nabla G|^2)&=B_{11}|\nabla G|^2+2B_{12}\nabla G\cdot s\nabla|\nabla G|^2+B_{22}\big|s\nabla|\nabla G|^2\big|^2+B_1\Delta G+B_2s\Delta|\nabla G|^2.
\end{aligned}$$

Notice that $\Delta|\nabla G|^2=2\nabla G\cdot\nabla\Delta G+2\operatorname{Tr}(\operatorname{Hess}G)^2$. Therefore

$$\begin{aligned}
&\Delta B(G,s|\nabla G|^2)-B_1\Delta G+2B_2|\nabla G|^2-2B_2s\nabla G\cdot\nabla\Delta G\\
&\quad=B_{11}|\nabla G|^2+2B_{12}\nabla G\cdot s\nabla|\nabla G|^2+B_{22}\big|s\nabla|\nabla G|^2\big|^2+2B_2|\nabla G|^2+2B_2s\operatorname{Tr}(\operatorname{Hess}G)^2\\
&\quad\geq B_{11}|\nabla G|^2-2|B_{12}||\nabla G|\big|s\nabla|\nabla G|^2\big|+B_{22}\big|s\nabla|\nabla G|^2\big|^2+2B_2|\nabla G|^2+2B_2s\operatorname{Tr}(\operatorname{Hess}G)^2.
\end{aligned} \tag{2-14}$$

First we want to consider the case when $|\nabla G|=0$. We recall that $M(x,y)=B(x,y^2)$. Therefore $B_2(x,0)$ exists and is equal to $\frac{1}{2}M_{yy}(x,0)$ (due to the fact that $M_y(x,0)=0$). Also

$$\lim_{y\to 0}B_{12}(x,y^2)y=\tfrac{1}{2}M_{xy}(x,0)$$

and

$$\lim_{y\to0}B_{22}(x,y^2)y^2=\lim_{y\to0}\tfrac{1}{4}(M_{yy}(x,|y|)-2B_2(x,y^2))=0.$$

Therefore, if $|\nabla G| = 0$, then due to the inequality

$$\mathrm{Tr}(\mathrm{Hess}\, G)^2 |\nabla G|^2 = \sum_{j=1}^{n} |\nabla D_j G|^2 |\nabla G|^2 \geq \sum_{j=1}^{n} (\nabla D_j G \cdot \nabla G)^2 = \tfrac{1}{4}\big|\nabla |\nabla G|^2\big|^2, \tag{2-15}$$

the expression (2-14) simplifies to

$$2B_2(G,0)s\, \mathrm{Tr}(\mathrm{Hess}\, G)^2 = \tfrac{1}{2} M_{yy}(G,0)s\, \mathrm{Tr}(\mathrm{Hess}\, G)^2 \geq 0,$$

where the last inequality holds true by assumption (2-11), hence (2-14) is nonnegative.

If $|\nabla G| > 0$ then we proceed as follows: Assumption (2-11) implies $yM_{xx} + M_y \geq 0$. In particular, taking $y = 0$ we obtain $M_y(x,0) \geq 0$. Also it follows from (2-11) that $M_{yy} \geq 0$. Thus $M_y(x,y) \geq 0$ for all $y \geq 0$. In particular, $B_2(x,y^2) \geq 0$ for all $y > 0$ (and also for $y = 0$ as we just noticed that $B_2(x,0) = \frac{1}{2} M_{yy}(x,0) \geq 0$). Therefore, using inequality (2-15), we may estimate the last term in (2-14) from below as $B_2 \big|s\nabla|\nabla G|^2\big|^2/(2s|\nabla G|^2)$. Finally,

$$\begin{aligned} &B_{11}|\nabla G|^2 - 2|B_{12}|\,|\nabla G|\,\big|s\nabla|\nabla G|^2\big| + B_{22}\big|s\nabla|\nabla G|^2\big|^2 + 2B_2|\nabla G|^2 + B_2 \frac{\big|s\nabla|\nabla G|^2\big|^2}{2s|\nabla G|^2} \\ &\qquad = \begin{pmatrix} -|\nabla G| & \frac{\sqrt{s}|\nabla|\nabla G|^2|}{2|\nabla G|} \end{pmatrix} \begin{pmatrix} B_{11}+2B_2 & 2\sqrt{s}|\nabla G|\,|B_{12}| \\ 2\sqrt{s}|\nabla G|\,|B_{12}| & 4s|\nabla G|^2 B_{22} + 2B_2 \end{pmatrix} \begin{pmatrix} -|\nabla G| \\ \frac{\sqrt{s}|\nabla|\nabla G|^2|}{2|\nabla G|} \end{pmatrix} \geq 0 \end{aligned}$$

by assumption (2-13) and the fact that B is evaluated at the point $(G, s|\nabla G|^2)$. □

Proof of Theorem 1.1. Notice that $U_s g(y) = \mathbb{E} g(y + \sqrt{s}X)$. Therefore, comparing the values of the map (2-12) at the endpoints $s = 0$ and $s = 1$ we obtain

$$M(\mathbb{E} g(X), 0) = U_0 M(U_1 g, \sqrt{0}|U_1 \nabla g|)(0) \leq U_1 M(U_0 g, |U_0 \nabla g|)(0) = \mathbb{E} M(g(X), |\nabla g(X)|).$$

In particular, for $g = e^f$ where $f \in C_0^\infty(\mathbb{R}^n)$, we obtain

$$\log \mathbb{E} e^{f(X)} \leq \mathbb{E} f(X) + \mathbb{E} F(|\nabla f(X)|).$$

The pointwise inequality $F(x) \leq 10 e^{x^2/2}(1+x)^{-1}$ from Lemma 2.2 finishes the proof of Theorem 1.1 □

Proof of Corollary 1.2. Let $d\mu = e^{-u(x)}\, dx$ be the density of the log-concave random vector X with $\mathrm{Hess}\, u \geq R I_{n\times n}$ for some $R > 0$. It follows from [Caffarelli 2000] that there exists a convex function $\psi : \mathbb{R}^n \to \mathbb{R}$ such that the Brenier map $T = \nabla\psi$ pushes forward the Gaussian measure

$$d\gamma_n(x) = \frac{e^{-|x|^2/2}}{\sqrt{(2\pi)^n}}\, dx$$

onto $d\mu$ and such that $0 \leq \mathrm{Hess}\, \psi \leq (1/\sqrt{R}) I_{n\times n}$. Next, apply the inequality

$$\log \int_{\mathbb{R}^n} e^{f(x)}\, d\gamma_n(x) \leq \int_{\mathbb{R}^n} f(x)\, d\gamma_n(x) + \int_{\mathbb{R}^n} F(|\nabla f(x)|)\, d\gamma_n(x) \tag{2-16}$$

with $f(x) = h(\nabla\psi(x))$ for an arbitrary $h \in C_0^\infty(\mathbb{R}^n)$. Then notice that

$$|\nabla f(x)| = |\mathrm{Hess}\, \psi\, \nabla h(\nabla\psi(x))| \leq \frac{1}{\sqrt{R}} |\nabla h(\nabla\psi(x))|.$$

Since $F' > 0$, we conclude that

$$\begin{aligned} F(|\nabla f(x)|) &\le F\left(\frac{1}{\sqrt{R}}|\nabla h(\nabla\psi(x))|\right) \\ &\le 10 e^{|\nabla h(\nabla\psi(x))|^2/(2R)}(1+R^{-1/2}|\nabla h(\nabla\psi)|)^{-1}. \end{aligned}$$

The preceding inequality together with (2-16) implies that

$$\log\int_{\mathbb{R}^n} e^{h(x)}\,d\mu(x) \le \int_{\mathbb{R}^n} h(x)\,d\mu(x) + 10\int_{\mathbb{R}^n} e^{|\nabla h(x)|^2/(2R)}(1+R^{-1/2}|\nabla h(x)|)^{-1}\,d\mu(x)$$

for all $h \in C_0^\infty(\mathbb{R}^n)$. This finishes the proof of Corollary 1.2. ☐

Remark 2.5. The transport map $T(x_1,\dots,x_n) = (\Phi(x_1),\dots,\Phi(x_n))$ pushes forward the standard Gaussian measure onto the uniform measure on $[0,1]^n$, and it is $(2\pi)^{-1/2}$ Lipschitz. Therefore, the inequality (2-16) applied to $f(x) = h(T(x))$ for a smooth $h : [0,1]^n \to \mathbb{R}$ implies that

$$\begin{aligned} \log \mathbb{E} e^{h(Y)-\mathbb{E}h(Y)} &\le \mathbb{E} F((2\pi)^{-1/2}|\nabla h(Y)|) \\ &\le \mathbb{E} e^{|\nabla h(Y)|^2/(4\pi)}(1+(2\pi)^{-1/2}|\nabla h|), \end{aligned}$$

where $Y \sim \mathrm{unif}([0,1]^n)$. We thank an anonymous referee for this remark.

3. Applications: the proofs of Theorem 1.3 and estimate (1-15)

Let us recall the definition of dyadic martingales. For each $n \ge 0$ we denote by $\mathcal{D}_n$ the dyadic intervals belonging to $[0,1)$ of level n, i.e.,

$$\mathcal{D}_n = \left\{ \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right),\ k = 0,\dots,2^n-1 \right\}.$$

Given $\xi \in L^1([0,1], dx)$, define a *dyadic martingale* $\{\xi_k\}_{k\ge 0}$ as

$$\xi_n(x) := \sum_{I\in\mathcal{D}_n} \langle\xi\rangle_I \mathbb{1}_I(x), \quad n \ge 0,$$

where

$$\langle\xi\rangle_I = \frac{1}{|I|}\int_I \xi\,dx;$$

here $|I|$ denotes the Lebesgue length of I. If we let $\mathcal{F}_n$ be the σ-algebra generated by the dyadic intervals in $\mathcal{D}_n$, then $\xi_n = \mathbb{E}(\xi \mid \mathcal{F}_n)$ is the martingale with respect to the increasing filtration $\{\mathcal{F}_k\}_{k\ge 0}$. Next we define the quadratic variation

$$[\xi] = \sum_{n\ge 0} d_n^2,$$

where $d_n := \xi_n - \xi_{n-1}$ is the martingale difference sequence. In what follows, to avoid the issues with convergence of the infinite series we will be assuming that all but finitely many d_n are zero, i.e., $\xi_N = \xi_{N+1} = \cdots = \xi$ for N sufficiently large. Such martingales we call simple dyadic martingales; they are also known as Walsh–Paley martingales [Hytönen et al. 2016].

Let

$$N(p,t) := \log(p) + G(p/\sqrt{t})$$

for $p > 0,\ t \geq 0$, where

$$G(s) = \int_s^\infty \int_r^\infty u^{-2} e^{(r^2-u^2)/2}\, du\, dr, \quad s > 0.$$

Lemma 3.1. *For all real numbers $p,\ a,\ t$ we have*

$$N(p+a, t+a^2) + N(p-a, t+a^2) \geq 2N(p,t), \tag{3-1}$$

provided that $p \pm a > 0$ and $t \geq 0$.

Proof. First we verify that $N(p,t)$ satisfies the backward heat equation

$$\tfrac{1}{2} N_{pp} + N_t = 0. \tag{3-2}$$

Indeed, we have

$$\begin{aligned} N_{pp} + 2N_t &= -\frac{1}{p^2} + \frac{G''(p/\sqrt{t})}{t} - G'(p/\sqrt{t})\, p t^{-3/2} \\ &= \frac{1}{p^2}(-1 + s^2 G''(s) - s^3 G'(s)), \end{aligned} \tag{3-3}$$

where $s = p/\sqrt{t}$. Direct calculations show that

$$G'(s) = -e^{s^2/2} \int_s^\infty u^{-2} e^{-u^2/2}\, du \quad \text{and} \quad G''(s) = s^{-2} - s e^{s^2/2} \int_s^\infty u^{-2} e^{-u^2/2}\, du. \tag{3-4}$$

Substituting (3-4) into (3-3) we see that the expression in (3-3) is zero.

Next, we claim that $t \mapsto N(p,t)$ is concave. Indeed,

$$\begin{aligned} N_t &= -\frac{1}{2} p t^{-3/2} G'(p/\sqrt{t}), \\ N_{tt} &= \frac{1}{4} p^2 t^{-3} G''(p/\sqrt{t}) + \frac{3}{4} p t^{-5/2} G'(p/\sqrt{t}) \\ &= \frac{1}{4t^2}[s^2 G''(s) + 3s G'(s)]. \end{aligned}$$

Since $N_{pp} + 2N_t = 0$, we have $G'' = s^{-2} + sG'(s)$ by (3-3). Therefore, the sign of N_{tt} coincides with the sign of $1 + (s^3 + 3s)G'(s)$. Using (3-4), it suffices to show that

$$\varphi(s) := \frac{e^{-s^2/2}}{s^3+3s} - \int_s^\infty u^{-2} e^{-u^2/2}\, du \leq 0 \quad \text{for all } s \geq 0.$$

We have $\varphi(\infty) = 0$ and

$$\varphi'(s) = e^{-s^2/2}\left[-\frac{1}{3+s^2} - \frac{3+3s^2}{(3s+s^3)^2} + \frac{1}{s^2}\right] = \frac{6e^{s^2/2}}{(3s+s^2)^2} \geq 0,$$

thereby $\varphi(s) \leq 0$, and hence $t \mapsto N(p,t)$ is concave for $t \geq 0$.

Next, consider the process

$$X_s = N(p + B_s, t + s),$$

where B_s is the standard Brownian motion starting at zero. It follows from Ito's formula that X_s is a martingale. Indeed, we have

$$dX_s = N_s ds + N_p dB_s + \tfrac{1}{2} N_{pp} ds \overset{\text{(3-2)}}{=} N_p dB_s.$$

Define the stopping time

$$\tau = \inf\{s \geq 0 : B_s \notin (-a, a)\}.$$

Set $Y_s = X_{\min\{s,\tau\}}$ for $s \geq 0$. Clearly Y_s is a martingale. On the one hand $Y_0 = N(p, t)$. On the other hand

$$\begin{aligned}
\mathbb{E}Y_\infty &= \mathbb{E}N(p+B_\tau, t+\tau) \\
&= \mathbb{E}(N(p-a, t+\tau) \mid B_\tau = -a)\mathbb{P}(B_\tau = -a) + \mathbb{E}(N(p+a, t+\tau) \mid B_\tau = -a)\mathbb{P}(B_\tau = -a) \\
&\leq \tfrac{1}{2}[N(p-a, t+\mathbb{E}(\tau \mid B_\tau = -a)) + N(p+a, t+\mathbb{E}(\tau \mid B_\tau = a))], \quad \text{(by concavity of } t \mapsto N(p, t)).
\end{aligned}$$

Finally, as $B_s^2 - s$ is a martingale, we have that $0 = \mathbb{E}(B_\tau^2 - \tau) = a^2 - \mathbb{E}\tau$. By symmetry we obtain $\mathbb{E}(\tau \mid B_\tau = -a) = \mathbb{E}(\tau \mid B_\tau = a) = a^2$. Thus the lemma follows from the optional stopping theorem. $\square$

Before we complete the proof of Theorem 1.3 let us make a remark. If $N(p, t)$ is an arbitrary smooth function satisfying the backwards heat equation (3-2) and inequality (3-1), then $t \mapsto N(p, t)$ must be concave. In other words, the concavity of $t \mapsto N(p, t)$ is necessary and sufficient for the inequality (3-1) to hold provided that N solves the backwards heat equation. Indeed, let $r(a) = N(p + a, t + a^2)$. Then

$$\begin{aligned}
r'(a) &= N_p + 2aN_t, \\
r''(a) &= N_{pp} + 4aN_{pt} + 2N_t + 4a^2 N_{tt} \overset{\text{(3-2)}}{=} 4aN_{pt} + 4a^2 N_{tt}, \\
r'''(a) &= 4N_{pt} + 4aN_{ppt} + 8a^2 N_{ptt} + 8aN_{tt} + 4a^2 N_{ttp} + 8a^3 N_{ttt} \overset{\text{(3-2)}}{=} 4N_{pt} + 12a^2 N_{ptt} + 8a^3 N_{ttt}, \\
r''''(a) &= 4N_{ptp} + 8aN_{ptt} + 24aN_{ptt} + 12a^2 N_{pttp} + 24a^3 N_{pttt} + 24a^2 N_{ttt} + 8a^3 N_{tttp} + 16a^4 N_{tttt} \\
&\overset{\text{(3-2)}}{=} 4N_{ppt} + 32aN_{ptt} + 32a^3 N_{pttt} + 16a^4 N_{tttt}.
\end{aligned}$$

By Taylor's formula we have

$$\begin{aligned}
N(p + a, t + a^2) + N(p - a, t + a^2) &= r(a) + r(-a) = 2r(0) + r''(0)a^2 + r''''(0)\frac{a^4}{12} + o(a^4) \\
&= 2N(p, t) + N_{ppt}(p, t)\frac{a^4}{3} + o(a^4) \\
&\overset{\text{(3-2)}}{=} 2N(p, t) - N_{tt}(p, t)\frac{2a^4}{3} + o(a^4).
\end{aligned}$$

Thus it follows from (3-2) that

$$\lim_{a \to 0} \frac{N(p + a, t + a^2) + N(p - a, t + a^2) - 2N(p, t)}{a^4} = -\tfrac{2}{3} N_{tt}$$

is nonnegative, i.e., $t \mapsto N(p, t)$ is concave.

Now we are ready to complete the proof of Theorem 1.3. Let $N \geq 0$ be such that $\xi_N = \xi_{N+1} = \cdots = \xi$. We have

$$\begin{aligned}\mathbb{E} N(\xi, [\xi]) &= \mathbb{E} N(\xi_N, [\xi_N]) \\ &= \mathbb{E}(\mathbb{E}(N(\xi_0 + (\xi_1 - \xi_0) + \cdots + (\xi_N - \xi_{N-1}), (\xi_1 - \xi_0)^2 + \cdots + (\xi_N - \xi_{N-1})^2) \mid \mathcal{F}_{N-1})).\end{aligned}$$

Notice that the random variables

$$\eta = \xi_0 + (\xi_1 - \xi_0) + \cdots + (\xi_{N-1} - \xi_{N-2}) \quad \text{and} \quad \zeta = (\xi_1 - \xi_0)^2 + \cdots + (\xi_{N-1} - \xi_{N-2})^2$$

are $\mathcal{F}_{N-1}$ measurable. Yet on each atom Q of $\mathcal{F}_{N-1}$ the random variable $\xi_{N-1} - \xi_N$ takes values $\pm A$ with equal probabilities $\frac{1}{2}|Q|$. Then it follows from (3-1) that

$$\mathbb{E} N(\xi_N, [\xi_N]) \geq \mathbb{E} N(\xi_{N-1}, [\xi_{N-1}]).$$

Iterating this inequality and using the boundary value $N(p, 0) = \log p$ for $p > 0$, we obtain

$$\mathbb{E} N(\xi, [\xi]) \geq \mathbb{E} N(\xi_0, 0) = \ln \mathbb{E}\xi.$$

This finishes the proof of Theorem 1.3. □

Inequality (1-15) follows from the following lemma.

Lemma 3.2. *We have that*

$$\log(1 + y^{-2}) \geq \int_y^\infty \int_x^\infty e^{(-t^2+x^2)/2} t^{-2}\, dt\, dx \geq \tfrac{1}{3} \log(1 + y^{-2}) \tag{3-5}$$

for all $y > 0$.

Proof. For a positive constant $C > 0$, consider a map

$$h(y; C) = \int_y^\infty \int_x^\infty e^{(-t^2+x^2)/2} t^{-2}\, dt\, dx - C \log(1 + y^{-2}), \quad y > 0.$$

Notice that $h(\infty; C) = 0$. To prove the second inequality in (3-5) (or the first inequality in (3-5)) it suffices to show that

$$h_y(y; C) = -\int_y^\infty \frac{e^{(-t^2+y^2)/2}}{t^2}\, dt + 2C \frac{1}{y^3 + y} \leq 0 \tag{3-6}$$

for $C = \frac{1}{3}$ (or $h_y(y; C) \geq 0$ for $C = 1$). Next, consider

$$\psi(y; C) = e^{-y^2/2} h'(y; C) = -\int_y^\infty e^{-t^2/2} t^{-2}\, dt + 2C \frac{e^{-y^2/2}}{y^3 + y}.$$

Clearly $\psi(\infty; C) = 0$. To show (3-6) for $C = \frac{1}{3}$ (or its reverse inequality when $C = 1$), it suffices to verify that

$$\psi_y(y; C) = \frac{Ce^{-y^2/2}}{y^2} \left(\frac{1}{C} - 2 \frac{(3y^2 + 1)}{(y^2 + 1)^2} - \frac{y^2}{y^2 + 1} \right) = \frac{Ce^{-y^2/2}}{y^2} \left(\frac{1}{C} - 1 + \frac{4}{(y^2 + 1)^2} - \frac{5}{y^2 + 1} \right) \geq 0$$

for $C = \frac{1}{3}$ (or the reverse inequality for $C = 1$). Let $s = (y^2+1)^{-1} \in [0, 1]$. Then $-1 + 4t^2 - 5t$ is minimized on $[0, 1]$ when $t = \frac{5}{8}$ and its minimal value is $-\frac{41}{16}$ (or maximized on $[0, 1]$ when $t = 0$ and its maximal value is -1). The lemma is proved. □

4. Concluding remarks

One may ask how we guessed $N(p, t)$ which played an essential role in the proof of Theorem 1.3. There is a general argument [Ivanisvili et al. 2018] which informally says that *any estimate in Gauss space* (*or more generally on the hamming cube*) *involving f and its gradient has a corresponding* ***dual estimate*** *for a stopped Brownian motion and its quadratic variation* (*or more generally dyadic square function*). For example, to prove inequality (1-2), there was a certain function $M(x, y)$ used in the proof. This function satisfies the Monge–Ampere type PDE

$$\det\begin{pmatrix} M_{xx}+\frac{M_y}{y} & M_{xy} \\ M_{xy} & M_{yy}\end{pmatrix} = 0 \tag{4-1}$$

with a boundary condition $M(x, 0) = \log(x)$, so that the matrix in (4-1) is positive definite. Suppose we would like to solve the PDE (4-1) in general. Using exterior differential systems (see details in [Ivanisvili and Volberg 2018]), the PDE may be linearized to the backwards heat equation; namely, locally the solutions can be parametrized as

$$\begin{cases} M(x, y) = -px + \sqrt{t}\,y + u(p, t), \\ x = -u_p(p, t), \\ y = 2\sqrt{t}\,u_t(p, t), \end{cases}$$

where u satisfies the backwards heat equation

$$\begin{cases} u_t + \frac{1}{2}u_{pp} = 0, \\ u(M_x(x, 0), 0) = M(x, 0) - xM_x(x, 0), \end{cases}$$

with $t \geq 0$ and $p \in \Omega \subset \mathbb{R}$. An important observation is that if u happens to satisfy

$$u(p+a, t+a^2) + u(p-a, t+a^2) \geq 2u(p, t),$$

then under some additional assumptions on u, one expects an identity

$$M(x, y) = \sup_t \inf_p\{-px + \sqrt{t}\,y + u(p, t)\} = \inf_p \sup_t\{-px + \sqrt{t}\,y + u(p, t)\},$$

which, if true, implies that M satisfies the 4-point inequality (2-8); see [Ivanisvili et al. 2017; 2018] for more details. These functions $M(x, y)$ and $u(p, t)$ we call dual to each other. One may verify that for our particular M defined by (1-11), the corresponding dual $u(p, t)$ is

$$u(p, t) = 1 + \log(-p) + \int_{-p/\sqrt{t}}^{\infty}\int_s^{\infty} r^{-2}e^{(-r^2+s^2)/2}\,dr\,ds, \quad p < 0, \quad t \geq 0,$$

which coincides with $N(p, t)$ after subtracting 1 and reflecting in the variable p.

Using this approach, one could try to prove the 4-point inequality (2-8), which would imply

$$\mathbb{E}M(g, |Dg|) \geq M(\mathbb{E}g, 0) \quad \text{for all } g : \{-1, 1\}^n \to \mathbb{R}_+.$$

So, one may hope to obtain Theorem 1.1 on the hamming cube after substituting $g = e^f$. However, we did not proceed with this path on the unfortunate grounds that the chain rule misbehaves on the hamming cube, i.e., the identity $|De^f|/e^f = |Df|$ does not hold. Therefore, to prove (1-2) on the hamming cube perhaps different ideas are needed.

Our last remark is that one may provide another proof of (1-15) using a simpler function compared to N (what we call *the supersolution*). Indeed, consider

$$N^{\sup}(p, t) = \tfrac{1}{2}\log(p^2 + t), \quad t \geq 0, \quad p > 0.$$

Notice that

$$N^{\sup}(p, 0) = \log(p), \tag{4-2}$$

$$\frac{N^{\sup}_{pp}}{2} + N^{\sup}_t = \frac{t + t^2}{2(p^2 + t)^2} \geq 0, \tag{4-3}$$

$$N^{\sup}_{tt} = -\frac{1}{2}\frac{1}{(t + p^2)^2} \leq 0. \tag{4-4}$$

Using the same argument as in the proof of (1-14) we verify that $N^{\sup}(p, t)$ satisfies (3-1). Notice that $N^{\sup}$ does not solve the backwards heat equation; however, due to inequality (4-3) the stochastic process Y_s constructed in the proof of (3-1) will be a submartingale which is sufficient for the proof of (3-1). Thus we obtain

$$\log \mathbb{E}\xi_\infty - \mathbb{E}\log \xi_\infty \leq \frac{1}{2}\mathbb{E}\log\left(1 + \frac{[\xi_\infty]}{\xi_\infty^2}\right),$$

which improves on (1-15) by a factor of $\frac{1}{2}$.

The supersolution $N^{\sup}(p, t)$ was guessed from the form of the inequality (1-15) by considering

$$\log(p) + C\log(1 + t/p^2)$$

and choosing an optimal constant C (in our case $C = \frac{1}{2}$ worked well). It was a good coincidence that such an $N^{\sup}$ satisfies (4-2), (4-3), and (4-4). However, if one tries to construct a supersolution to the inequality (1-2) one may hope that, by chance,

$$M(x, y) = \log(x) + Ce^{y^2/(2x^2)}(1 + y/x)^{-1}$$

may work for some positive C. A direct calculation shows that there is no positive constant C such that inequality (2-11) holds.

Acknowledgments

We would like to thank the referees for helpful suggestions and remarks. The authors also thank Sergey Bobkov and Changfeng Gui for helpful discussions and comments. Ivanisvili was supported in part by NSF grants DMS-2152346 and CAREER-DMS-2152401.

References

[Bobkov and Götze 1999] S. G. Bobkov and F. Götze, "Exponential integrability and transportation cost related to logarithmic Sobolev inequalities", *J. Funct. Anal.* **163**:1 (1999), 1–28. MR Zbl

[Caffarelli 2000] L. A. Caffarelli, "Monotonicity properties of optimal transportation and the FKG and related inequalities", *Comm. Math. Phys.* **214**:3 (2000), 547–563. MR Zbl

[Efraim and Lust-Piquard 2008] L. B. Efraim and F. Lust-Piquard, "Poincaré type inequalities on the discrete cube and in the CAR algebra", *Probab. Theory Related Fields* **141**:3-4 (2008), 569–602. MR Zbl

[Hytönen et al. 2016] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach spaces, I*: *Martingales and Littlewood–Paley theory*, Ergebnisse der Mathematik (3) **63**, Springer, 2016. MR Zbl

[Ivanisvili and Volberg 2018] P. Ivanisvili and A. Volberg, "Isoperimetric functional inequalities via the maximum principle: the exterior differential systems approach", pp. 281–305 in 50 *years with Hardy spaces*, edited by A. Baranov and N. Nikolski, Oper. Theory Adv. Appl. **261**, Birkhäuser, Cham, 2018. MR Zbl

[Ivanisvili and Volberg 2020] P. Ivanisvili and A. Volberg, "Poincaré inequality $3/2$ on the Hamming cube", *Rev. Mat. Iberoam.* **36**:1 (2020), 79–97. MR Zbl

[Ivanisvili et al. 2017] P. Ivanisvili, F. Nazarov, and A. Volberg, "Hamming cube and martingales", *C. R. Math. Acad. Sci. Paris* **355**:10 (2017), 1072–1076. MR Zbl

[Ivanisvili et al. 2018] P. Ivanisvili, F. Nazarov, and A. Volberg, "Square functions and the Hamming cube: duality", *Discrete Anal.* **2018** (2018), art. id. 1. MR Zbl

[Moser 1971] J. Moser, "A sharp form of an inequality by N. Trudinger", *Indiana Univ. Math. J.* **20** (1971), 1077–1092. MR Zbl

[Onofri 1982] E. Onofri, "On the positivity of the effective action in a theory of random surfaces", *Comm. Math. Phys.* **86**:3 (1982), 321–326. MR Zbl

[Stolyarov et al. 2022] D. Stolyarov, V. Vasyunin, P. Zatitskiy, and I. Zlotnikov, "Sharp moment estimates for martingales with uniformly bounded square functions", *Math. Z.* **302**:1 (2022), 181–217. MR Zbl

Received 19 May 2021. Revised 25 Nov 2021. Accepted 20 Dec 2021.

PAATA IVANISVILI: pivanisv@uci.edu
Department of Mathematics, University of California, Irvine, Irvine, CA, United States

RYAN RUSSELL: rrussell14@g.ucla.edu
Current address: *Irvine, CA, United States*
Department of Mathematics and Statistics, California State University, Long Beach, CA, United States

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ANALYSIS & PDE

Volume 16 No. 5 2023

Nonuniform stability of damped contraction semigroups 1089
RALPH CHILL, LASSI PAUNONEN, DAVID SEIFERT, REINHARD STAHN and YURI TOMILOV

Long time solutions for quasilinear Hamiltonian perturbations of Schrödinger and Klein–Gordon equations on tori 1133
ROBERTO FEOLA, BENOÎT GRÉBERT and FELICE IANDOLI

An extension problem, trace Hardy and Hardy's inequalities for the Ornstein–Uhlenbeck operator 1205
PRITAM GANGULY, RAMESH MANNA and SUNDARAM THANGAVELU

On the well-posedness problem for the derivative nonlinear Schrödinger equation 1245
ROWAN KILLIP, MARIA NTEKOUME and MONICA VIŞAN

Exponential integrability in Gauss space 1271
PAATA IVANISVILI and RYAN RUSSELL